Original Investigations JAMIA

Research Paper ■

Variables That May Enhance Medical Students' Perceived Preparedness for Computer-based Testing

Deirdre C. Lynch, RhD, Theodore W. Whitley, PhD, David A. Emmerling, EdD, Jack E. Brinn, PhD

Abstract **Objective:** To identify variables that may enhance medical student's preparedness for computer-based administration of the United States Medical Licensing Examination (USMLE).

Design: A cross-sectional survey of 301 medical students who completed a self-administered questionnaire.

Measurements: The questionnaire was designed to obtain information about students' computer resources, personal experience with computers, computer expertise, opinions about computers, experience with computer-based testing, perceived preparedness for the computer-based USMLE, and demographic variables. Variables related to students' perceived preparedness for the computer-based USMLE were identified by ordinal logistic regression.

Results: A significant regression model yielded four significant predictors: perceived preparedness for USMLE content ($P < 0.0001$), opinions about computers ($P < 0.0012$), gender ($P < 0.0001$), and a gender by computer-based testing experience interaction ($P < 0.0004$). Computer resources, personal experience with computers, computer expertise, age, race, and year of medical school were not significant predictors.

Conclusion: Students' perceived preparedness for computer-based administration of high-stakes examinations may be facilitated by preparing them for examination content, by enhancing their opinions about computers, and by increasing their computer-based testing experiences.

■ **J Am Med Inform Assoc.** 2000;7:469–474.

Affiliation of the authors: East Carolina University, Greenville, North Carolina.

This work was supported in part by the Generalist Physician Initiative of The Robert Wood Johnson Foundation.

Correspondence and reprints: Deirdre C. Lynch, RhD, 493 Sheridan, Apt. 2E, Evanston, IL 60202;
e-mail: ⟨dlynch@medscape.com⟩.

Received for publication: 11/15/99; accepted for publication: 5/10/00.

Computer-based testing (CBT), defined as the use of computers to administer tests,[1] is now used in high-stakes examinations such as the Graduate Record Examination and nursing licensure examinations. Instant scoring, flexibility in scheduling, and enhanced security are some of the advantages of CBT.[2] Computer-based versions of Steps 1, 2, and 3 of the United States Medical Licensing Examination (USMLE) have just begun to replace paper-and-pencil forms.[3]

Recent research suggests that most North American medical students have access to, and have used, computers in some capacity.[4,5] From 1994 to 1999, there have been increases in the number of graduating medical students in the United States who reported that adequate time was allocated toward instruction in the use of computers during medical school.[6] In addition, more than half the respondents indicated that computer-based programs had been a part of course instruction.

The transition to computer-based administration of the step examinations, however, has prompted questions about the differential impact of paper-and-pencil versus computer-based testing on test performance. These differential effects are referred to as mode effects.[7] Furthermore, there is some concern that mode effects may be related to examinee characteristics such as computer experience, opinions about computers, and demographics.

While speeded tests have stringent time limits to measure processing speed, power tests are designed to measure examinee knowledge of a specific content area. Providing unlimited time for examinees to take the USMLE is impractical; therefore, time limits are imposed for this power test. Research suggests that computer versions of timed power tests can be constructed to measure the same trait as corresponding paper forms. A meta-analysis of 21 studies that examined timed power tests completed by young adults and adults revealed a correlation of 0.97 between these modes of administration.[8] The mean and standard deviation of the scaled differences was −0.03, which indicated that the computer forms tended to be slightly harder than the paper forms, with relatively small variation from study to study.

The American Psychological Association has guidelines for the development, use, and interpretation of computerized tests. These guidelines indicate that paper and computer versions of the same test may be considered equivalent if the rank order of scores of individuals tested in both modes are similar and if the means, dispersions, and shapes of the score distributions are approximately the same.[9] To date, field tests of computer versions of the USMLE indicate that CBT does not have a differential effect on test performance.[10]

Research into the relationship between mode effects and examinee characteristics is less conclusive. It appears that neither level of computer experience nor anxiety about using computers affects performance on computer tests compared with paper tests.[11,12] On the other hand, findings with regard to mode effects due to gender, age, and race have been mixed. Results obtained from 1,114 examinees who completed computer and paper versions of the Graduate Record Examination (GRE) indicated that on the analytic scale, Asian examinees performed better than their paper scores predicted on the computer version, compared with African-Americans. With regard to the verbal scale of the GRE, nonwhite male examinees performed better on the computer version than their paper scores predicted, compared with nonwhite female examinees. On the same scale, younger African-American examinees performed better on the computer version than older African-American examinees. With regard to the quantitative scale, the age pattern held, and the overall performance of male examinees on the computer version was better than would be expected from their paper scores, whereas female examinees performed better on the paper version.[7]

As worldwide interest in CBT increases, it would be useful to know how to help medical students prepare for this transition. To date, variables that may influence student preparation for CBT have not been reported in the medical education literature. The specific purpose of this study was to examine the relationship between perceived preparation for computer-based administration of the USMLE step examinations and the following predictor variables: computer resources, computer expertise, personal experiences with computers, experience with CBT, opinions about computers, perceived preparedness for USMLE content, and demographics.

Method

Design and Instrument

A cross-sectional survey of 301 students who attended a public medical school in the southeastern United States was conducted. First-, second-, third-, and fourth-year students completed a self-administered questionnaire designed to obtain information about several aspects of computer use. Information about personal computer resources and computer expertise at the time of entering medical school was assessed by summing four and five differentially weighted experiences, to yield composite ''resource'' (e.g., own a computer, have an e-mail address) and ''expertise'' (e.g., self-taught, degree in computer science) variables, respectively (Table 1). Specific computer experiences were assessed by the sum of 17 differentially weighted experiences to yield a composite ''experience'' variable. Experience with CBT was the sum of four equally weighted experiences, and opinions about computers was the sum of scores obtained from the Opinions about Computers scale.

The Opinions about Computers scale consisted of eight items derived from two reliable and valid instruments, the Computer Opinion Survey[13] and the Computer Anxiety Scale.[14] It used a five-point Likert scale ranging from "strongly disagree" to "strongly agree." An internal consistency reliability coefficient (Cronbach alpha) of 0.79 was observed in this study, compared with previously observed alphas of 0.94 and 0.95 for the Computer Opinion Survey and the Computer Anxiety Scale, respectively. Information about participant preparedness for completing the USMLE on computer and preparedness for USMLE content was obtained from the questionnaire, as were demographic data including student gender, age, ethnicity, and year in medical school. Degree of preparedness was assessed on a five-point scale ranging from "very unprepared" to "very prepared."

Data Analysis

Based on a review of the literature, the following predictor variables were identified: computer resources; expertise; experiences; CBT experience; opinions about computers; perceived preparedness for USMLE content; gender; age, race, year of medical school; and interactions between gender, age, race and personal computer experiences, CBT experience, and opinions about computers (nine interactions). Degree of perceived preparedness for completing the USMLE step examinations on computer was the criterion variable. A stepwise ordinal logistic regression model (SAS System for Windows, version 6.14) was used to analyze the data. An alpha level of 0.05 was the criterion for statistical significance of the relationship between the criterion and predictor variables and individual predictor variables.

Results

Eighty-seven percent (69 of 79), 87 percent (62 of 71), 99 percent (74 of 75), and 82 percent (62 of 76) of first-, second-, third-, and fourth-year students, respectively, completed the questionnaire. Male students made up 52 percent of the group, and the mean ages of the first-, second-, third-, and fourth-year students were 26, 27, 27, and 29 years, respectively. Eighty percent were white, and 20 percent were African-American, American Indian, Asian, or Hispanic. With the exception of race (there are proportionally more under-represented minority students at this school), these demographics are comparable with those of students in other U.S. medical schools.[15]

Most respondents owned a computer (70.4 percent), had access to a computer (98.5 percent), had an e-mail address (94.8 percent), and were connected to the Internet (48.3 percent) (Table 1). With regard to specific computer experiences, the data were consistent with national findings. The majority of students had used e-mail (97 percent), educational software (75 percent), and conducted online literature searches (88 percent). Twenty-seven percent had used computers to do statistics. Similarly, 96 percent, 79 percent, 85 percent, and 29 percent of students nationwide have experience using e-mail, using educational software, conducting online literature searches, and using computers to do statistics, respectively.[4]

Table 1 ■

Computer Resources, Expertise, and Experiences of Medical Students

Questionnaire Item (Weight)	% Students (N = 267)
Resources:	
1. Own a computer? (2)	70.4
2. Computer connected to the Internet? (3)	48.3
3. Have access to a computer at school? (1)	98.5
4. Have an e-mail address? (1)	94.8
Level of computer expertise at beginning of medical school:	
1. Self-taught or informal experience (1)	68.9
2. A course or course work in using a particular program (2)	26.2
3. A course or course work in writing particular programs (3)	9.4
4. A degree in computer science (4)	0.0
5. Other (1)	1.5
Computer experience:	
1. Word processing (2)	97.8
2. Statistics (3)	27
3. Graphics (3)	46.1
4. Use a database management/filing program (3)	29.2
5. Use a spreadsheet (3)	61.8
6. Write a computer program (3)	19.9
7. Play a game (1)	81.3
8. Use e-mail (1)	96.6
9. Access the Internet (1)	89.9
10. Teleconferencing (2)	5.2
11. Use an Internet chat room (2)	30.7
12. Use educational software (2)	75.3
13. Access class materials (2)	64.4
14. Do an examination or quiz for a class (3)	53.6
15. Play a CD-ROM (1)	83.9
16. Do an on-line literature search (1)	87.6
17. Other (1)	2.6
Experience with computerized testing:	
1. Answered practice questions on a CD-ROM (1)	73.8
2. Answered practice questions on a floppy disk (1)	35.6
3. Attended a presentation on computerized testing (1)	12.0
4. Other (1)	4.5

Table 2 ■

Results of Stepwise Ordinal Logistic Regression Analysis to Predict Preparedness for Computerized Administration of the U.S. Medical Licensing Examination (USMLE)

Variable	Coefficient (Beta)	Standard Error	Wald Chi-square	*P* Value	Odds Ratio	95% CI*
Intercept 1	−7.4	0.91	66.22	0.0001	–	–
Intercept 2	−5.48	0.85	41.1	0.0001	–	–
Intercept 3	−3.64	0.82	19.78	0.0001	–	–
Intercept 4	−2.48	0.8	9.54	0.002	–	–
Prepared for USMLE content	0.71	0.11	38.5	0.0001	2.034	1.63–2.55
Opinions about computers	0.08	0.02	10.53	0.0012	1.08	1.03–1.13
Gender	−1.76	0.37	22.38	0.0001	0.172	0.08–0.36
Gender by CBT experience	0.83	0.23	12.77	0.0004	2.292	1.45–3.61

NOTE: CBT indicates computer-based testing.
*The 95% confidence interval for the estimated odds ratio.

The proportional odds assumption, required for the ordinal logistic regression model, was tested for goodness of fit and met ($\psi^2 = 4.0087$, $P = 0.2605$). The relationship between the criterion and predictor variables ($P < 0.0001$) was significant. Significant predictors of perceived preparedness for completing the USMLE on computer were perceived preparedness for USMLE content ($P < 0.0001$), positive opinions about computers ($P < 0.001$), gender ($P < 0.0001$), and a gender by CBT experience interaction ($P < 0.0001$) (Table 2). Computer resources, expertise, experiences, age, race, year of medical school, and the interactions of perceived preparation for USMLE content, CBT experience, and opinions about computers with age and race were not significant predictors.

Bivariate correlation analyses among the predictor variables indicated very few substantial associations. While year of medical school and computer experience were significantly related to four other variables, exclusion of these two variables from the analysis did not change the statistical significance of the remaining variables or the final regression model.

Discussion

This study extends previous research by examining variables related to medical student preparedness for the computer-based USMLE. Such information is very relevant to test-taking skill development and to designing medical informatics curricula,[16] now considered a requisite component of medical education.[17] The results indicate that students who felt more prepared for the computer-based version of the USMLE step examinations were also more likely to feel more prepared for USMLE content, be male, be female with CBT experience, and have more positive opinions about computers.

The findings of this study are consistent with research that indicates a weak relationship between CBT performance and computer expertise.[11] At first glance, the results seem incongruent with studies that suggest that mode of test administration (i.e., paper versus computer) is not influenced by opinions about, or anxiety toward, computers.[11,12] It has been suggested, however, that these variables might become influential as computer tasks become more complex,[18,19] a phenomenon labeled "situation-specific anxiety." Because it is a licensing examination for physicians, the USMLE is both a high-stakes and complex computer task. The specific anxiety associated with preparing for it could be characterized, therefore, as "USMLE anxiety." Thus, students who do not feel prepared for USMLE content, and who have negative opinions about computers, may experience additional anxiety that makes them feel relatively more unprepared for the computer-based versions of the USMLE.

Previous research suggests that mode effects due to demographic variables are dependent also on the nature of computer tasks. Although gender did not interact with mode of test administration for relatively low-stakes tests such as psychological assessments,[20] gender differences in perceived self-efficacy regarding complex computer tasks were evident in other studies.[7,18] It may be possible, however, to decrease gender differences by preparing female students for complex computer tasks. Consequently, female students with more experience in relevant computer tasks might feel more prepared for high-stakes, complex computer tasks, which was observed in the present study. Specifically, female students with CBT experience were more likely to feel prepared for the computer-based USMLE.

The results suggest variables that could be examined to identify students who may be apprehensive about

the transition of high-stakes examinations to CBT and, consequently, allow schools to tailor interventions for these students. The findings indicate that medical schools can prepare students for computerized administration of the USMLE by readying them for USMLE content and by increasing student exposure to CBT. The latter may be achieved by increasing the integration of CBT into basic science and clinical courses. Current data indicate that 31 percent of medical students nationwide experience CBT as part of the medical education curriculum.[16] There is, therefore, ample room for improvement in this area.

Although the external validity of this study is limited by the use of a single medical school, this concern is mitigated by similarities in demographics and computer experiences between the sample of students in this study and the national population of medical students.[4] At the same time, we recognize that each school has its own culture, and additional studies would be useful to confirm the generalizability of the findings.

The present study was concerned with assessing a subjective variable that could be measured only by respondent self-report. The psychology literature suggests that self-report data can be reliable and valid under the following conditions: when an emotional state is being assessed (e.g., how do you feel?) rather than a request for causal attribution (why do you feel this way?)[21]; when the questions are not threatening, which may not be the case, for instance, when one is being asked about sexual behavior[22]; and when questions are simply stated.[23] We believe that the item "To what extent do you feel prepared to take the USMLE on computer?" meets these criteria.

The psychology literature also suggests a strong positive link between perception of ability to perform specific tasks and subsequent performance.[24] This appears to be the case for computer use.[25] In addition, the medical education literature suggests positive relationships between medical students' perceptions of skills and evaluations and those rendered by more objective benchmarks. With regard to students' reports of readiness, for instance, Fincher et al.[26] found a correlation of 0.58 between reported readiness for internship training, as assessed by a single item, and subsequent evaluation by internship directors. Furthermore, positive relationships between medical students' perceptions of their clinical skills and ensuing clerkship performances have been reported.[27–29]

Finally, we believe that the self-report data obtained from the participants in the present study are conceptually sound. To ascertain the latter, data obtained from the fourth-year students in our study were compared with responses by the same students to items on the Association of American Medical Colleges (AAMC) Graduating Questionnaire. In terms of the relationship between actual computer experiences and confidence in having specific skills and knowledge, there were several consistencies between the information obtained via the questionnaire used in our study and that used by the AAMC. Ninety-seven percent, 90 percent, and 26 percent of students in our study stated that they had used e-mail, done an online literature search, and done statistics on computer, respectively. On the AAMC Graduating Questionnaire,[6] 93 percent, 96 percent, and 24 percent of the same cohort of students indicated that they were "confident they had the appropriate knowledge and skills to, a) use electronic mail, b) carry out reasonably sophisticated searches of the medical education databases, and c) use a statistical software package," respectively. Evidence of consistency between students' reported confidence in having specific computer skills and knowledge in conjunction with their reported computer experiences suggests that students would be consistent, too, in reporting their perceived preparedness for computer-based testing of the USMLE and their actual preparedness.

A multi-item scale would probably be a more desirable approach to assessing the outcome variable, but we found no precedent in the literature for a scale with enough questions to improve reliability and validity given the specific nature of the outcome variable (i.e., perceived preparedness for taking the USMLE by computer).

In conclusion, this study indicates that some students do not feel prepared for the computer-based version of the USMLE. However, implementation of activities designed to increase preparation for examination content, improve opinions about computers, and increase CBT experiences may help to mitigate this perceived lack of preparedness.

References ■

1. Bugbee, AC. The equivalence of paper-and-pencil and computer-based testing. J Res Comput Educ. 1996;28:282–99.
2. Wise SL, Plake BS. Computer-based testing in higher education. Measure Eval Counseling Develop. 1990;23:3–10.
3. National Board of Medical Examiners. Plans for Administering the Medical Licensing Examination on Computer. Philadelphia, Pa: NBME, 1998.
4. Association of American Medical Colleges. Contemporary Issues in Medical Education: Keeping Up with Technology and the Changing Role of Medicine. Washington, DC: AAMC, 1999.
5. Haynes RB, McKibbon A, Bayley E, Walker CJ, Johnston ME. Interest in knowledge and use of information technol-

ogy by entering medical students at McMaster University in successive annual surveys. Proc 17th Annul Symp Comput Appl Med Care. 1994:560–3.
6. Association of American Medical Colleges. Medical School Graduation Questionnaire. Washington, DC: AAMC, 1994–1999.
7. Parshall CG, Kromrey JD. Computer testing versus paper-and-pencil testing: an analysis of examinee characteristics associated with mode effect. Presented at: The Annual Meeting of the American Educational Research Association; Atlanta, Ga; April 12–16, 1993.
8. Mead AD, Drasgow F. Equivalence of computerized and paper-and-pencil cognitive ability tests: a meta-analysis. Psychol Bull. 1993;114:449–58.
9. American Psychological Association. Guidelines for Computer-based Tests and Interpretations. Washington, DC: APA, 1986.
10. Luecht RM, Hadadi A, Swanson DB, Case SM. A comparative study of a comprehensive basic sciences test using paper-and-pencil and computer-based formats. Acad Med. 1998; 73(10 suppl):S51–S53.
11. Powers DE, O'Neill K. Inexperienced and anxious computer users: coping with a computer-administered test of academic skills. The Praxis Series: Professional Assessments for Beginning Teachers. Princeton, NJ: Educational Testing Service, 1992.
12. Webster J, Compeau D. Computer-assisted versus paper-and-pencil administration of questionnaires. Behav Res Methods Instrum Compute. 1996;28:567–76.
13. Lloyd BH, Gressard CP. Reliability and factorial validity of computer attitude scales. Educ Psychol Measure. 1984;44: 501–5.
14. Mauser M, Simonson M. Development and validation of a measure of computer anxiety. Ames, Iowa: Instructional Resources Center, 1984.
15. Association of American Medical Colleges. Medical School Matriculating Student Questionnaire. Washington, DC: AAMC, 1996–1999.
16. Association of American Medical Colleges. Contemporary Issues in Medical Education: Computers and Medical Informatics in the Curriculum. Washington, DC: AAMC, 1998.
17. Association of American Medical Colleges. Medical School Objectives Report I: Learning Objectives for Medical Student Education. Washington, DC: AAMC, 1998.
18. Busch T. Gender differences in self-efficacy and attitudes toward computers. J Educ Comput Res. 1995:147–58.
19. Bernt FM, Bugbee AC, Arceo RD. Factors influencing student resistance to computer administered testing. J Res Comput Educ. 1990;22:265–75.
20. Miles EW, King WC. Gender and administration mode effects when paper-and-pencil personality tests are computerized. Educ Psychol Measure. 1998;58:68–76.
21. Hollon SD, Bemis KM. Self-report and the assessment of cognitive functions. In: Hersen M, Bellack AS (eds). Behavioral Assessment: A Practical Handbook. New York: Pergamon, 1981:125–74.
22. Schaeffer NC. Asking questions about threatening topics: a selective overview. In: Stone AA, Turkkan JS (eds). The Science of Self-report: Implications for Research and Practice. Mahwah, NJ: Lawrence Erlbaum, 1999:105–21.
23. Fowler FJ. Survey Research Methods. Newbury Park, Calif: Sage, 1993.
24. Bandura A. Social Foundations of Thought and Action : A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall, 1986.
25. Harrison AW, Rainer RK Jr., Hochwarter WA, Thompson KR. Testing the self-efficacy–performance linkage of social-cognitive theory. J Soc Psychol. 1997;137:79–87.
26. Fincher RM, Lewis LA, Kuske TT. Relationship of interns' performance to their self-assessments of their preparedness for internship and to their academic performances in medical school. Acad Med. 1993;68(2 suppl):S47–S50.
27. Herbert WN, McGaghie WC, Droegemueller W, Riddle MH, Maxwell KL. Student evaluation in obstetrics and gynecology: self- versus departmental. Obstet Gynecol. 1990;76: 458–61.
28. Kaiser S, Bauer JJ. Checklist self-evaluation in a standardized patient exercise. Am J Surg. 1995;169:418–20.
29. Detering R, Kamin C, Barley G, Adams L, Dwinnell B, Merenstein G. Effect of a longitudinal course on student performance in clerkships. Arch Pediatr Adolesc Med. 1999; 153:755–60.