

How Much Content Knowledge Do Graduating Medical Students Retain from Their Pre-Clinical and Clinical Courses?

Kenneth D. Royal¹ · Mari-Wells Hedgpeth¹ · Debra Bynum² · Cristin Colford²

Published online: 9 December 2015

© International Association of Medical Science Educators 2015

Abstract Because medical students are exposed to so much material during their 4 years of medical school, many have questioned how much knowledge one might reasonably retain. Further, physicians and medical educators have long been particularly critical of basic science knowledge retention and have asserted that a significant amount of basic science information learned is lost by the time students graduate. The purpose of this study was to investigate knowledge retention and decaying effects by utilizing a quasi-experimental design that compared responses from a sample of exiting fourth year medical students to their previously rendered responses to the exact same items 2 to 3 years prior. Results indicated exiting fourth-year medical students performed about equally well on both pre-clinical and clinical items. Additionally, it was found that 53 % of responses were deemed stable, correct knowledge over time as indicated by students providing the same correct response. Conversely, approximately 31 % of responses went from right to wrong when assessed again, indicating students likely forgot prior information or guessed correctly the first time.

Keywords Medical education · Knowledge retention · Knowledge decay · Quasi-experimental research · Measurement · Psychometrics · Repeat items

Knowledge retention has been a long-standing problem in medical education [1]. Because medical students are exposed to so much material during their 4 years of medical school, many have questioned how much knowledge one might reasonably retain. Due to its pre-clinical focus, many have been particularly critical of basic science instruction and have long asserted that a significant amount of basic science information is lost by the time students graduate [2–7].

Previous research on discipline-specific courses (e.g., anatomy, physiology, biochemistry, etc.) has often presented conflicting findings. Although most published studies indicate knowledge decay does occur over time [8, 9], some studies have revealed surprisingly high rates (>90 %) of knowledge retention [10–13], whereas other studies have reported that most students would fail the course if tested again [5, 14, 15].

To our knowledge, no research has rigorously investigated holistic medical knowledge retention as measured by exiting fourth-year students, nor has any studies involved a quasi-experimental design that compares students' current responses to previous responses on the same items captured 2 to 3 years prior. Thus, the purpose of this study was twofold. First, we sought to determine how well exiting fourth-year medical students could perform on an examination consisting of randomly selected items measuring basic science and clinical science content previously covered in the medical curriculum. Second, we sought to determine the extent to which fourth-year medical students' responses remain the same or differed when presented the same items 2 to 3 years after they were originally answered, as this information would be valuable for discerning the stability of content knowledge over time. It was hypothesized that because students had received basic science instruction during the first year (2010–2011) and very limited exposure thereafter, but significantly more, and more recent, instruction and training in the clinical sciences (2011–2014) that students would likely perform better on clinical items and less well on basic science items.

✉ Kenneth D. Royal
kdroyal2@ncsu.edu

¹ North Carolina State University, Raleigh, NC, USA

² University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Methods

Institutional Context

At the time this study was performed, the medical school curriculum at our university was based on a first year that focused on the basic sciences and a second year that focused on clinical sciences. During the first academic year, the curriculum was comprised of four blocks or courses (Molecules to Cells, Structure and Development, Integrative Function, and Host Defense) each consisting of 8 weeks of instruction. This curriculum focused on the basic sciences of microbiology, anatomy, physiology, and immunology.

During the second program year, the curriculum consisted of one course with the objective of introducing pre-clinical students to the tools used for diagnosis and treatment (radiology and pathology), and nine organ-based courses or blocks that focused on abnormal findings, and were more clinical in nature: Hematology-Oncology, Pathophysiology of the Cardiovascular System, Respiratory, Gastrointestinal, Renal-Urinary, Brain and Behavior, Endocrinology, Reproductive-Genetics, and Musculoskeletal. Each of these organ-based blocks consisted of 2 to 6 weeks of instruction.

Study Design

This study utilized a quasi-experimental design. In an effort to determine how well fourth-year medical students could perform on an examination consisting of items measuring basic science and clinical science content previously covered in the medical curriculum, two primary factors were necessary. First, items needed to be selected at random. Second, the randomly selected items needed to possess similar psychometric characteristics as the other collective items from the larger pool from which they were drawn.

Because we wanted to determine the extent to which fourth-year medical students' responses remained the same or differed when presented the same items at two different points in time, it was necessary to recruit an appropriate sample of students who could provide two sets of responses to a series of common items. We opted to target fourth-year students primarily for two reasons. First, they were approaching graduation, so the idea of measuring students' knowledge at the time of departure from medical school was appealing for multiple reasons (e.g., assessment purposes, understanding the "forgetting curve", etc.). Second, a significant amount of time had lapsed since the students were initially exposed to each of the items, thus providing plenty of time for students to be exposed to additional instruction, particularly clinical instruction, which could have a significant impact on their responses when tested at a second point in time. It was

hypothesized that because students had received basic science instruction during the first year (2010–2011) and very limited exposure thereafter, but significantly more and more recent instruction and training in the clinical sciences (2011–2014) that students would likely perform better on clinical items and less well on basic science items.

Data Collection and Sample Frame

To recruit students to participate in the study, an email message was sent to all of the fourth-year students ($n=161$) participating in the required capstone course. Participants were informed that if they agreed they would be re-tested on examination items that they had encountered during their first 2 years of the program. In exchange for participation, students were awarded 2 h of credit toward the required capstone course, and they were eligible to partake of pre-test study test breakfast of doughnuts. Another incentive was that students who completed the test were provided with an individualized score report.

A total of 36 fourth-year medical students (about 22.22 % of the fourth-year class) comprised the sample frame for this study. With regard to demographic characteristics, the average age was 28.82 ($SD=3.39$), with 22 (61.10 %) females and 14 (38.90 %) males. With regard to race, 28 (77.78 %) students were White, 4 (11.11 %) were Black, and 4 (11.11 %) were classified as Other.

It should be noted that recruitment of students who were enrolled in the first two program years from 2010–2012 was difficult as many students elect to pursue a dual degree (MPH, MBA, etc.) or complete a research year between their second and third years. For the purpose of this study, we specifically wanted participants who completed the first-year curriculum in 2010–2011 and the second-year curriculum in 2011–2012 with no breaks. The recruitment efforts were aimed at the fourth-year class knowing that some of those students did not complete the curriculum in that order due to having decelerated or taken a leave of absence or were returning from pursuing a dual degree. However, if those students were interested in participating, they were allowed to take the examination, but their results were excluded from these analyses.

Thus, when the demographic characteristics of the students in this study was compared to the population of students in the 2014 graduating class, the sample characteristics were somewhat disproportionate with regard to gender and race, but neither difference was deemed statistically significantly different. In particular, the graduating class of 2014 consisted of 81 (50.31 %) females and 80 (49.69 %) males, with 101 (63.34 %) reported as White, 18 (11.11 %) as Black, and 42 (25.93 %) as Other. A chi-squared test indicated the sample was not statistically significantly different from the population

of fourth-year students with regard to gender, $X^2=1.37$, $df=1$, $p=0.241$, or race, $X^2=3.84$, $df=2$, $p=0.146$.

Examination and Item Selection

Examination construction began by selecting items from two primary pools of items. The first item pool consisted of 1029 basic science items available during the 2010–2011 year. The second item pool consisted of 1192 clinical science items available during the 2011–2012 year. A number was assigned to each item in each item pool and then randomly sorted. A systematic sampling procedure was then employed in which every n^{th} item was selected to create a sample of 30 items from each of the first- and second-year item pools. This process resulted in an examination consisting of 60 total items. The selection of 30 items per course year was intentional as

psychometric research has indicated examinations consisting of as few as 25 items are likely to yield robust measures provided the items are of sufficient psychometric quality [16, 17]. Further, we recognized that exceeding 60 items would increase the likelihood of fatigue for examinees, thus potentially resulting in careless responses on latter items given there were no stakes associated with this examination.

All items appearing on the fourth-year student examination were randomly selected from the larger population of items administered during the 2010–2011 year. The blueprints of first-year and second-year items are presented in Tables 1 and 2. To ensure the item samples were representative of the population, the difficulty and discrimination values were compared for items in both the sample and the population (see Table 3).

Table 1 Blueprint of basic science items included on examination

Item #	Course	Content area
1	Histology	Urinary system
2	Histology	Urinary system
3	Histology	Endocrine system
4	Host Defense and Microbial Pathogens	Major histocompatibility complex
5	Host Defense and Microbial Pathogens	Antigen-antibody
6	Host Defense and Microbial Pathogens	Viral replication
7	Host Defense and Microbial Pathogens	Respiratory bacteria
8	Host Defense and Microbial Pathogens	Innate immunity
9	Host Defense and Microbial Pathogens	Immunity and cancer
10	Integrative Function and Cellular Basis	Cognition
11	Integrative Function and Cellular Basis	Gastric physiology
12	Molecules to Cells	Hyperlipidemia
13	Molecules to Cells	Cell to Cell Interaction
14	Molecules to Cells	Oncogenes and Tumor Suppression
15	Molecules to Cells	Ketone bodies
16	Molecules to Cells	Dose-response relationships (GP)
17	Molecules to Cells	Glycoproteins, glycosaminoglycans, glycolipids
18	Molecules to Cells	Dose-response relationships (GP)
19	Molecules to Cells	Drug elimination
20	Physiology	Integrated: sodium balance, hypertension
21	Physiology	Acid-base balance
22	Physiology	Integrated: intestinal electrolytes, intestinal physiology/pathology
23	Physiology	Renal vasculature/vasa recta
24	Physiology	Gastric physiology/pathophysiology
25	Structure and Development	Head: anatomy
26	Structure and Development	Abdomen: embryology
27	Structure and Development	Head: anatomy
28	Structure and Development	Head: embryology
29	Structure and Development	Abdomen: radiology
30	Structure and Development	Pelvis: anatomy

Table 2 Blueprint of clinical science items included on examination

Item #	Course	Content area
1	Endocrine System and Nutrition	Diabetes/pancreas
2	Endocrine System and Nutrition	Pituitary
3	Endocrine System and Nutrition	Salt and hypertension
4	Gastrointestinal System	Hepatitis
5	Gastrointestinal System	Abdominal pain
6	Gastrointestinal System	Gastric neoplasms
7	Hematology/Oncology	Hemolytic anemia
8	Hematology/Oncology	Platelets
9	Hematology/Oncology	Hemolytic anemia
10	Musculoskeletal System	Vascular disorders of bone
11	Musculoskeletal System	Tumors and tumorlike conditions
12	Musculoskeletal System	MSK practical exam questions
13	Musculoskeletal System	Biochemistry and physiologic processes of bone
14	Nervous System	Otolaryngology
15	Nervous System	Clinical Neurology
16	Nervous System	Otolaryngology
17	Nervous System	Psychiatry
18	Pathology	Breast pathology
19	Pathophysiology of Heart Disease	Valvular heart disease
20	Pathophysiology of Heart Disease	Heart failure
21	Reproductive Medicine	Contraception
22	Reproductive Medicine	Normal menstrual cycle
23	Respiratory System	Palliative Care
24	Tools for Diagnosis and Therapy	Radiology: tools for diagnosis and therapy
25	Urinary System	Glomerular disease
26	Urinary System	Acid-base
27	Urinary System	Glomerular disease
28	Urinary System	Kidney stones
29	Urinary System	Acid-base
30	Urinary System	Glomerular disease

An independent samples *t* test revealed each sample of items did not statistically differ for basic science items with regard to difficulty, $t(30)=0.00$, $p=1.000$ and discrimination

$t(30)=-0.39$, $p=0.702$ with alpha set at 0.05. Clinical science items also revealed no statistically significant differences with regard to difficulty, $t(30)=0.00$, $p=1.000$ and discrimination $t(30)=-0.83$, $p=0.413$.

Table 3 Characteristics of sample items and population items

Item subset	Difficulty	Discrimination
Basic science population ($n=1029$)	M=0.83, SD=0.16	M=0.18, SD=0.13
Basic science sample ($n=30$)	M=0.83, SD=0.17	M=0.19, SD=0.14
Clinical science population ($n=1192$)	M=0.85, SD=0.16	M=0.16, SD=0.13
Clinical science sample ($n=30$)	M=0.85, SD=0.15	M=0.18, SD=0.13

Administration

In an effort to minimize the sources of error potentially stemming from the administration of the examination, participants completed the examination under the same conditions from previous years. Participants were allowed 60 min to complete the examination (about 1 min per item), which is comparable to the average length of first- and second-year final examinations at this institution. Approval to conduct this study was obtained by the institution's Institutional Review Board (IRB).

Data Analysis

Data were initially analyzed according to group level performance across various item subsets. Data were then analyzed by comparing each student’s response to the same item at two different points in time. Basic science items were administered 3 years prior, and clinical science items were administered 2 years prior. A schema for discerning results was used to aid interpretation of results. SPSS statistical software was used to perform various statistical analyses.

Results

Group Level Performance

Students’ performance on the examination containing repeat items from previous years is reported in Table 4.

An independent samples *t* test comparing basic science and clinical science scores indicate students’ performance was not statistically significantly different on the two content areas: $t(69)=-0.59, p=0.556$ with alpha set at 0.05. Results indicate students performed comparably on basic and clinical science content when administered a randomly selected group of items previously completed during their first 2 years of medical school.

Students’ Performance: Time 1 Versus Time 2

We compared individual student responses from their initial attempt at the item (Time 1) to their repeat attempt (Time 2) years later. The testing software program used retains students’ responses for every item they have ever completed on any of our examinations. This permitted us to pull each student’s response to each individual item (regardless of where it occurred in the curriculum). The intention of this comparison was to gain insights as to how well students might be reasonably expected to retain content knowledge over time. We believe five different scenarios were possible with regard to students’ performance across two points in time. These scenarios are presented in Table 5.

Table 6 presents a breakdown of students’ performance according to the scenarios presented in Table 5.

Table 4 Students’ performance (percent correct) during 4th year

Item subset	Mean (%)	SD (%)	Min. (%)	Max. (%)	Median (%)
Overall	59.46	7.82	40.68	81.67	59.13
Basic science	58.76	10.16	37.93	80.00	60.00
Clinical science	60.14	9.65	31.03	83.33	60.00

Table 5 Possible responses and explanations

Scenario	Potential explanation
Right to right	Stable correct content knowledge
Right to wrong	Either forgot information, or guessed the first time
Wrong to right	Either learned information, or guessed better second time
Wrong to same wrong	Stable incorrect knowledge, or guessed the same
Wrong to different wrong	Likely guessed both times

A comparison of student responses provided at Time 1 and Time 2 indicated most responses were stable, correct answers (52.56 %). Interestingly, a significant number of responses also indicate students may have either forgotten information, or guessed better the first time the item was presented to them (30.95 %) as their response went from right to wrong. About 6 % of responses indicate students may have learned or guessed better since the first attempt, about 5 % of responses suggest students have stable incorrect knowledge about some content, and about 6 % of responses likely indicate students never learned the content at all.

Discussion

We hypothesized students would perform much better on the clinical items due to recency effects and significantly more training in this area. We were particularly surprised to see that students performed about equally well on basic and clinical science items. Given the response change schema utilized in Table 5, it was disconcerting to know that approximately 31 % of responses went from right to wrong when measured years later. We believe this finding is important for two reasons: (1) it serves as a useful approximation of the “forgetting curve” and (2) it might provide clues regarding the degree to which measurement error was present during initial measurements.

A significant limitation of this study, as well as most any study that compares students’ scores collected in a similar manner, is the potential for increased measurement error

Table 6 Results as measured from Time 1 to Time 2

Scenario	Count	Percent (%)
Right to right	1211	52.56
Right to wrong	713	30.95
Wrong to right	137	5.95
Wrong to same wrong	116	5.03
Wrong to different wrong	127	5.51

stemming from initial measurements. Given the students likely knew what general content to expect on their examinations, had received direct instruction pertaining to that content (instructional sensitivity) [18, 19], and had devoted specific study in preparation, it is entirely plausible that initial measures of performance were overestimated due to instructional sensitivity, recall, and study effects. This potential score contamination could invalidate the baseline measure and distort any genuine frame of reference for understanding knowledge retention and decay.

With regard to curricular affairs, many medical schools have been revising their curricula in recent years to create integrated learning opportunities for the basic and clinical sciences. Given this study's finding that students tend to recall both basic and clinical science material equally well, it seems adding clinical content earlier in the curriculum could help anchor basic science material so that is perceived as more relevant to students. Similarly, it seems adding basic science content later into the curriculum, where relevant, could also help with basic science content retention. Future research should replicate this study on medical schools with different curricula to discern knowledge retention and decaying effects.

Conclusion

In conclusion, our quasi-experimental study revealed medical students typically retain knowledge obtained from basic and clinical science courses equally well. Overall, students were able to answer nearly 60 % of the same items correct when administered 2 to 3 years later. About 53 % of responses were deemed stable, correct knowledge, but approximately 31 % of responses went from right to wrong when assessed again. We believe these findings are particularly useful for establishing reasonable expectations for knowledge retention obtained from classroom instruction. Future research should identify the extent to which various instructional strategies may improve knowledge retention over time.

Acknowledgments The authors would like to thank Drs. Kurt Gilliland, Edward Kernick, and Kenya McNeil-Trice for their logistical assistance in making this study possible, and Ms. Julia Xin for computer programming assistance that made data acquisition possible.

Compliance with Ethical Standards Approval to conduct this study was obtained by the institution's Institutional Review Board (IRB).

References

1. Custers EJFM. Long-term retention of basic science knowledge: a review study. *Adv Health Sci Educ.* 2010;15:109–28.
2. Bethe A. Kritische Betrachtungen u"ber den vorklinischen Unterricht. [Critical observations on preclinical education]. *Klin Wochenschr.* 1928;7(31):1481–3.
3. Cole L. What is wrong with the medical curriculum? VIII. *Lancet.* 1932;110(30):253–4.
4. Miller GE, Graser HP, Abrahamson S, Harnack RS, Cohen IS, Land A. *Teaching and learning in medical school.* Cambridge: Harvard University Press; 1961.
5. Blizard PJ, Carmody JJ, Holland RAB. Medical students' retention of knowledge of physics and biochemistry on entry to a course in physiology. *Br J Med Educ.* 1975;9:249–54.
6. Neame RL. The preclinical course of study: help or hindrance? *J Med Educ.* 1984;59(9):699–707.
7. Norman GR. Where we want to be by 2010: the essential role of basic science in medical education: the perspective from psychology. *Clin Invest Med.* 2000;23(1):47–51.
8. Donovan JC, Salzman LF, Allen PZ. Patterns of learning in medical school. *J Med Educ.* 1969;44(7):589–94.
9. Rico E, Galindo J, Marset P. Remembering biochemistry: a study of the patterns of loss of biochemical knowledge in medical students. *Biochem Educ.* 1981;9:100–2.
10. Swanson DB, Case SM, Luecht RM, Dillon GF. Retention of basic science information by fourth-year medical students. *Acad Med.* 1996;71:S80–2.
11. Rodriguez R, Campos-Sepulveda E, Vidrio H, Contreras E, Valenzuela F. Evaluating knowledge retention of third-year medical students taught with an innovative pharmacology program. *Acad Med.* 2002;77(6):574–7.
12. Saffran M, Kennedy WB, Kelley PR. Retention of knowledge of pharmacology by U.S. and Canadian medical students. *Trends Pharmacol Sci.* 1982;3:461–3.
13. Kennedy WB, Kelley PR, Saffran M. Use of NBME examinations to assess retention of basic science knowledge. *J Med Educ.* 1981;56:167–73.
14. D'Eon MF. Knowledge loss of medical students on first year basic science courses at the University of Saskatchewan. *BMC Med Educ.* 2006;6:5.
15. Levine HG, Forman PM. A study of retention of knowledge of neurosciences information. *J Med Educ.* 1973;48:867–9.
16. Wright BD, Stone MH. *Best test design.* Chicago: MESA Laboratory; 1979.
17. Linacre JM. Sample size and item calibration stability. *Rasch Meas Trans.* 1994;7(4):328.
18. Popham JW. Instructional insensitivity of tests: accountability's dire drawback. *Phi Delta Kappan.* 2007;89(2):146–50.
19. Royal KD, Hedgpeth MW, Smith KW, Kirk D. A method for investigating "instructional familiarity" and discerning authentic learning. *Ann Med Health Sci Res.* 2015.