Learning and retention of quantum concepts with different teaching methods

Louis Deslauriers

Department of Physics & Astronomy, and Carl Wieman Science Education Initiative, University of British Columbia,
Vancouver, British Columbia, Canada V6T 1Z3

Carl Wieman

Department of Physics & Astronomy, and Carl Wieman Science Education Initiative, University of British Columbia,
Vancouver, British Columbia, Canada V6T 1Z3,
and Department of Physics, University of Colorado, Boulder, Colorado 80309-0390, USA
(Received 20 May 2010; published 31 January 2011)

We measured mastery and retention of conceptual understanding of quantum mechanics in a modern physics course. This was studied for two equivalent cohorts of students taught with different pedagogical approaches using the Quantum Mechanics Conceptual Survey. We measured the impact of pedagogical approach both on the original conceptual learning and on long-term retention. The cohort of students who had a very highly rated traditional lecturer scored 19% lower than the equivalent cohort that was taught using interactive engagement methods. However, the amount of retention was very high for both cohorts, showing only a few percent decrease in scores when retested 6 and 18 months after completion of the course and with no exposure to the material in the interim period. This high level of retention is in striking contrast to the retention measured for more factual learning from university courses and argues for the value of emphasizing conceptual learning.

PACS numbers: 01.40.−d

I. INTRODUCTION

What is important in education is not what students know the day of the final exam, but rather what learning they retain and can apply months and years later. However, nearly all assessments of learning, whether carried out in educational research or as part of the regular educational activities in a course, measure performance during or at completion of a course, rather than long-term performance. Research on memory and retention suggests that many standard educational practices, such as noncumulative exams and great emphasis on final exams that encourages studying by cramming, are likely to lead to enhanced short-term performance at the expense of poor long-term retention [1]. The data on retention of learning from university courses vary considerably between different studies, likely as a result of difficulties in having well-controlled studies given the variety of complicating variables. However, the preponderance of evidence is that the majority of the factual information learned in a single course is lost within a year if there are no further opportunities for relearning or review, with most of that forgetting occurring within the first three months [2,3]. This matches the common belief among faculty and students alike that much of the knowledge acquired in a university physics course (or most courses for that matter) is soon forgotten after the final exam.

Here we carried out a relatively well-controlled study looking at students' retention of concepts in quantum mechanics (QM) up to 1.5 years later, when taught by two very different pedagogical approaches. We show that the teaching methods used have a substantial impact on the initial mastery, but retention of these concepts remains nearly perfect over a long period, independent of teaching methods.

There have been a few studies [4–8], most quite recent, suggesting that conceptual knowledge in physics may be retained better than the primarily factual knowledge measured in retention studies such as Ref. [3]. In Ref. [3] it was seen that the majority of the material was forgotten after only three months. However, the data on retention of conceptual learning in physics and similar subjects are very limited. It is difficult to do longitudinal studies in general, and in this case there are the further confounding variables of possible exposure to the material during the retention study interval and the difficulties in distinguishing conceptual and factual knowledge.

An early study by Francis *et al.* [4] showed little drop in scores on the Force Concept Inventory after some years for a self-selected set of volunteers. It is unknown how much the volunteers may have reencountered the material during the retention interval, and the absence of uncertainties or standard deviations for the sample populations in the publication prevent quantitative conclusions. Shaffer and McDermott report somewhat contradictory results on retention when testing students' basic understanding of electric circuits. They show that the pedagogical approach clearly impacts initial learning, but for one test group the scores dropped by approximately 50% relative to the base-

line when tested in the next school quarter [5], whereas in a second test group there was zero decrease after a year interval [6]. For the latter case, however, there is no discussion given as to what other possible exposure to the material the students may have had in the intervening year. Recently Pollock [7] showed that scores on the Brief Electricity and Magnetism Assessment (BEMA) test that measures understanding of basic concepts of electricity and magnetism (E&M) dropped $5 \pm 2\%$ over a period of two years and that the delayed BEMA scores reflected differences in pedagogical approach encountered two years earlier. However, he cautions that students likely will have had several occasional encounters with E&M in courses in the intervening period. Kolhmyer *et al.* obtained somewhat similar results to Pollock, again using the BEMA test, although with rather larger ($\sim 25\%$) drops (more forgetting) over time [8]. However, it is difficult to make a direct comparison, because a large range of retention intervals are aggregated in the data of Ref. [8], and there is no indication of relearning opportunities that might have been encountered during these intervals. Also, the student populations are different, being physics majors in [7] and a broader undergraduate population in [8]. In summary, these results are suggestive that introductory physics knowledge measured by conceptual tests such as the Force Concept Inventory (FCI) and BEMA is retained relatively well, but it is clear that it would be useful to have further studies.

The ideal retention study is to have a measure of student performance immediately following the learning period, and then subsequent measures for those same students at later times without any further exposure to the material during the retention interval. We came fairly close to that in this study; we measured students' performance on the Quantum Mechanics Concept Survey (QMCS) [9] when they completed a second-year modern physics course and then measured the same students using the QMCS again at periods of 6 and 18 months later. They had no exposure to the material in any courses during this retention interval and when polled could not think of anything they had encountered in the intervening time that would have helped them answer these questions. In this study it would have been preferable to have the two pedagogical treatments measured for retention after the same time interval, but that was balanced against other practical realities. Using two different retention intervals made it possible to retest a large fraction of both of the original sample groups over substantial retention intervals before they were "contaminated" by further exposure to the material covered on the QMCS.

There are some unique benefits in using quantum mechanics for studying retention of conceptual learning. First, it is a challenging conceptual subject and so is much more representative of the learning we value for long-term educational outputs than are the necessarily simpler memory tasks used in cognitive psychology studies of retention, such as remembering novel words from a foreign language. Second, there is a common belief that the concepts of QM are so challenging that one cannot grasp them immediately, and they must be "digested" over a large fraction of a physics student's education for them to be understood. This belief implies that conceptual mastery of QM would be quite unstable after only one semester of exposure and thus would be expected to change substantially with time after completion of an introductory course. Testing this hypothesis is one of our research questions. A third benefit to studying learning of QM is that a student will very seldom encounter the subject outside of a physics course,[1] so by knowing the schedule of courses that students take during the retention interval in the study, it is possible to greatly reduce the likelihood that students will have "relearning" events where they encounter the material during this interval. Conceptual questions in introductory physics often involve everyday objects such as the motion of a car, circuits with electric light bulbs, or the behavior of magnets. It is not unlikely that encountering these objects could trigger students to reflect on the material and "self-test" their understanding during the retention interval. Such retrieval has been shown to have an important effect on retention [1].

Finally, although there are considerable data showing that different pedagogical approaches produce different results on physics tests of conceptual mastery of introductory mechanics and electricity, such as the FCI (and similar instruments such as The Force and Motion Conceptual Evaluation, etc.) and BEMA tests, there are much less data on conceptual learning of other subjects. QM is somewhat different from introductory mechanics and E&M, in that students have encountered those latter subjects before college regularly in daily life and at least somewhat in their school courses. They have encountered QM far less and so are likely to have fewer preconceptions about the material. So a hypothesis we have heard a number of faculty express is that, for more advanced material of this sort, students are much more the proverbial "blank slates," so the traditional lecture method is better suited for teaching this material than it is for introductory material. A related hypothesis that is often heard from many non-physics-education-research faculty members is that the interactive engagement techniques that achieve superior results on BEMA and FCI tests work for the nonphysics students who dominate the first-year physics courses, but these techniques are either unnecessary or detrimental for serious physics majors. Here we are able to test both of these hypotheses.

---

[1]Students may encounter references to quantum mechanics in television shows or in reading popular science books, but it is very unlikely that they will encounter material in such sources that would have a bearing on answering questions such as are on the QMCS.

The final advantage to using quantum mechanics in this study is that it gives us a uniquely clean opportunity to measure retention of "conceptual knowledge." The QMCS was carefully designed to be insensitive to the relatively large variations in the particular factual knowledge that is covered in introductory modern physics courses [9]. It tests mastery of very general concepts that will be encountered in every introductory QM course, even though presentation of many other aspects of the subject may vary. For this study we used the 12-question version of the QMCS that includes those 12 questions that have been thoroughly tested and shown to be consistently interpreted as intended by students.

One complication in any study of retention is the type of knowledge being retained, whether it is "factual" or "conceptual" and how and if this distinction impacts retention. A full discussion of these issues is beyond the scope of this paper, particularly as many aspects of this are still subjects of extensive debate in the literature. An examination of the distinctions in types of knowledge and their respective levels of retention will be the subject of our further work. However, for this paper our working definition of conceptual knowledge is knowledge that is connected with and transferred from multiple topics covered in a course, so if students cannot figure out the answer based on what they know about one specific topic, they could get the answer based on knowledge about other topics and/or things learned at other times in a course. A test question that a normal student could answer only by remembering that particular piece of information as presented in a single particular context, we would label as testing retention of "factual knowledge." In the language of transfer, a purely factual question would involve no transfer and could not be answered by using a transfer process. In contrast, we label questions that involve significant transfer, particularly if it is possible to transfer from multiple contexts, as testing conceptual knowledge. We believe that all the QMCS questions fall under this definition of conceptual. Our preliminary indications support the utility of this definition in that we see that knowledge we would characterize as conceptual by these criteria is retained much better than knowledge we characterize as factual.

II. CONTEXT OF THE STUDY

The study was carried out using two consecutive years of students in the engineering physics program at the University of British Columbia (UBC). There are about 60 students in each year's class of engineering physics students, and the admissions standards and their backgrounds are essentially identical from one year to the next. The UBC engineering physics curriculum is highly structured so each year's class takes nearly all of the same courses as a cohort, and, prior to this study, those courses had remained quite static in both the material covered and the teaching methods used. The UBC engineering physics program is arguably the most selective and demanding program at UBC, and the engineering physics students generally score above the regular UBC physics majors and honors students when they take the same courses. UBC in turn is among the most selective institutions in Canada, and although there are no measures like SAT scores that can be used for comparison with U.S. institutions, in terms of position in graduating high school class, UBC science students are comparable to those at the most elite U.S. public institutions. Nearly all UBC students graduate from British Columbia high schools, and in the most recent PISA [10] international comparison, BC high school students scored higher on the PISA science test than did students from any country except Finland.

These students were administered the QMCS at the completion of the Introduction to Modern Physics course for engineering physics majors. Engineering physics students take this course during the summer at the end of their second year. This course covers standard topics in quantum mechanics and special relativity with more emphasis on the quantum mechanics portion of the course. The stated primary goal of this course is to introduce students to the basic concepts of quantum mechanics. The first cohort in this study (C1) took the course in the summer of 2008 and had an enrollment of 57 students. The second cohort (C2) was in the summer of 2009 and had an enrollment of 67 students, 4 of whom were not in the engineering physics program and whose data are not included in the retention comparison.

The general structure of the course was the same in both years. The class met twice a week for 1 hour and 50 minutes and once per week for 1.5 hours, for a period of 11 weeks. There were two in-class midterm exams and a final exam, as well as weekly homework assignments. The homework assignments for C1 were typically four to five substantial word problems that encompassed both conceptual ideas and quantitative calculations, whereas for C2 the problems tended to be more broken up into a number of smaller pieces that would have a specific conceptual or quantitative focus, but the total amount of time required to do the homework was similar for both groups. Both cohorts had similar weekly optional problem-solving sessions where students would gather in a room with teaching assistants to work on homework problems. The material covered was quite similar in the two years, except C2 covered some additional material, primarily on applications of quantum mechanics. Standard widely used but different textbooks were used for both courses. The lecture notes were posted online for C2 but were not for C1.

C1 was taught using a traditional lecture approach by an instructor who is widely recognized by both the faculty and students alike as being a superb lecturer and was a recent recipient of the annual award given to the most outstanding teachers at the university. The lecturing style was very well received by the students, resulting in outstanding student evaluations.

The in-class pedagogical approach for C2 was quite different. There were preclass reading assignments with quizzes on the reading, and class time was highly interactive and largely taken up with clicker questions with peer discussion and small group activities such as completing worksheets or concept maps, with follow up "minilectures." It was not unusual to have 30 student questions or comments in a 1 hour and 50 minute class with 1/3 or more of the students making such contributions. Instruction was guided by research on student learning of quantum mechanics. A number of PhET simulations [11] were used in the in-class activities and occasionally in homework assignments. Several articles presenting research on how people learn and how to study for most effective long-term learning were posted for the students to read, although we do not know how many did so.

As we have observed in other newly transformed courses, a few students in C2 initially expressed concern that the course did not involve more traditional lecturing. By the third or fourth week of class, however, most or all of those concerned students told us that they had changed their minds. On the end-of-course survey, the class was overwhelmingly in favor of the teaching methods used in this course, and many of the students have since publically expressed their frustration that *all* their lecture courses were *not* being transformed in this manner. Student evaluations of this course were also exceptionally high.

III. MEASUREMENTS

A. Initial learning

The QMCS was given in class in both courses, one week prior to the final examination. It was presented as an opportunity to start reviewing for the final exam, with the promise that areas where they scored the lowest would be reviewed in the last class. In both courses, students appeared to be working diligently on the exam, and several anecdotally commented that completing the QMCS was a useful exercise in studying for their final exam. Upon completion of the test, students were required to turn in both their answer sheets and the QMCS exam itself. No copies of the QMCS questions were retained by students, and they were never told what they scored.

In keeping with common practice for the QMCS [9], we did not give the students a pretest of the QMCS in these courses. During the development of the QMCS it was given as a pretest, but the students' scores were only a few percent above random guessing, and they indicated it was a demoralizing experience. Random guessing would give a score of 29% on this version of the QMCS. All observations of the two cohorts during the terms they took modern physics supported this assumption that these students had no useful knowledge about quantum mechanics, other than a hazy view of the Bohr model of the hydrogen atom that they had gotten from high school and their first-year chemistry course, but this would have provided them with no help on the QMCS.

In C1, 48 out of the 57 students completed the QMCS. As shown in Table I, the average score was 67 with a standard deviation of $\sigma = 18\%$. In C2, 62 out of the 67 students enrolled completed the QMCS. Their average score was 85 with a standard deviation of $\sigma = 14\%$. This distribution is noticeably distorted from a normal distribution due to the ceiling effect, as a significant fraction of the class scored near or at 100%. The fraction of the students who completed the exam was a measure of attendance at class that day, and generally the weaker students miss more class. Thus, the score for the C1 cohort may be slightly biased upward relative to the C2 score, because a smaller fraction of the C1 cohort completed the exam (84% vs 92.5%). The 67% obtained in the traditional class is well above what has been obtained at the University of Colorado in the counterpart of this course with traditional instruction and comparable to the scores obtained in the transformed pedagogy version of this course at Colorado. The 85% obtained in the transformed UBC course is 16% above the highest score obtained in the transformed version of this course at Colorado and is comparable to the scores students obtain after taking the graduate quantum course.

These results demonstrate that mastery of basic concepts of quantum mechanics by strong physics students is af-

TABLE I. QMCS results for the two classes showing (1) averages for all students who completed the QMCS at the end of their respective course, along with number of students N taking the exam compared to total number enrolled; (2) end-of-course average for that subset for whom we have matched data for both the end of the course and after the corresponding retention interval; (3) and (4) average at the end of the respective 6- and 18-month retention intervals; and (5) "Forget" is the average of the differences for each student between $t = 0$ score and $t = 6$- or 18-month retention score with the standard error in the difference.[a] N is 29 for the top-right four columns and 44 for the bottom-right four columns.

Year	All students $t = 0$	$t = 0$ matched	$t = 6$ months	$t = 18$ months	Forget
08 (C1)	$67 \pm 3\%$ [$N = 48/57$]	$67 \pm 3\%$ [$N = 29$]	—	65%	$2.3 \pm 2.7\%$
09 (C2)	$85 \pm 2\%$ [$N = 62/67$]	$88 \pm 3\%$ [$N = 44$]	85%	—	$3.4 \pm 2.2\%$

[a]Note that this standard error in the difference is smaller than what one would get if one (incorrectly) combined the errors in quadrature of the average values for the initial and retention tests to get the uncertainty in the difference. It is incorrect because the two results are not statistically independent as they involve data from the same students.

fected by the pedagogy used. This is similar to what has been observed with introductory physics material.

B. Retention

There is normally an 18-month gap between the completion of the modern physics course and the second exposure of UBC engineering physics students to quantum mechanics. In order to measure retention of their QM conceptual knowledge, the QMCS was given again early in 2010 in two courses taken largely by the same two cohorts of students. For the C1 cohort, this was an interval of 18 months since they had taken the QMCS, whereas for the C2 cohort it was an interval of 6 months. To test the C1 group, the QMCS was administered in class on the first day of the upper division quantum class, which is taken by many but not all engineering physics students in their fourth year. Students in this course were told that this would serve as a diagnostic of what quantum mechanics they already knew, and so it was important for them to put in an effort to answer as well as they could. Twenty-nine of the original C1 cohort took the exam both times. These 29 students averaged $67.5 \pm 3\%$ ($\sigma = 18\%$) on the QMCS when they took it the first time and hence were highly representative of the original C1 cohort.

The QMCS was administered a second time to much of the C2 cohort during an open class period of an unrelated physics course they were taking. Forty-four students from the C2 cohort took the test both times, with a 6-month interval in between. They were again requested to make their best effort to answer the questions and were told that these results would be used to know how well their courses were being taught and how they might be improved. They were also told that studies on retention of learning show that by testing themselves in this way, it would help preserve their learning of this material. All indications from observing the group was that they also took the exam quite seriously. These 44 students also were a representative sample of the full C2 cohort, as their average initial QMCS score was $88 \pm 3\%$ ($\sigma = 17\%$). The ceiling effect noted in the results of this group when they first took the exam is a concern in a retention study. However, we think it unlikely that it changes the conclusion of near perfect retention for this group. Although the distribution is distorted by the ceiling effect, the averages are still well below 100%, and the standard deviations for this group are essentially the same on the first and second tests.

It should be emphasized that this is an unusually well-controlled measurement of retention in a formal educational setting. These students have a rigid schedule of courses, and the courses they take during this 6–18-month retention period provide no exposure to quantum mechanics. Just after they had completed the QMCS for the second time, we also surveyed all the students in both cohorts, asking them if they had been exposed to anything since the final exam in their previous modern physics course that

would have helped them answer any of the QMCS questions. All students said they had not been exposed to any such material.

There is one small uncontrolled variable in this study, which is that we do not know how much learning students may have done of the material tested by the QMCS between the time it was administered on the next-to-last day of class and when they completed the final exam a week later. It seems unlikely that the amount of learning was significant. These concepts were covered repeatedly in class and on homework throughout the term, and for the final exam they needed to study a great deal of other material involving quantitative problem solving, special relativity (the material covered at the end of the term and not covered on any previous midterm exam), and detailed facts about the experimental basis of quantum mechanics, etc.

The well-known "testing effect" [12], where retention is improved by repeated testing, is likely a negligible effect here. As discussed in [12], the testing effect studies are typically carried out with retention intervals of a week or less, and the impact of a single test improves retention by only several percent. When examined in a classroom setting, a single multiple choice test with feedback showed negligible effect [12]. Tests without feedback, as was the case in this study, have a smaller impact on improving retention.
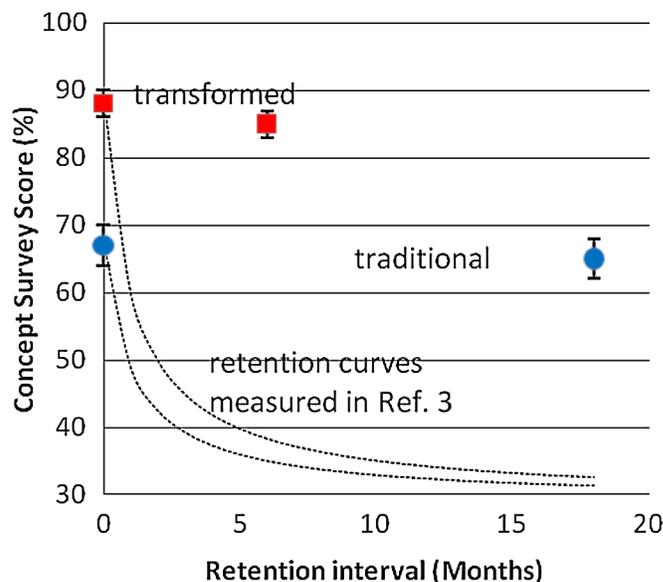
FIG. 1 (color online). Scores on the Quantum Mechanics Concept Survey (QMCS) vs retention interval. Blue circles: traditional lecture retention study group. Red squares: interactive engagement retention study group. The dashed lines are the retention curves for the average student from the university consumer behavior course measured in Ref. [3] for comparison. The curves have been scaled to a baseline of 29%, which is the QMCS score achieved by random guessing.

As shown in Table I and illustrated in Fig. 1, the average decrease for the C1 cohort upon taking the QCMS after an interval of 18 months was only $2.3 \pm 2.7\%$. For the C2 cohort the average score after an interval of 6 months decreased by $3.4 \pm 2.2\%$. Both of these are statistically consistent with zero at the level of $P = 0.1$. For comparison, we plot in Fig. 1 the corresponding retention curves derived from measurements after a university course on consumer behavior, which primarily involved recalling information [3]. These results clearly demonstrate that concepts in quantum mechanics are very strongly retained, far better than most other types of learning that have been measured in university courses. These data also show that the different pedagogical approaches do not result in significant differences in the fraction of this conceptual learning that was retained.

C. Conclusions and future work

This study shows that physics conceptual knowledge is well maintained by students over relatively large segments of their university career, consistent with the results of Pollock [7]. This work also adds further confirmation to the idea that instruction that has the students actively engaged and interacting in the classroom, and provides ongoing formative assessment, results in better conceptual learning than is achieved by even a very highly rated traditional lecturer. This work also extends previous work by demonstrating that these ideas about effective instruction apply both to advanced material for which students have had no previous exposure and to very high level students who are pursuing degrees in physics.

One implication of these and related results is that when shortcomings in the conceptual knowledge of students are encountered in subsequent courses in a physics program, the fault, and hence the solution, likely lies primarily with the teaching and learning in the earlier courses. It is less likely that students originally learned the desired knowledge but then forgot it.

This work has very specifically focused on conceptual mastery. In future work we plan to look at retention of more specific factual knowledge in physics and explore more deeply how one can better distinguish factual and conceptual knowledge. Our hypothesis is that more factual material will be forgotten much more rapidly, and the amount of forgetting will be dependent on the pedagogy used, but this remains to be tested and goes counter to some previous claims [2].

[1] R. A. Bjork, Memory and Metamemory Considerations in the Training of Human Beings, in *Metacognition: Knowing about Knowing*, edited by J. Metcalfe and A. Shimamura (MIT Press, Cambridge, MA, 1994).

[2] M. A. Conway, G. Cohen, and N. Stanhope, Very Long-Term Memory for Knowledge Acquired at School and University, Appl. Cog. Psych. **6**, 467 (1992).

[3] D. R. Bacon and K. A. Stewart, How Fast Do Students Forget What They Learn in Consumer Behavior? A Longitudinal Study, J. Mar. Educ. **28**, 181 (2006).

[4] G. E. Francis, J. P. Adams, and E. J. Noonan, Do They Stay Fixed? Phys. Teach. **36**, 488 (1998).

[5] P. S. Shaffer and L. C. McDermott, Research as a Guide for Curriculum Development: An Example from Introductory Electricity. Part II: Design of Instructional Strategies, Am. J. Phys. **60**, 1003 (1992).

[6] L. C. McDermott, P. S. Shaffer, and C. P. Constantinou, Preparing Teachers to Teach Physics and Physical Science by Inquiry, Phys. Educ. **35**, 411 (2000).

[7] S. J. Pollock, Longitudinal Study of Student Conceptual Understanding in Electricity and Magnetism, Phys. Rev. ST Phys. Educ. Res. **5**, 020110 (2009).

[8] M. Kohlmyer *et al.*, Tale of Two Curricula: The Performance of 2000 Students in Introductory Electromagnetism, Phys. Rev. ST Phys. Educ. Res. **5**, 020105 (2009).

[9] S. B. McKagan and C. E. Wieman, *Exploring Student Understanding of Energy through the Quantum Mechanics Conceptual Survey*, Physics Education Research Conference Proceedings 2005 (AIP, Melville, NY, 2006); S. B. McKagan, K. K. Perkins, and C. E. Wieman, Phys. Rev. ST Phys. Educ. Res. **6**, 020121 (2010); http://per.colorado.edu/QMCS.

[10] http://www.pisa.gc.ca/eng/pdf/81-590-E.pdf; http://www.cwsei.ubc.ca/resources/files/selectivity_report.pdf.

[11] www.phet.colorado.edu.

[12] M. A. McDaniel, H. L. Roediger III, and K. B. McDermott, Generalizing Test-Enhanced Learning from the Laboratory to the Classroom, Psychonomic bulletin and review **14** (2), 200 (2007).