

BRIEF

Students' Collective Memory to Recall An Examination

Adam M. Persky, PhD,^{a,b} Kathryn A. Fuller, PharmD^a

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

^b Associate Editor, *American Journal of Pharmaceutical Education*, Arlington, Virginia

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

Submitted March 15, 2021; accepted July 6, 2021; ePublished July 2021

Objective. To validate a faculty observation and 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. For accuracy and extent of reproduction, the results were compared to the original examination version.

Results. The students were able to collectively recreate 90% of questions on the two examinations. For most questions (51%), students also recreated the question in combination with a correct response and at least one incorrect response. A large majority of the questions recreated were considered 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 allows them to recreate almost an entire completed examination from memory. Based on the findings of the study and tenets of social psychology, we discuss the potential implications for examination security, providing feedback 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 completed examinations to students, faculty generally fall into three categories: those that release the exams, those that do not release the exams but offer examination reviews/provide feedback, and those that do not release the exams and provide no review. Often, students want to see their completed examinations 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 examinations, it seems natural they would want to know the results of this effort.^{4,5} As educators, faculty should provide feedback as evidence suggests feedback is an essential part of the learning process.⁶⁻⁸ However, faculty often protect their examinations and do not release them due to the resources required to construct quality examinations. Faculty believe that by not releasing completed examinations they can safely reuse questions and ensure future exam security and accuracy of evaluating learning. Often, faculty may see students discussing answers to examination questions after completing examinations or students may use their study groups to recreate the exam. Inherent curiosity in combination with an investment in learning inclines students to seek answers. The question then arises whether completed examinations not released to students for feedback are truly safe from reproduction.

In theory, a group of students should be able to recreate a shared experience (ie, an examination) due to *collective memory*. The social psychology literature is rich with examples of collective memory, which is the phenomenon that groups of people should better recall information than any one individual.⁹⁻¹² Brown et al 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 approximately five advertisements. If this information is translated to examination questions and each student can recall 5 questions, then it would take approximately 20 students to recall a 100-question examination, assuming no redundancies. Additionally, due to *collaborative facilitation*, recall of one individual item may cue the recall of other items, allowing students to remember more items as they discuss the examination questions.¹²

As mentioned earlier, students compare their responses after an examination which serves two purposes: to determine their "correctness" and to communicate. We communicate to remember and share experiences. Harber et al conducted a study where 33 students completed a field trip to the morgue.¹³ In 3 days, over 800 people had heard about

the field trip. The spread of the news of the field trip was attributed to the compulsion to disclose traumatic experiences; that is, people tend to share emotional events with others with little prompting. Examinations may also serve as an emotional event and sharing those details, is a natural human response.¹⁴ Thus, if students can recall information from an examination, they may naturally want to share details.

To date, there is negligible research on the ability of students to recreate assessments. Anderson and McDaniel examined students' memory for a quiz.¹⁵ Here, students were asked to recall questions they remembered, and the authors concluded that students had a poor memory of the quiz. While this indicates that individual students did not have a good memory of the quiz, it does not conclude or predict their collective memory of the entire examination. The objective of this study is to validate a faculty observation that students can recreate a recently completed examination from memory. By answering this question, faculty may alter their views about assessment and releasing examinations and therefore, focus more on using assessment to promote learning rather 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-terms with each exam assessing new content and a cumulative final. The three midterms included 40 to 70 single answer, multiple choice questions ranging in Bloom's Taxonomy from lower level (knowledge/comprehension) to middle level (application/analysis).¹⁶ Each question had 1 correct response and 2 to 3 distractors. Examinations were administered using ExamSoft. Because of the COVID-19 pandemic, examinations were virtually proctored to minimize academic dishonesty.¹⁷ Immediately after the completion of the second and third midterm, the instructor emailed the students a link to a single GoogleDoc with instructions to recreate the examination. For midterm 2, students were instructed to be as accurate as possible with content, order, and phrasing; for midterm 3, the instructions were modified to be more explicit and include the correct and incorrect answer choices. Students were given 48 hours after each examination and offered extra credit, in the form of one extra percentage point towards the final examination, for attempting to complete the activity. In previous years of the course, examinations were not released to students as they only received their examination score and a report of performance on question categories. That is, they did not have access to examination questions from previous years. In the current year, the same procedures were used: students only received their score and performance on question categories.

The primary outcome measures were the number of questions recreated, the number of questions recreated where the correct answer was identified, the number of questions recreated where at least one distractor was correctly identified, and the ordinal scoring of question specificity. For this latter metric, recreated questions that captured the general 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 more specific or exact phrasing of the target question, parts of the case vignette or other details were considered high specificity. In the analysis, we included the examination question item difficulty and discrimination index to assess if question difficulty impacted students' ability to recreate the examination. Based on the data distribution, we tested for differences between generated versus non-generated items using a Mann-Whitney U-test. Statistical significance was set at $p < .05$. This was a retrospective study exempted by the University of North Carolina Institutional Review Board.

RESULTS

Within 48 hours after two midterm examinations, the class recreated 90% of the questions and the respective correct answer (103/115, 90% for midterm 2 and 89% for midterm 3). Of these questions, students recreated on average 51% of the questions with the correct answer in combination with at least one distractor. However, the percentage of questions with distractors recreated varied between the two examinations with 31% recreated on the second midterm and 82% recreated on the third midterm. Of the 51% of questions with the correct answer in combination with at least one distractor, 10% had 1 distractor, 12% had 2 distractors, and 29% had all 3 distractors recreated correctly. Cumulatively, 8% of questions recreated were deemed low specificity meaning students captured the very minimal of general content or gist of the question without specifics; 38% were medium specificity as they captured the stem of the question and the remaining 54% were deemed high specificity based on further question details. Figure 1 depicts an example for low and high specificity questions.

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

DISCUSSION

The goal of this study was to assess the collective memory of students for an examination and therefore determine if students can recreate a recently completed examination from memory. The results demonstrate that soon after completing an examination, students can recall a very large portion.

The students were not able to recreate the entire examination. Interestingly, the 10% of questions that were not recreated seem to be less difficult or challenging questions. There are several explanations for this phenomenon: encoding variability, attentional failure; or recollection failure. Generally, recall of information is enhanced when there are multiple routes to the target information rather than a singular route.^{18,19} Encoding variability theory suggests that the more cues a person has to recall information, the better the recall.¹⁸ 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 those contextual details are lacking, in part by increased familiarity. As an example, if a question was discussed in class it may be familiar and “old” and old items in memory may be harder to recall.^{20, 21} In this study, the examination was recalled as a collaborative group, asynchronously. If a group can divide a memory task amongst themselves, this distributes the burden of remembering as each person uses another 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 could serve a cue for another student and promote recall of other facts and this would be consistent with the idea of *collaborative facilitation*.⁹ The reverse could also be true, and it is noted that collaborative brainstorming and recall often lead to inhibition known as *collaborative inhibition*.⁹

We did note some differences between the two examinations which may reflect the specificity of instructions. For the third midterm, more specific instructions were provided to students asking them to recreate question distractors. This did result in an increased number of distractors listed (32% vs 81% of questions). Therefore, the instructions provided to students to recreate their collective memory may be influential. This is consistent with the literature on eyewitness testimony in that how questions are asked lead to varying degrees of answer specificity.²²⁻²⁴

One strength of this study includes that it was conducted in an authentic classroom environment. Additionally, students were unaware they were going to be asked to reproduce the examinations and therefore did not prepare or plan ahead to memorize these questions. Since these were first year, first semester students, they were unfamiliar with the instructors and their question writing style. Finally, previous years examinations were not released and there were no practice examinations, 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 overall ability to recreate these questions may be underestimated. An additional and final limitation, may be that students with higher academic ability, such as those in a professional program, may have a higher ability to retain information and thus can recreate examinations more easily compared to those with lower academic abilities.

An argument may be made whether recalling an examination falls into academic dishonesty or the unauthorized use or attempted use of material, information, notes, study aids, devices or communication in the learning environment.²⁵ However, students recalling from memory an experience does not fall into academic dishonesty. The real question is what students do with the recalled information. Some might argue if they generate an examination and share it with future classes, this may constitute academic dishonesty. However, if a student generated an exam 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.²⁶⁻²⁹ Additionally, sharing of experiences should deepen learning based on the theory of *shared in-group attention*.³⁰ 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.³⁰ Thus, the sharing of examination experiences should promote learning.

If we assume students can reproduce examinations from memory and their natural tendency is to share information, which also may facilitate learning, there are a few recommendations. One recommendation is that the academic culture should reduce the high stakes nature of examinations or minimize the transactive nature of education (ie, students should reproduce what faculty teach). In doing so, we can minimize the student inclination to cheat and, hopefully, promote learning.^{31,32} In the learning process, we need to remember feedback is important. As such, we recommend learning and feedback should 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 restudy them, there is minimal need

for feedback. However, if quizzes are not returned, corrective feedback is necessary.¹⁵ Thus, simply returning students' assessments would reduce the need and time for corrective feedback. If assessments are to be used for summative judgements of learning which 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.³³ However, when high stakes examinations are necessary (eg, licensure), then best practices for examination development should be followed.³⁴

CONCLUSION

By understanding this collective memory, instructors may be challenged in their views in releasing completed examinations to students or providing feedback under the guise of exam security. The results of the study should not be surprising as 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, Burianek 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. Harber KD, Cohen DJ. The Emotional Broadcaster Theory of Social Sharing. *J Lang Soc Psychol.* 2005;24(4):382-400.
14. 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.* 2021/02/01/ 2021;73:101706.
15. 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.
16. 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.
17. 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.
18. Postman L, Knecht K. Encoding variability and retention. *Higher Educ* 1983;22(2):133-52.
19. Nairne JS. The myth of the encoding-retrieval match. *Memory.* 2002;10(5-6):389-95.
20. 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.
21. 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.
22. Ginet 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.
23. Thorley C. Enhancing individual and collaborative eyewitness memory with category clustering recall. *Memory.* 2018;26(8):1128-39.

24. Verkampt F, Ginet 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.
25. McCabe DL, Trevino LK, Butterfield KD. Cheating in Academic Institutions: A Decade of Research. *Ethic Behav*. 2001;11(3):219-32.
26. 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.
27. Loda T, Erschens R, Loenneker H, et al. Cognitive and social congruence in peer-assisted learning - A scoping review. *PloS one*. 2019;14(9):e0222224-e.
28. 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.
29. Williams B, Reddy P. Does peer-assisted learning improve academic performance? A scoping review. *Nurse Educ Today*. 2016;42:23-9.
30. 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.
31. McCabe D. Cheating: Why Students Do It and How We Can Help Them Stop. *Am Educator*. 2001;25(4):38.
32. 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.
33. Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. Twelve Tips for programmatic assessment. *Med Teach*. 2015;37(7):641-6.
34. Council NR. *High Stakes: Testing for Tracking, Promotion, and Graduation*. . Washington, DC: The National Academies Press.; 1999.

AJPE
Accepted Draft

Table 1. Summary of Question Specificity on Recreation and the Question Psychometrics.

Specificity	Count (%)	Mean Question Difficulty (SD)	Mean Question Discrimination Index (SD)
Low	8 (7%)	.82 (.22)	.20 (.11)
Medium	39 (35%)	.78 (.22)	.21 (.16)
High	55 (49%)	.81 (.21)	.18 (.14)
Not recreated	11 (10%)	.91 (.08)	.13 (.09)

Data summarized as mean and standard deviation.

Accepted Draft
 AJPE

Figure 1. Examples of Low, Medium, and High Specificity Questions That Were Recreated

Category	Recreated Question	Original
Low	<p>Anxiety and worry neural circuits</p> <ul style="list-style-type: none"> • CSTC loop^a 	<p>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?</p> <ul style="list-style-type: none"> • Amygdala-centered circuit • Mesocortical dopamine pathway • Nigrostriatal dopamine pathway • Cortico-striato thalamo-cortical (CSTC) circuit
Medium	<p>What lab test result would you find in someone with CKD?</p> <ul style="list-style-type: none"> • Low phosphate • Low calcium^a • High 1,25 OH vitamin D • Low PTH 	<p>In CKD patients with mineral bone disorders, which of the following would you expect to see when looking at the lab values below?</p> <ul style="list-style-type: none"> • low blood calcium • low blood parathyroid • high blood 1,25 diOH vitamin D • low blood phosphate
High	<p>A Patient with a PMH of HTN, afib, hypotension, goutwhat type of thromboembolism is at risk and why?</p> <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 	<p>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...</p> <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

^adenotes the correct response