

Long-term understanding of basic science knowledge in senior medical students

Niklas Wilhelmsson¹, Klara Bolander-Laksov¹, Lars O. Dahlgren², Håkan Hult², Gunnar Nilsson³, Sari Ponzer⁴, Lars Smedman⁵, Anna Josephson⁶

¹Centre for Medical Education, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Sweden

²Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden

³Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Sweden

⁴Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Sweden

⁵Department of Women's and Children's Health, Karolinska Institutet, Sweden

⁶Department of Neuroscience, Karolinska Institutet, Sweden

Correspondence: Niklas Wilhelmsson, Centre for Medical Education, Department of Learning, Informatics, Management and Ethics, Berzelius väg 3, Karolinska Institutet, 171 77 Stockholm, Sweden. Email: niklas.wilhelmsson@karolinska.se

Accepted: September 12, 2013

Abstract

Objectives: The purpose of this study is to explore the relation between basic science knowledge and the ability to understand and make use of basic science in explaining a clinical scenario in the final year of medical school.

Methods: A sample of senior medical students was re-assessed using the same test they had taken 3 years earlier. This was followed by an in-depth interview on one of the topics taken from the test. Their respective level of knowledge was compared with their performance in the interview. The test was analysed according to the revised Bloom's Taxonomy, and the interviews carried out according to the phenomenographic approach.

Results: Performance was around 60% (n=19) of the original performance, with no significant correlation between original test and re-test ($r = 0.258$, $p = 0.29$) and large interpersonal variation. A high performance in the original test did not predict a good performance; rather, the

reverse seemed probable. None of the students who achieved high grades in the original test displayed a stable long term understanding that was measured in the interview. The test comprised questions of a generally low taxonomical level, but could not explain the mismatch between test-result and level of understanding.

Conclusions: Findings suggest substantial loss of basic science knowledge during medical training. Attention should be directed to designing examinations that are purposeful, when it comes to what kind of knowledge is desirable in medical graduates as well as how that knowledge should be acquired. Further larger-scale research is needed to corroborate these findings.

Keywords: Basic medical education, clinical reasoning, medical students' learning, basic science, transfer of knowledge

Introduction

Medical students' use of basic science knowledge in clinical settings has long been of interest. Studies have mostly focused on basic science knowledge¹ and its usefulness in clinical diagnosis,² less attention has been given to the transformation of acquired knowledge over time. As emphasized almost 80 years ago, it is critical to consider the longitudinal development of knowledge over time in order to avoid "disuse atrophy", disintegration due to inactivity, especially in lengthy educational programmes such as the

medical one.³ Knowledge of basic science has been of interest to medical educators since Miller found disappointingly low scores on a delayed test in anatomy, biochemistry and physiology among senior medical students, regardless of the students' initial scores.⁴ When the same test was administered to clinical faculty, the results were even more devastating. In subsequent studies, the magnitude of the loss ranges from levels not significantly different from random guessing,⁵ to more moderate decreases⁶ and finally

to almost no loss at all.⁷ Interestingly, the degree of knowledge decline does not seem to be related to the initial score.^{6,7}

The level of basic science knowledge used in clinical diagnosis differs depending on the discipline and mirrors to some extent the content of clinical clerk-ships: biochemistry and anatomy appear more susceptible to substantial decay; physiology and pathology are more resistant.^{8,9,10,11} In analogy, subject areas more frequently used during clerk-ships, e.g. pharmacology, are even more sustainable and have been shown to improve.⁷ Still, the implications of this are open to discussion, as answering examination questions is one thing, transforming knowledge into a clinical situation is another, and the former might have a shorter survival rate than the latter.¹²

While biochemistry and anatomy seem to suffer from severe long-term knowledge decay, physiology has proven to be particularly hard to understand from the very start.^{13,14,15,16} Misconceptions are common, even in central aspects of body regulation such as cardiac function,¹⁵ respiratory function¹⁴ and blood-pressure regulation.¹⁶ This has been attributed to the complex nature of the discipline, including causal reasoning, mathematics and far-reaching integration.¹⁷ In studying physiology, students often seem to rely more on situational descriptions than on underlying causality.¹⁶ Suggestions to focus educational efforts on general physiological principles have been put forward in order to facilitate students' understanding.^{18,19} It is assumed that by comprehending a limited number of principles, students will recognize patterns in body functions and similarities between organ systems.

This study was set up as an attempt to explore the relation between the level of basic science knowledge and the ability to understand and make use of basic science in explaining a clinical scenario measured with qualitative, phenomenographic interviews. Mixed methods²⁰ are used to explore possible links between results obtained from quantitative and qualitative perspectives. Physiology was chosen for its integrative features and for its ability to link basic science with clinical application.

Method

Participants

Ethical permission was applied for to the local board of ethics at Karolinska Institutet and to the Stockholm City Council regional Ethics Committee. Both bodies ruled that the study should proceed as there was no risk involved to the participating subjects.

Out of a year cohort of 120, 19 randomly chosen medical students sat a re-examination on basic science in their last year of study. They were invited to an "educational test", but had no knowledge of what the test was going to include, nor were they given any opportunity to prepare.

Study Design

At the time of the study, the medical programme consisted of a preclinical phase of 2 years comprising the basic sciences with minor auscultation sessions in a clinical setting, followed by clinical rotations for the remaining 3.5 years. Physiology was taught at the beginning of the second year for 16 weeks. The course work included lectures, laboratory exercises and seminars on clinical cases. At the end of that year, students took an integrated written examination on the basic sciences. This examination marks the closing of the pre-clinical phase of the curriculum and the start of the clinical rotations. It contained questions regarding all the subjects that had been studied: anatomy, cell-biology, biochemistry, physiology and neuroscience. It was divided into 4 themes, each with 12-14 short essay-type questions.

Data collection

The re-examination consisted of 2 parts, out of four in the original examination, which comprised 27 questions. Each part comprised a theme; a medical case with attached questions. All questions were rated, according to identical assessment-standards as used in 2003, by the same senior professor of physiology who had marked their initial examinations. Their results from this re-examination were then compared to their results from the original examination, which they had completed 36 months earlier. The comparison was based on the result on all questions in the two themes, as the questions were interrelated and would render a more substantial view of the student's level of knowledge.

Interviews were carried out with nine of the participating students, focusing on their understanding of the following scenario: "Consider a person riding an exercise bicycle. Why is the onset of fatigue so sudden, when the work load is increased linearly?" This question was copied from the pre-clinical examination the students had taken. It was chosen as it required deeper-level explanation and allowed for elaboration within a number of organ-systems. The interviews sought to clarify conceptions the subjects held about physiological fatigue, and how different aspects they focused on contributed to their understanding. The interviews were carried out by experienced senior consultants with research-experience in varying disciplines of the basic sciences. Pilot-interviews and seminars on interview-technique were performed prior to the start to ensure inter-interviewer reliability.

Data analysis

Phenomenographic analysis was performed on the transcribed interviews.²¹ This approach seeks to reveal possible qualitative differences in ways of understanding and different conceptions held about a phenomenon. The specific aspects of the phenomenon – in this case, physiological fatigue – that are being focused on are then related in

categories of description and ranked depending on completeness and qualitative level of understanding. The categories together form the outcome space of the phenomenon and are described by referring to the content of the phenomenon. The categories are thus descriptive, non-personal, together exhaustive, hierarchically related to each other and logically related to the phenomenon of physiological fatigue.

In order to clarify the comparison between the two ways of measuring the knowledge maturation – quantitative remembrance and qualitative understanding of a clinical scenario – an analysis of the examination was also performed according to the Revised Bloom's Taxonomy.²² To avoid a misleading comparison it seemed important to consider what the examination actually measured. The qualitative analysis was conducted to ensure validity,²³ interviews were performed by experienced attending clinicians and both educational and medical experts took part in the analysis. The interviews were open ended, lasting until the topic was exhausted, which enabled all information the student forwarded to be elaborated upon.

Results

Re-examination

The scores on the re-examination depict a substantial knowledge decline over a period of three years (Figure 1). The average score presented as a percentage of the initial score (second year examination) was 60%, and 39% if presented as a percentage of total score (last year examination). The level of retention varied substantially, between 28% and 105% (one student improved). No significant correlation between the score on original examination and re-examination was found, Pearson $r = 0.258$, $p = 0.29$, 95% CI $[-0.223, 0.637]$; Spearman rank correlation $r(s) = 0.213$, $p = 0.38$, 95% CI $[-0.267, 0.608]$. Of the 19 students that participated, only one would have passed the re-examination. It is especially noteworthy that all three top students, that performed the highest results in the original examination, lost more examination scores in the re-examination, compared to the other 16 students. The group investigated was representative of the whole population of second year medical students who took the original test; the original test's mean score was 66.1% for the whole year cohort, and 65.4% in the sample. Of the nineteen students that took the re-examination, nine random students were interviewed; one student displayed a high performance (more than 60% of the total score) in that test, two students a low performance (less than 40% of the total score), and five students were around the mean of the whole population.

Analysis

The students' explanations of the scenario evolved into

three categories of description depending on which aspects were focused on.²⁴ Category A focuses on causal relations between biochemical mechanisms and physiological responses, thus linking organ-systems together. Students assigned to this category are able to view the situation as a displaced equilibrium, discerning important thresholds in the chain of reactions that give rise to the displacement. Category B emphasizes the large number of factors involved; whilst exhibiting a horizontal juxtaposition of sections of physiological knowledge, it generally lacks clear cause and effect relations. Most relevant organ-systems are accounted for but the links between them are more associative than causal. As opposed to the notion of equilibrium in category A, this category is characterized by the idea of a finite amount of substrate, i.e. a limited supply of combustible chemical compounds to yield energy. Category C is an inversion of the former two since its main characteristic is fragmentation. The students tend to fill their gaps of knowledge with explanations from clinical medicine, which do not capture core elements of the situation. The answers given here revealed a poor understanding comprising isolated sections of physiological knowledge. Lactate production and some of its effects were correctly accounted for, but connecting sections of knowledge to explain an overall state of anaerobic metabolism was not achieved. The categories are hierarchically ordered, meaning that B and C are both subsets of A, and C a subset of B (Figure 2).

When the scores for each student were compared within each category, no general pattern was found, only large interpersonal variation. Only one student (student 9) displays a positive remembrance and is found in category A. Results of the re-examination in category B varied from 31% to 70% and in category C from 28% to 68%.

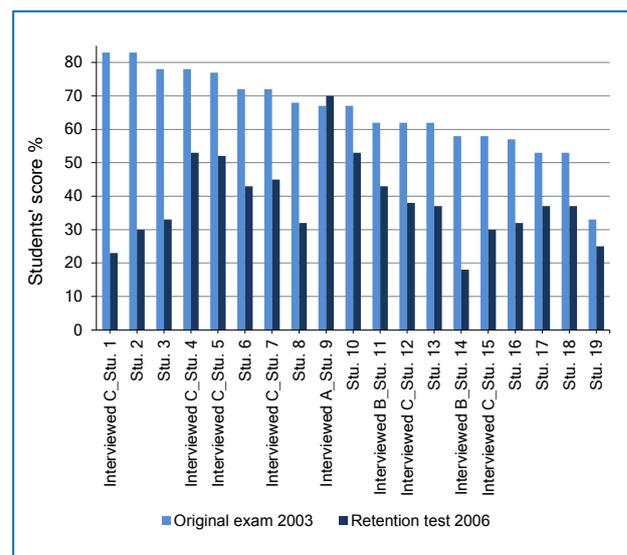


Figure 1. Students' (n=19) scores on original examination (second year of study) and re-examination (final year of study). The 9 interviewed students were categorised into A, B or C depending on their answer in the qualitative analysis.

Analysis of the examination

Some of the 27 questions contained more than one question. All learning tasks were analyzed by the use of Bloom's Revised Taxonomy and referred to one field in the two-dimensional outcome space. The majority ($n = 25$) of the questions asked belonged to the first cognitive dimension of this taxonomy ("Remember"), followed by the second ("Understand"), which contained 19 questions. In higher cognitive dimensions ("Apply", "Analyze", "Evaluate" and "Create") only 2 questions were found.

Discussion

Our results suggest that, over an interval of three years, a good initial achievement does not immunize against loss of knowledge, corroborating more large-scale studies.^{4,9} In our sample, the relative decline was even greater among (initially) top performing students. Neither did high initial marks seem to predict a sustainable long-term understanding. Considering that all four students with the highest scores in the original test offered explanations falling into category C, there may even be a negative correlation between long-term understanding and achievement in examination, a worrying finding also reported in other academic areas. The average level of basic science knowledge in our study was found in the lower interval compared with earlier findings^{5,7,9,10,11} presumably due to re-organization of this type of knowledge during clinical training. To reveal the students' understanding, the phenomenographic analysis focused on the different conceptions the students held of physiological fatigue. One approach was found more successful in terms of elaborated explanation (category A), because of the focus on the notions of equilibrium and threshold.

The analysis of the examination demonstrates an emphasis on recalling and reproducing, followed by tasks directed at interpreting, comparing and explaining. The types of knowledge demanded are mainly classifications, principles and models; specific facts are also frequently asked for. More questions related to a lower cognitive level such as "remembering" than "understanding". With this in mind, one might argue that the students' understanding was not adequately tested in the original examination, and in that sense any re-examination test is of very limited value. If the examination directs the students' course of study,²⁵ this might result in a sub-optimal approach to the learning tasks, leading them astray.

The validity of course examinations has been questioned because of their inability to provide opportunities for correction and feed-back,²⁶ and final examinations do not correlate with either clinical performance or previous results on course assessment.²⁷ In view of our findings it is therefore worth asking in what way basic science knowledge is represented in clinical medicine and how this should be measured. Putting too much trust in written examinations when it comes to long-term understanding clearly seems a perilous way to go. And in terms of a constructive align-

ment of intended learning outcomes and assessment procedures,²⁸ the examination does not seem to be of very much help either.

According to encapsulation theory,²⁹ knowledge is embedded in a network that changes as the student becomes more experienced and new knowledge is incorporated. When basic science knowledge is used in a clinical case it is used in an encapsulated mode comprising higher-level concepts. By its nature, encapsulated knowledge is used holistically as individual items have gradually formed an associative net from which the higher-level concepts derive. It is therefore essential that assessments target knowledge in this holistic and transformed way.

It has been suggested that emphasizing core principles¹⁹ and general models¹⁸ could promote understanding and transfer between physiological subsystems, although in a clinical setting, problem specific knowledge proved more important than generic knowledge in diagnostic performance.¹ If understanding is based on knowledge of general principles rather than on detailed factual knowledge, there is reason to ask what kind of understanding is desirable in a medical graduate. Understanding a discipline built on causality like physiology is clearly not the same cognitive task as understanding anatomy, where there is no causality to understand.³⁰

The comparably small number of participants constitutes an important limitation in the study, and naturally hampers generalization of the results. Nevertheless, it was necessary to include diverse levels of student performance in the sample for methodological reasons. It was deemed beyond reach to randomize enough subjects to achieve statistical power, and this way of conduct does not comply with phenomenographic methodology. On the other hand, the use of complementary methods revealed aspects of understanding in relation to remembrance that would not otherwise have been examined had just a single method been employed.

In the present study, delayed assessment of basic science and clinical understanding were employed as two contrasting ways of measuring knowledge maturation. While the test was centered around questions on a low taxonomical level, the interview was solely focused on deeper level processing, i.e. to realize connections and structures in the physiological situation at hand. The delayed re-test generated an average level of recall around 60% and the phenomenographic analysis yielded a delayed understanding that can be generally classified, with a few exceptions, as poor. High performance in the original test did not predict good remembrance; in fact, the reverse seemed probable. Furthermore, none of the students who achieved high grades in the original test displayed a stable long-term understanding. In view of these outcomes, the focus on detailed factual knowledge in the test does not seem to favour long-term understanding. Examining students on a higher taxonomical level might result in a more stable long-term under-

standing. If the misconceptions were already there during the first test, this was not revealed in the original assessment. Remembrance of knowledge seems to be a sub-optimal approach since the knowledge has been transformed in a process of encapsulation rather than retained. These findings highlight important issues concerning which examinations to use in medical training, and what these examinations actually measure.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Coderre S, Jenkins D, Mclaughlin K. Qualitative differences in knowledge structure are associated with diagnostic performance in medical students. *Advances in Physiology Education*. 2009;14:677-684.
2. Norman G. Research in clinical reasoning: past history and current trends. *Medical Education*. 2005;39:418-427.
3. Cole L. What is wrong with the medical curriculum? *The Lancet*. 1932;220(5683):253-254.
4. Miller GE. An inquiry into medical teaching. *Journal of Medical Education*. 1962;37(3):185-191.
5. Sinclair DC. An experiment in the teaching of anatomy. *Journal of Medical Education*. 1965;40:401-413.
6. Rico E, Galindo J, Marset P. Remembering biochemistry: a study of the patterns of loss of biochemical knowledge in medical students. *Biochemistry Education*. 1981;9:100-102.
7. Kennedy WB, Kelley PR, Saffran M. Use of NBME examinations to assess retention of basic science knowledge. *Journal of Medical Education*. 1981;56:167-173.
8. Mateen FJ, D'Eon M. Neuroanatomy: a single institution study of knowledge loss. *Medical Teacher*. 2008;30:537-539.
9. D'Eon M. Knowledge loss of medical students on first year basic science courses at the University of Saskatchewan. *BMC Medical Education*. 2006;6(5) DOI:10.1186/1472-6920-6-5.
10. Ling Y, Swanson DB, Holtzman K, Bucak SD. Retention of basic science information by senior medical students. *Academic Medicine*. 2008; 83(Suppl):82-85.
11. Swanson DB, Case SM, Leucht RM, Dillon GF. Retention of basic science information by fourth-year medical students. *Academic Medicine*. 1996;71(Suppl):80-82.
12. Watt ME. Retention of preclinical knowledge by clinical students. *Medical Education*. 1987;21:119-124.
13. Michael JA. Students' misconceptions about perceived physiological responses. *Advances in Physiology Education*. 1998;19:90-98.
14. Michael JA, Richardson D, Rovick A, Modell H, Bruce D, Horwitz B, et al. Undergraduate students' misconceptions about respiratory physiology. *Advances in Physiology Education*. 1999;22:127-135.
15. Michael JA, Wenderoth MP, Modell HI, Cliff W, Horwitz B, McHale P, et al. Undergraduates' understanding of cardiovascular phenomena. *Advances in Physiology Education*. 2002;26:72-84.
16. Fyrenius A, Silén C, Wirell S. Students' conceptions of underlying principles in medical physiology: an interview study of medical students' understanding in a PBL curriculum. *Advances in Physiology Education*. 2007;31:364-369.
17. Michael J. What makes physiology hard for students to learn? Results of a faculty survey. *Advances in Physiology Education*. 2007;31:34-40.
18. Modell H. How to help students understand physiology? Emphasize general models. *Advances in Physiology Education*. 2000;23:101-107.
19. Michael J, Modell H, McFarland J, Cliff W. The "core principles" of physiology: what should students understand? *Advances in Physiology Education*. 2009;33:10-16.
20. Schifferdecker KI, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. *Medical Education*. 2009;43:637-644.
21. Marton F. Phenomenography – describing conceptions of the world around us. *Instructional Science*. 1981;10:177-200.
22. Andersson LW, Krathwohl DR. Taxonomy for learning, teaching and assessing: a revision of bloom's educational objectives. Boston: Allyn & Bacon; 2001.
23. Giacomini MK, Cook DJ. Users' guide to the medical literature XXIII. Qualitative research in health care A. Are the results valid? *Journal of the American Medical Association*. 2008;284(3):357-362.
24. Wilhelmsson N, Dahlgren L-O, Hult H, Wirell S, Ledin T, Josephson A. A phenomenographic study of basic science understanding – senior medical students' conceptions of physiological fatigue. 2013, Unpublished paper.
25. Ramsden P. Learning to teach in higher education. London: Routledge; 1992.
26. van der Vleuten. Validity of final examinations in undergraduate medical training. *British Medical Journal*. 2000;321: 1217-1219.
27. McManus IC, Richards P, Winder BC, Sproston KA. Clinical experience, performance in final examinations, and learning style in medical students: prospective study. *British Medical Journal*. 1998;316: 345-350.
28. Biggs J, Tang C. Teaching for quality learning at university. Buckingham: SRHE & Open University Press; 2007.
29. Schmidt HG, Boshuizen HPA. On the origin of intermediate effects in clinical case recall. *Memory & Cognition*. 1993;21(3):338-351.
30. Wilhelmsson N, Dahlgren LO, Hult H, Josephson A. On the understanding of anatomy. *Studies in Higher Education*. 2011;2(36):153-165.