Teaching for understanding in medical classrooms using multimedia design principles

Nabil Issa,[1] Richard E. Mayer,[2] Mary Schuller,[1] Edward Wang,[1] Michael B. Shapiro[1] & Debra A. DaRosa[1]

OBJECTIVES In line with a recent report entitled *Effective Use of Educational Technology in Medical Education* from the Association of American Medical Colleges Institute for Improving Medical Education (AAMC-IME), this study examined whether revising a medical lecture based on evidence-based principles of multimedia design would lead to improved long-term transfer and retention in Year 3 medical students. A previous study yielded positive effects on an immediate retention test, but did not investigate long-term effects.

METHODS In a pre-test/post-test control design, a cohort of 37 Year 3 medical students at a private, midwestern medical school received a bullet point-based PowerPoint™ lecture on shock developed by the instructor as part of their core curriculum (the traditional condition group). Another cohort of 43 similar medical students received a lecture covering identical content using slides redesigned according to Mayer's evidence-based principles of multimedia design (the modified condition group).

RESULTS Findings showed that the modified condition group significantly outscored the traditional condition group on delayed tests of transfer given 1 week ($d = 0.83$) and 4 weeks ($d = 1.17$) after instruction, and on delayed tests of retention given 1 week ($d = 0.83$) and 4 weeks ($d = 0.79$) after instruction. The modified condition group also significantly outperformed the traditional condition group on immediate tests of retention ($d = 1.49$) and transfer ($d = 0.76$).

CONCLUSIONS This study provides the first evidence that applying multimedia design principles to an actual medical lecture has significant effects on measures of learner understanding (i.e. long-term transfer and long-term retention). This work reinforces the need to apply the science of learning and instruction in medical education.

Medical Education 2013: ***47****: 388–396*
doi:10.1111/medu.12127

[1]Department of Surgery, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA
[2]Department of Psychology, University of California Santa Barbara, Santa Barbara, California, USA

Correspondence: Nabil Issa, Northwestern University Feinberg School of Medicine, Department of Surgery, Division of Trauma and Surgical Critical Care, 676 N. Saint Clair Street, Suite 650, Chicago, Illinois 60611, USA. Tel: 00 1 312 695 4835;
E-mail: nissa@northwestern.edu

INTRODUCTION

There is growing interest in applying the science of learning to medical education.[1,2] The Association of American Medical Colleges Institute for Improving Medical Education report, entitled *Effective Use of Educational Technology in Medical Education*,[3] recommends the use of Mayer's multimedia principles[4] when designing educational materials for medical students as the use of these principles has been shown to improve retention (remembering what was presented) and transfer (applying what was learned to solve novel problems) in undergraduate college students.[5–7]

The term 'multimedia' refers to the presence of both visual material (i.e. static images such as illustrations or dynamic images such as animation) and verbal material (i.e. printed words or spoken words), which, in this context, are used to inform and explain new knowledge to learners. With the advancement of technology and the increasing use of computers, multimedia material has become versatile and is easily accessed by the masses through widely available technologies, including computer software, thumb-drives, web-based education programmes and smart phone applications, to name a few. Educators have recognised the importance of using these rapidly evolving technologies as new potential environments for education.[8–10]

The design of effective multimedia educational materials should enhance the innate cognitive processes involved in how humans learn and retain information.[8–10] Effective designs help learners attend to relevant information, organise it into a coherent mental representation, and integrate it with prior knowledge. Poorly designed multimedia materials may overwhelm and confuse learners.[11] The cognitive theory of multimedia learning provides principles for developing effective, evidence-based multimedia messages. The theory was developed by Mayer[4] based on cognitive load theory[12] and dual-channel coding theory,[13,14] and is backed by empirical research[15] and real-world applications.[8,16,17]

The effectiveness of instructional design can be evaluated by assessing learners' understanding of the educational materials using appropriately designed tests. Early research[18,19] showed that rote learning leads to good performance on immediate retention post-tests, whereas meaningful learning leads to good performance on delayed retention post-tests and on immediate and delayed transfer post-tests.[20,21] Researchers have used delayed post-tests of transfer to measure whether students remember over time what was taught, largely when they want to assess the impact of a new instructional method aimed at promoting meaningful learning.[8]

In a previous study,[5] we found that revising lecture slides according to Mayer's multimedia design principles significantly improved medical students' performance on a post-test of retention given within an hour of lecture completion.[4] Our aim in the current study is to extend that research by investigating whether the use of slides designed according to Mayer's multimedia principles impacts on long-term retention and transfer of learning among medical students.[4]

METHODS

Participants and design

The participants were 80 third year medical school students at a private, midwestern medical school progressing through their surgery clerkship. Total class size was 156 students, with all students having to rotate through the 12 week surgery clerkship. The class was divided into 4 cohorts by the medical school registrar, with each cohort having about 40 students proceeding through the surgery rotation at any given time. A pre-test/ post-test control group design was used with instructional method (traditional versus modified) as the between subjects factor and time of test (pre-test, immediate post-test, 1-week post-test, and 4-week post-test) as the within subject factor. Forty three students progressing through the surgical clerkship in July, 2011 served in the modified group while 37 students progressing through the surgery clerkship in September, 2011 served in the traditional group. All students were in good academic standing. An Institutional Review Board approval was obtained (STU00016328) and informed consent was obtained from all participating students.

Materials and procedure

The required weekly core curriculum didactic lecture series served as the venue for this research. The focus was a 1-hour lecture on the topic of shock. The traditional slides for this lecture had been developed by the lecturer before learning about Mayer's multimedia design principles and included mostly text in a bullet point format.[4] The modified slides were developed by the instructor and changes were then reviewed with another content expert (MSh), as well as with an authority in multimedia principles (REM) and an educator (DD). The instructor analysed the key point or message of each slide in the traditional lecture and then determined how it could be changed to enhance understanding through the application of relevant multimedia principles. The modification process focused on adapting Mayer's design principles without changing each slide's teaching points as both lectures were premised on the same learning objectives.[4] The content of the shock lecture in its two forms, traditional and modified, was based on the objectives stated in the syllabus of the surgery clerkship. These objectives were drawn from the textbook *Essentials of General Surgery*,[22] which is the official textbook for the surgery clerkship at our institution. Learning objectives for the 50-minute lecture included: definition of shock; mechanisms of cellular dysfunction in shock; aetiology and classification of shock, and principles of diagnosis and management of shock.

Figure 1 shows a representative slide for the modification process using Mayer's design principles. Bulleted text was replaced by visual representation.[4] This is known as the multimedia principle. All pictures and text not directly related to the content were deleted based on Mayer's coherence principle. We highlighted important teaching points by using a larger font and a different colour based on Mayer's signalling principle. Graphs and related text appeared contiguously on the screen in accordance with the spatial contiguity principle. Slides reflecting complex materials, such as the electric waves generated when floating a pulmonary artery catheter, or the relationships between oxygen delivery, consumption and tissue extraction, were presented as a visual representation in the form of a picture or graph, and used vocal narration to explain these complex phenomena based on the modality principle. The process of slide modification helped to decrease the total slide count from a total of 35 slides in the traditional slide set to 28 slides in the modified slide set (Fig. 2).

A coin was tossed to determine which lecture type would be presented to the first or second clerkship. As a consequence, the modified lecture was presented to the first clerkship rotation of the year and the traditional lecture format was used in the second clerkship rotation of the year.

The lecturer used a conversational style of delivery during both sessions to help manage essential processing and understanding of the material. In both conditions, at the end of the lecture, the lecturer presented a clinical vignette in which data were progressively disclosed. After each set of data had been presented, students were asked to reach a diagnosis or provide an intervention and give appropriate justification for their answer. In-time feedback was provided by the lecturer to help students foster generative understanding of the topic. In a manner similar to that used in our previous research,[5] both lecture formats were given by the same lecturer to control for lecturer style differences.

The paper-based test materials included the previously developed five-item retention test and five-item transfer test.[5] Retention was assessed using open-ended questions, such as: 'Define shock'; 'Enumerate the different types of shock', and 'What are the numerical variables provided by the pulmonary artery catheter?' Transfer test questions were also open-ended and required respondents to apply the presented information to a clinical vignette of a patient in septic shock secondary to cholangitis. Students were asked to identify the clinical condition (septic shock) and explain the reasons for their answers based on the data provided to them in the question stem. They were also asked to institute appropriate initial therapies. The scenario then changed to describe the same patient developing a complication causing oxygen consumption to increase; students were asked to provide possible reasons for this clinical finding. The retention and transfer test questions were administered as one examination just before (pre-test) and 1 hour after the lecture (immediate post-test). The tests were re-administered 1 week and 4 weeks later (delayed post-tests). The same examination was used throughout. Students were given 10 minutes to complete the examination. The test format and question types were developed using the question format published in Mayer's studies on the effects of multimedia learning.[4] One point, half a point or no point was given for each of the questions. A tally of the five retention questions yielded the student's retention score and a tally of the five transfer questions yielded the student's transfer score.

Both the tests and corresponding answer keys were developed and field-tested in our previous study.[5] The two content experts (NI and MSc) who scored the tests in the original study[5] also scored the tests in the current study using the scoring rubric they had developed for the first study. Inter-rater reliability between the two surgeons during the first study demonstrated agreement for each question of ≥ 93%.[5] In the current study, all test questions were scored by one researcher and a randomly chosen sample of 20% of the scored tests were re-scored independently by another clinician-researcher to

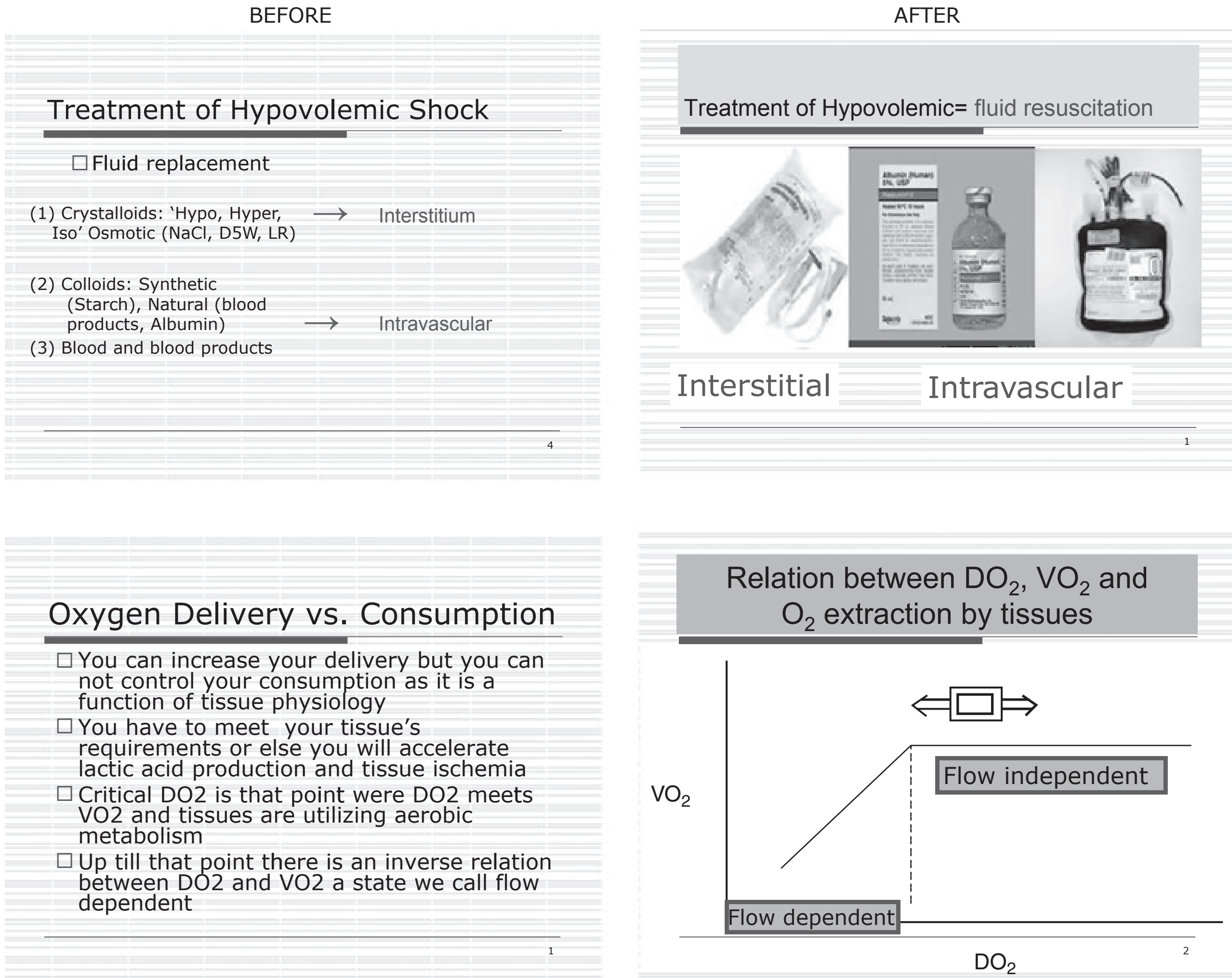

Figure 1 The AFTER slides illustrate the use of several multimedia design principles to modify the traditional bullet point-based (BEFORE) slides. (a) The multimedia principle is applied: bullet points are deleted and visual representations used. The signalling principle is also applied: important concepts are highlighted in different fonts and colours. (b) The modality principle is applied: visual representation and narration are used in place of bullet point text to explain complex phenomena. The spatial contiguity principle is used: printed words are displayed next to related portions of the graph

verify inter-rater agreement. Tests were bundled into envelopes by rotation. The scoring process did not start until all examinations had been completed. Scores were tallied by a researcher who was not involved in grading the examinations.

Statistical analysis

Data for participants who had missed one or more tests were eliminated, yielding final sample numbers of 40 participants in the modified condition group and 31 in the traditional condition group. There was no significant difference in dropout between the two experiment groups. Test scores for each group at each of the four data collection points were summarised using descriptive statistics including means and standard deviations (SDs). The comparison of scores between the two groups on the pre-test was accomplished using an independent samples *t*-test, whereas the comparison of scores within each group was accomplished using paired *t*-tests. We also compared students included and excluded from the study and found no significant difference in their pre-test scores. We further conducted analysis of covariance (ANCOVA) to compare the group differences in transfer and retention post-test scores using the pre-test score as a covariate. Finally, we employed a generalised mixed-effects model using group, time and the group*time interaction as fixed effects, and intercept as a random effect to compare the trajectory of change in test scores over time between the modified and the traditional condition groups. Inter-rater

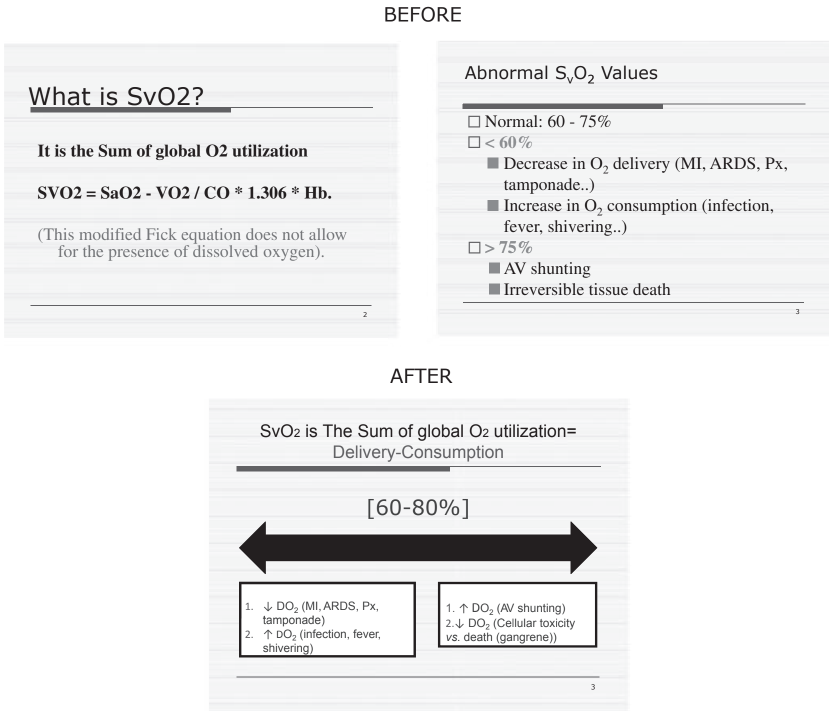

Figure 2 Example of how modification of slide design decreased slide redundancy: in this instance, two (BEFORE) slides were replaced by one (AFTER) slide

reliability was studied by calculating Pearson's correlation coefficients and internal test stability was assessed using Cronbach's alpha. Throughout the study, the effect size was computed using Cohen's *d*.Statistical analyses were performed using SAS Version 9.2 (SAS, Inc., Cary, NC, USA).[23] Statistical significance was set at $\alpha \leq 0.5$.

RESULTS

The overall inter-rater reliabilities for retention questions, transfer questions and total test scores were 0.72, 0.68 and 0.74, respectively. Internal consistency analysis of the test yielded a Cronbach's alpha of 0.66.

Table 1 shows the mean ± SD scores on the transfer test for the two groups on the pre-test, immediate post-test, 1-week post-test and 4-week post-test. The traditional condition group and the modified condition group did not differ significantly in scores on the pre-test transfer questions ($t[69] = 0.82$, $p = 0.4142$, $d = 0.20$). In the subsequent post-tests, however, the modified condition group outperformed the traditional condition group on the immediate transfer post-test ($t[69] = 3.21$, $p = 0.0020$, $d = 0.76$), 1-week post-test ($t[69] = 3.45$, $p = 0.0006$, $d = 0.83$) and 4-week post-test ($t[69] = 4.92$, $p < 0.0001$, $d = 1.17$). Using the pre-test score as a covariate, the ANCOVAS yielded similar results. On the immediate transfer post-test, the modified condition group achieved higher scores than the traditional condition group ($F[1,68] = 12.4$, $p = 0.0008$). A comparison of the groups on the delayed transfer post-tests also indicates that the modified condition group significantly outperformed the traditional condition group on the 1-week post-test ($F[1,68] = 15.8$, $p = 0.0004$) and the 4-week post-test ($F[1,68] = 28.6$, $p < 0.0001$).

Table 1 also shows the mean ± SD scores on the retention tests achieved by the two groups on the pre-test, immediate post-test, 1-week post-test and 4-week post-test. The groups did not differ significantly in scores on the pre-test ($t[69] = 0.82$, $p = 0.4156$, $d = 0.22$). In the subsequent post-tests, however, the modified condition group outperformed the traditional condition group on the immediate retention post-test ($t[69] = 6.29$, $p < 0.0001$, $d = 1.49$), 1-week post-test ($t[69] = 3.60$, $p = 0.0016$, $d = 0.83$) and 4-week post-test ($t[69] = 3.49$, $p = 0.0009$, $d = 0.79$). Using the pre-test score as a covariate, the ANCOVAS yielded similar results, with the modified condition group significantly outperforming the traditional

Table 1 Scores achieved by the modified and traditional condition groups on transfer and retention tests at four time-points

	Pre-test	Immediate post-test	1-week post-test	4-week post-test
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
Transfer test				
Modified condition (n = 40)	1.99 ± 1.18	3.71 ± 1.13	3.51 ± 1.16	3.71 ± 0.93
Traditional condition (n = 31)	2.21 ± 1.05	2.94 ± 0.83	2.61 ± 0.99	2.52 ± 1.12
t-test p-value*	0.4142	0.0020	0.0006	< 0.0001
ANCOVA p-value†	–	0.0008	0.0004	< 0.0001
Cohen's d‡	− 0.20	0.76	0.83	1.17
Retention test				
Modified condition (n = 40)	2.41 ± 0.65	4.41 ± 0.47	3.69 ± 0.77	3.51 ± 0.81
Traditional condition (n = 31)	2.27 ± 0.62	3.73 ± 0.44	3.13 ± 0.53	2.95 ± 0.54
t-test p-value*	0.4156	< 0.0001	0.0010	0.0009
ANCOVA p-value†	–	< 0.0001	0.0016	0.0022
Cohen's d‡	0.22	1.49	0.83	0.79

* Two sample t-test
† ANCOVA test adjusted using the pre-test score as a covariate
‡ Cohen's effect size
SD = standard deviation

condition group on the immediate ($F[1,68] = 38.1$, $p < 0.0001$), 1-week ($F[1,68] = 10.9$, $p = 0.0016$) and 4-week ($F[1,68] = 10.1$, $p = 0.0022$) post-tests.

As Fig. 3 shows, immediately after lectures, the traditional condition group showed a significant increase in transfer ($d = 0.61$, $p < 0.01$) and retention ($d = 1.89$, $p < 0.01$) scores; similarly, the modified condition group showed a significant increase from the pre-test to the immediate post-test on transfer ($d = 1.25$, $p < 0.01$) and retention ($d = 2.67$, $p < 0.01$). These improvements were significantly greater in the modified condition group than in the traditional condition group, yielding a significant group*time interaction effect for transfer ($p = 0.0019$) and retention ($p = 0.0049$) in the mixed-effect models. On the 1-week post-test, students in both the traditional and modified condition groups showed slight decays in transfer scores ($d = -0.31$, $p = 0.094$ and $d = -0.19$, $p = 0.240$, respectively). On the 4-week test, students in the modified condition group exhibited a slight increase in their transfer score ($d = 0.23$, $p = 0.156$), whereas those who attended the traditional lecture continued to show a decline ($d = -0.09$, $p = 0.604$). However, neither trend was statistically significant. Students in both groups illustrated significant decay in their retention scores on the 1-week post-test (traditional condition group: $d = -1.05$, $p < 0.001$; modified condition group: $d = -1.21$, $p < 0.001$). These declining trends continued for both the traditional ($d = -0.32$, $p = 0.086$) and modified ($d = -0.21$, $p = 0.189$) condition groups on the 4-week post-test, but were not statistically significant; however, the modified condition group's scores were still higher than those achieved by the traditional condition group.

DISCUSSION

This study builds on the work in instructional design pioneered by cognitive psychologists by applying the science of learning and instruction to medical education.

The current study showed that applying design principles based on the cognitive theory of multimedia learning improves medical students' long-term retention and transfer. Our findings regarding short-term retention and transfer are consistent with those noted in Mayer's earlier work with college students instructed using computer-based multimedia educational material in a controlled laboratory-like envi-

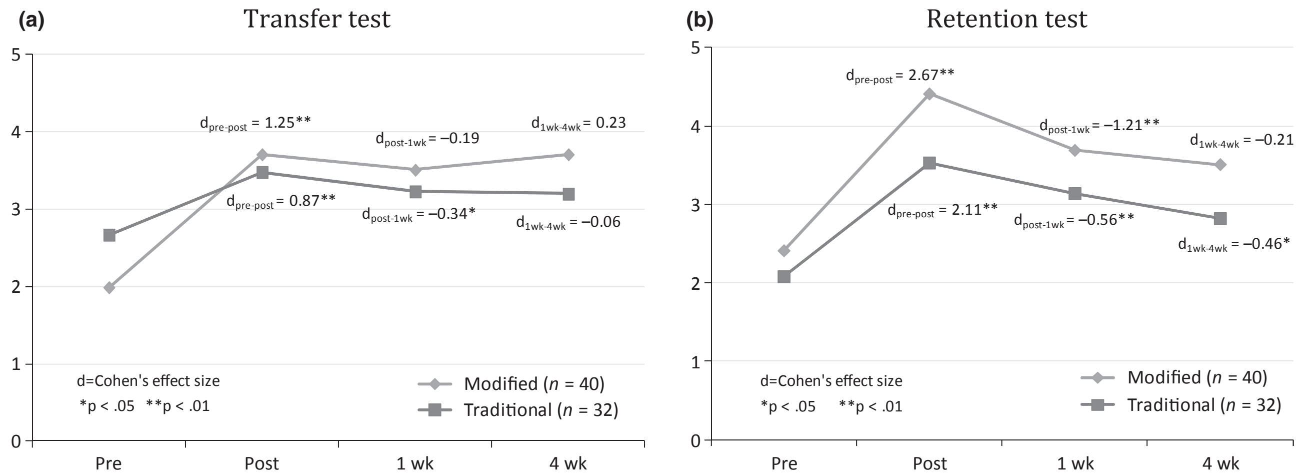

Figure 3 Scores on (a) transfer and (b) retention tests plotted by group over time. *d* = Cohen's effect size; * $p < 0.01$

ronment[4] and the findings of our previous study conducted in a classroom of medical students.[5] The current study expands on those early findings by showing that applying Mayer's design principles[4] to a multimedia lecture given to medical students improves not just short-term retention, but also long-term transfer of the study material, which researchers have identified as an important indicator of meaningful learning among students.

There were several limitations to the current study. The study was carried out at a single academic institution using a single lecture topic and a single lecturer, which may have introduced unconscious bias into the lecture presentations. Ratings of the lectures by the students, however, did not differ; ratings of 4.73 and 4.67 out of 5.00 were obtained from the modified condition group and the traditional condition group, respectively (the median overall faculty rating across topics was 4.24). This suggests that the lecturer was perceived as positively by students in the traditional condition as by students attending the modified lecture. We also did not verify whether students in either group actually read about the lecture topic prior to their attendance as instructed. The students' tests were bundled in envelopes marked by clerkship rotation date, which may have biased the graders' scoring of the tests. The answer key developed and used by both scorers in the original study[5] and in this study required fixed answers in order for an item to be scored as 'correct'. However, we cannot claim that the scoring was completely bias-free because the tests were bundled and marked by clerkship rotation. Researchers interested in building or replicating this research should ensure that the tests from both groups are coded to designate group type by a researcher in order to ensure graders are completely blinded. Another limitation of this study was that the process of lecture modification was performed by a team that included a multimedia specialist (REM), an educator (DD) and the instructor (NI). We do not know from this study the extent to which having a team modify a traditional lecture or the use of multimedia design principles truly yielded differences in learning outcomes. It is plausible that any kind of review and feedback on any teaching intervention will cause differences in learning outcomes, regardless of the extent to which multimedia principles were employed.

The current study design remained faithful to the methodology and materials used in the earlier study[5] in order to enhance the factor of time as the independent variable. Thus, the test questions did not vary between the data collection points, which may have allowed some students to remember their own answers and choose not to revise them. Using the same test material in a repeated fashion for the same cohort of learners may have introduced a testing effect. Learners have been shown to perform better on delayed tests (for both retention and transfer functions) after studying the material and taking a practice test than after studying the material twice.[24,25] The findings of those studies[24,25] suggest that tests have an enhancing effect on learning and advocate for the adoption of tests as a teaching tool. In our study design, both cohorts of students were exposed to the same testing effect, which would theoretically negate its effect on the results.

The move toward building electronic, self-learning modules for today's students suggests the need for audiovisual materials to reflect empirically based multimedia principles that are shown to support the retention of information and transfer of learning. In the present study, students in both groups remembered what they had learned in the session, but students in the modified lecture condition demonstrated increased and sustained learning to a greater extent than students in the traditional lecture condition. These data should provide convincing evidence to motivate faculty staff to learn about multimedia principles and to include them in their classroom lectures or distance learning modules.

CONCLUSIONS

Applying multimedia design principles to lecture slides given to medical students helps to improve both long-term retention and long-term transfer of the lecture material, in addition to having positive short-term retention effects, as discovered in previous research.[5] Medical educators should consider multimedia principles when designing lecture slides for medical students and abandon the use of word- or bullet point-based presentations.

Faculty development programmes for clinician-educators should adopt courses on evidence-based instructional design for their faculty members. These programmes should be evaluated by determining the degree to which they improve clinical instructors' ability to create courses that improve student learning outcomes better than traditional courses.

Contributors: NI contributed to forming of hypothesis, conception & design of the study, conducted the lectures, acquisition and interpretation of data, drafted the article. REM contributed to forming of hypothesis, conception & design of the study, analysis and interpretation of data, critically revised article. MSc contributed to forming of hypothesis, conception & design of the study, data acquisition and interpretation, drafted the article. EW contributed to design of the study, power analysis, data analysis and interpretation, drafted the article. MSh contributed to design of the study, wrote test questions, scored tests, data acquisition and data interpretation, drafted the article. DD contributed to forming of hypothesis, conception & design of the study, analysis and interpretation of data, critical revised the article. All authors approved the final manuscript for publication.
Acknowledgement: this study was completed as part of the Association for Surgical Education's Surgical Education Fellowship (SERF) programme. We appreciate the help and support offered by the organisation and extend particular thanks to Kimberly Schenarts, PhD, programme director of SERF.
Funding: none.
Conflict of interest: none.
Ethical approval: this study was approved by the Institutional Review Board, Northwestern University (Chicago, IL, USA) (ref. STU00016328).

REFERENCES

1 Mayer RE. Applying the science of learning to medical education. *Med Educ*, 2010;**44**:543–9.
2 Levinson AJ. Where is evidence-based instructional design in medical education curriculum development? *Med Educ* 2010;**44** (60):536–7.
3 Association of American Medical Colleges. Effective Use of Educational Technology in Medical Education. Candler C, ed. *Institute for Improving Medical Education.* Washington, DC: AAMC, 2007.
4 Mayer RE. *Multimedia Learning.* Cambridge: Cambridge University Press 2001.
5 Issa N, Schuller M, Santacaterina S, Shapiro M, Wang E, Mayer RE, DaRosa DA. Applying multimedia design principles enhances learning in medical education. *Med Educ* 2011;**45** (8):818–26.
6 Najjar LJ. Principles of educational multimedia user interface design. *Hum Factors* 1998;**40**:311–23.
7 Mayer RE. The promise of multimedia learning: using the same instruction design methods across different media. *Learn Instr* 2003;**13** (2):125–39.
8 DiGiacinto D. Using multimedia effectively in teaching-learning process. *J Allied Health* 2007;**36**:176–9.
9 Sweller J, Chandler P. Why some material is difficult to learn. *Cogn Instr* 1994;**12** (3):185–233.
10 Sweller J, Chandler P. Cognitive load theory and the format of instruction. *Cogn Instr*, 1991;**8** (4):293–332.
11 Clark J, Paivio A. Dual coding theory and education. *Educ Psych Rev* 1991;**3** (3):149–210.
12 Baddeley AD. *Human Memory: Theory and Practice.* Hove: Psychology Press 1997.
13 Sweller J. *Instructional Design in Technical Areas.* Adelaid, SA: National Center for Vocational Education Research 1999.
14 Starbek P, Starcic Erjavec M, Peklaj C. Teaching genetics with multimedia results in better acquisition of knowledge and improvement in comprehension. *J Comput Assist Learn* 2010;**26**:214–24.
15 Katona G. *Organizing and memorizing.* New York, NY: Columbia University Press 1940.
16 Wertheimer M. *Productive thinking.* New York, NY: Harper & Row 1959.
17 Anderson LW, Krathwohl DR, Airasian PW, Cruikshank KA, Mayer RE, Pintrich PR, Raths J, Wittrock MC. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* New York, NY: Longman 2000.
18 Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR. *Taxonomy of Educational Objectives: Cognitive Domain.*

New York, NY: David McKay 1956.
19 Mayer RE *Learning and instruction*, 2nd edn. Upper Saddle River, NJ: Pearson 2008.
20 Mayer R, Wittrock M Problem-solving transfer. In: Berliner D, Calfee R, eds. *Handbook of Educational Psychology*. Mahwah, NJ: Erlbaum 2006; 47–62.
21 Lawrence PF. *Essentials of General Surgery*. Baltimore, MD: Lippincott Williams & Wilkins 2006.
22 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum 1988.
23 Roediger HL III, Karpicke JD. The power of testing memory: Basic research and implications for educational practice. *Perspect Psychol Sci* 2006; (1):181–210.
24 Roediger HL III, Karpicke JD. Test-enhanced learning: Taking tests improves long-term retention. *Psychol Sci* 2006; (17):249–55.
25 Johnson CI, Mayer RE. A testing effect with multimedia learning. *J Educ Psychol* 2009;**101** (3):621–9.

Received 5 April 2012; editorial comments to authors 31 July 2012; accepted for publication 19 November 2012