Comparison of Certification and Recertification Examinee Performance on Multiple-Choice Items in Forensic Psychiatry

Dorthea Juul, PhD, Jennifer Vollmer, BS, Linjun Shen, PhD, and Larry R. Faulkner, MD

Research on the association between age and performance on tests of medical knowledge has generally shown an inverse relationship, which is of concern because of the positive association between measures of knowledge and measures of clinical performance. Because the certification and maintenance of certification (MOC) examinations in the subspecialty of forensic psychiatry draw on a common item bank, performance of the two groups of examinees on the same items could be compared. In addition, the relationship between age and test performance was analyzed. Performance on items administered to certification and MOC examinees did not differ significantly, and the mean amount of time spent on each item was similar for the two groups. Although the majority (five of eight) of the correlations between age and test score on the certification and MOC examinations were negative, only three were significant, and the amount of variance explained by age was small. In addition, examination performance for those younger than 50 was similar to those 60 and older, and diplomates recertifying for the second time outperformed those doing so for the first time. These results indicate that in this subspecialty, there is no clear evidence of an age-related decline in knowledge as assessed by multiple-choice items.

J Am Acad Psychiatry Law 44:91–95, 2016

Individual medical expertise is rooted in a broad and deep knowledge base. As Dr. Geoffrey Norman states, ". . . it is now clear that an expert possesses superior knowledge of many kinds, both formal and informal, and any or all may be brought to bear on the solution of a particular problem" (Ref. 1, p 425). Almost all medical licensure, certification, and recertification processes include assessments of knowledge, and a positive relationship between test scores and performance in practice has been demonstrated in several studies. Scores on the Canadian licensure examinations predicted the quality of care provided by general practitioners.[2–5] Holmboe and colleagues[6] found a positive correlation between scores on an internal medicine recertification examination and several quality-of-care measures (hemoglobin A1c monitoring, lipid testing, and retinal eye examinations for patients with diabetes and mammography screening for female patients). Hess and colleagues[7] found a positive relationship between performance on an internal medicine recertification examination and a composite measure of diabetes care that included process, outcome, and patient experience measures.

Hence, the evidence that performance on knowledge measures declines over time is of concern. Choudhry and colleagues[8] reviewed 12 studies that analyzed this relationship, and all of them reported decreasing test performance with increasing age or years in practice. Four of these studies[9–12] involved physicians who voluntarily participated in recertification (now referred to as maintenance of certification or MOC) examinations. Leigh and colleagues[13] found that the performance of family practice physicians steadily declined on mandatory recertification examinations taken 6, 12, and 18 years after initial certification. Rhodes and Biester[14] found that surgery recertification examination fail rates grew with increasing number of years since initial

Dr. Juul is Vice President for Research and Development, Ms. Vollmer is Manager, Test Scoring Services, Dr. Shen is Vice President for Test Development and Core Competencies, and Dr. Faulkner is President and Chief Executive Officer, American Board of Psychiatry and Neurology, Buffalo Grove, IL. Address correspondence to: Dorthea Juul, PhD, 2150 East Lake Cook Road, Ste. 900, Buffalo Grove, IL 60089. E-mail: djuul@abpn.com

Disclosures of financial or other potential conflicts of interest: None.

certification, and Lipner and colleagues[15] found that age correlated negatively with performance on mandatory MOC examinations for both internists and surgeons.

Possible explanations for this decline include general decline in cognitive ability associated with aging and failure to keep up with changes in medical practice. A study by Day and colleagues[16] of performance on a voluntary recertification examination in internal medicine lent support to the latter hypothesis. Examinees who were further out of training performed less well on items that tested new or changing knowledge, whereas performance on items that tested stable knowledge was relatively constant across age groups.

In a recent study of family physicians, however, those who took an examination for MOC performed better than those who took the same examination for certification, although there was some decline in performance beginning at about age 60.[17]

The purposes of this study were to compare performance in response accuracy and testing time on specific test items administered to candidates for certification and MOC in a psychiatric subspecialty, forensic psychiatry, and to examine the relationship between age and performance on certification and MOC examinations.

Subspecialty in Forensic Psychiatry

Forensic psychiatry is concerned with the application of psychiatry to legal matters, such as violence, criminal responsibility, competence (civil and criminal), child custody and visitation, psychic injury, mental disability, malpractice, confidentiality, involuntary treatment, correctional psychiatry, juvenile justice, and ethics and human rights.[18]

Beginning in the mid-1970s, certification in forensic psychiatry was available through the American Board of Forensic Psychiatry, cosponsored by the American Academy of Psychiatry and the Law, the American Academy of Forensic Sciences, and the Forensic Sciences Foundation; 260 certificates were awarded between 1976 and 1993, an average of 14 certificates per year.[19]

In 1992, the American Board of Psychiatry and Neurology (ABPN) received approval from the American Board of Medical Specialties to issue a subspecialty certificate in forensic psychiatry, and the first examination was administered in 1994. During the grandfathering period (1994–2001), applicants could qualify for the certification examination either by having completed fellowship training in the subspecialty or by meeting practice requirements. During this period, 1,384 certificates were issued by the ABPN. The requirement for an ACGME-accredited fellowship went into effect with the 2003 examination, and since then, 741 certificates have been issued, for a total of 2,125 certificates in forensic psychiatry by the ABPN.

The Accreditation Council for Graduate Medical Education (ACGME) approved the fellowship training requirements in 1996. At the time of application to the ABMS, there were 18 non–ACGME-accredited forensic psychiatry fellowship programs, with 32 filled positions.[20] In academic year 2013–2014 there were 39 accredited programs with 73 on-duty residents,[21] and it is estimated that most graduates seek ABPN certification.

Successful examinees received 10-year time-limited certificates, and the first MOC examination was administered in 2003. The same test blueprint (in terms of content categories) and item bank were used to develop both examinations, and hence it was possible to compare the performance of certification and MOC candidates on the same items. Of the diplomates who have needed to recertify, approximately 60 percent have done so.[20]

Methods

There were 127 items on three recent MOC examinations (administered in 2011, 2012, and 2013) that had been administered on prior certification examinations. Their mean difficulties on the two types of examinations were calculated, and a *t* test was performed to compare them. Item difficulty is the percentage of examinees who answer the item correctly and can range in value from 0 percent (no one answered correctly) to 100 percent (everyone selected the right answer).

The amount of time examinees spent on each item was also available for 122 of the 127 items. The early certification examinations were administered by paper-and-pencil rather than computer, and item times were not available for items on those tests. The mean times spent on the two types of examinations were calculated, and a *t* test was performed to compare them.

Four certification cohorts and four MOC cohorts were selected to study the relationship between age and test performance. The certification cohorts were

those who took the first two examinations, which were administered in 1994 and 1996 (ACGME-accredited fellowship not required), and those who took the first two examinations that were administered when ACGME-accredited fellowship training was required (administered in 2003 and 2005). The MOC cohorts were those who took the first two MOC examinations which were administered in 2003 and 2004 and those who took two recent MOC examinations (administered in 2012 and 2013). Because scores were not equated across the examinations, which differed somewhat in difficulty, Pearson correlation coefficients for age and test score were calculated for each cohort.

As part of the ABPN certification and MOC application processes, applicants signed a statement allowing the board to release information about examination results and examination scores, provided that such data were reported in the aggregate.

Results

For the 127 items used on both certification and MOC examinations, the mean item difficulty was 0.83 (SD = 0.12) when used for certification and 0.85 (SD = 0.12) when used for MOC. The mean item difficulties were not significantly different ($t = -1.73$; $df = 126$; $p = .09$).

Although the items did not function differently overall across the two examinee groups, eight items on which there were relatively large differences (difference in difficulty $> \pm 0.15$) were identified. The MOC mean difficulties were greater than the certification mean difficulties for six of them, and for two of them, it was the reverse. One of the authors who is a content expert (L.R.F.) reviewed the items to determine whether such factors as content area, cognitive level (factual recall versus problem solving), and changes in knowledge (new, changing, or stable) might account for the differences. He concluded that there were no readily identifiable characteristics that account for differences in performance related to certification versus MOC status and that all of the items were relevant to practice.

The mean amount of time spent on each item ($n = 122$) was 53 seconds/item for certification examinees and 56 seconds/item for MOC examinees, a significant difference ($t = 4.23$; $df = 121$; $p < .01$).

The number of examinees, mean test score, mean age, and correlation between age and test score are summarized in Table 1 for four certification cohorts and four MOC cohorts.

For the four certification cohorts, the mean examinee ages were 47 in 1994 and 46 in 1996. In 2003 and 2005, the mean examinee ages were 37 and 40, respectively, as would be expected when the grandfathering period had ended and ACGME accredited training had become a requirement. There were significant negative correlations between test score and age in two examination years (-0.18 ($r^2 = 0.03$) in 1996 and -0.53 ($r^2 = 0.28$) in 2005; $p < .01$ for both correlations). The other two correlations were not significantly different from zero.

For the four MOC cohorts, the mean examinee ages ranged from 54 in 2003 and 2013 to 58 in 2012. There was one significant negative correlation between test score and age (-0.40 ($r^2 = 0.16$) in 2004; $p < .01$). The other three correlations were not significantly different from zero.

Table 1 Mean Test Scores and Ages and Correlations Between Age and Test Score for Certification and MOC Cohorts

Examination	Examinees (*n*)	Test Score	Age	Age Range	Correlation (*r*)
ACGME: accredited fellowship training not required					
Cert 1994	316	0.81 (0.09)	47 (9)	32–74	−0.08
Cert 1996	292	0.73 (0.09)	46 (9)	31–72	−0.18*
ACGME: accredited fellowship training required					
Cert 2003	83	0.89 (0.04)	37 (5)	31–52	0.02
Cert 2005	127	0.83 (0.07)	40 (7)	30–61	−0.53*
First two MOC examinations					
MOC 2003	61	0.90 (0.04)	54 (8)	40–78	−0.14
MOC 2004	84	0.88 (0.05)	55 (8)	41–76	−0.40*
Two recent MOC examinations					
MOC 2012	23	0.81 (0.11)	58 (8)	44–75	0.32
MOC 2013	93	0.81 (0.08)	54 (9)	41–80	0.11

Data are the mean (SD), unless stated otherwise.
* $p < .01$.

Of the 23 examinees who sat for the 2012 MOC examination, 11 were recertifying for the first time, and 12 were recertifying for the second time. The mean ages and scores for the first-time recertifiers were 55 (SD = 7) and 0.73 (SD = 0.10), respectively. The mean ages and scores for the second time recertifiers were 62 (SD = 7) and 0.88 (SD = .05), respectively. The second-time recertifiers were significantly older than the first-time recertifiers ($t = -2.37$; $df = 21$; $p = .03$), and they scored significantly higher on the examination ($t = -5.00$; $df = 21$; $p < .01$).

Of the 93 examinees who sat for the 2013 MOC examination, 49 were recertifying for the first time, and 44 were recertifying for the second time. The mean ages and scores for the first time recertifiers were 47 (SD = 4) and 0.79 (SD = 0.07), respectively. The mean ages and scores for the second time recertifiers were 62 (SD = 7) and 0.83 (SD = 0.08), respectively. Like the 2012 examinees, the second time recertifiers were significantly older than the first time recertifiers ($t = -11.92$; $df = 91$; $p < .01$), and they scored significantly higher on the examination ($t = -2.55$; $df = 91$; $p = .01$).

Based on the recent family medicine experience with recertification,[17] the examinees were split into two groups: those younger than 50 and those 60 and older. (In 2003 and 2005 there were too few certification examinees in the older age group to make this comparison.) As shown in Table 2, the differences between the two groups on the 1994 and 1996 certification examinations were minor and favored the younger group. On two of the four MOC examinations (2012 and 2013), the 60 and older group performed slightly better than the younger group.

Table 2 Comparison of Test Scores by Age Group for Certification and MOC Cohorts

	Age <50		Age ≥60	
Year	Examinees (*n*)	Test Score	Examinees (*n*)	Test Score
Cert 1994	206	0.82 (0.09)	22	0.80 (0.07)
Cert 1996	194	0.74 (0.08)	23	0.70 (0.08)
MOC 2003	21	0.91 (0.04)	16	0.90 (0.03)
MOC 2004	25	0.90 (0.04)	25	0.85 (0.06)
MOC 2012	3	0.75 (0.08)	10	0.82 (0.09)
MOC 2013	33	0.79 (0.07)	26	0.83 (0.08)

Data are the mean (SD), unless stated otherwise.

Discussion

In this study, items administered to certification and MOC examinees did not differ significantly in difficulty (percentage of examinees answering the item correctly). Although the mean amount of time spent on each item by MOC examinees was greater than the time spent by certification examinees, the difference was not of practical significance, with both groups spending less than one minute per item. Although the majority (five of eight) of the correlations between age and test score on the certification and MOC examinations were negative, only three were significant, but the percentage of variance explained by age was small ($r^2 \leq 0.28$). Performance on both the certification and MOC examinations for those younger than 50 was similar to that of those 60 and older. For the two recent MOC examinations, those recertifying for the second time (about 20 years postinitial certification) were significantly older and scored significantly better than those recertifying for the first time (about 10 years postinitial certification).

These results indicate that in this subspecialty, there is no clear evidence of an age-related decline in knowledge as assessed by multiple-choice items. Review by a subject matter expert of items on which there were large differences in item difficulty for the two groups of examinees reinforced this conclusion. The findings from this study differ from those of most other studies and may reflect a rigorous approach to test development that emphasized assessing knowledge relevant to practice.

Another factor that may have affected the results is the self-selection of the MOC candidates. A significant proportion (about 40%) of those initially certified have chosen not to participate in MOC. Their decision may reflect career stage, given that many of the early diplomates were already in midcareer, or a determination that subspecialty certification was not needed for practice. The rate of recertification is similar to that for the ABPN subspecialties in addiction and geriatric psychiatry but lower than that for child and adolescent psychiatry (~75%).[20]

There is a great deal of public and professional interest in the effects of aging on various aspects of physician competence, and future research could include tracking the performance of individual diplomates over time, as well as performance on specific items and topics.

References

1. Norman G: Research in clinical reasoning: past history and current trends. Med Educ 39:418–427, 2005
2. Tamblyn R, Abrahamowicz M, Brailovsky C, *et al*: Association between licensing examination scores and resource use and quality of care in primary care practice. JAMA 280:989–996, 1998
3. Tamblyn R, Abrahamowicz M, Dauphinee WD, *et al*: Association between licensure examination scores and practice in primary care. JAMA 288:3019–3026, 2002
4. Tamblyn R, Abrahamowicz M, Dauphinee WD, *et al*: Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. JAMA 298:993–1001, 2007
5. Wenghofer E, Klass D, Abrahamowicz M, *et al*: Doctor scores on national qualifying examinations predict quality of care in future practice. Med Educ 43:1166–1173, 2009
6. Holmboe ES, Wang Y, Meehan TP, et al: Association between maintenance of certification examination scores and quality of care for Medicare beneficiaries. Arch Intern Med 168:1396–1403, 2008
7. Hess BJ, Weng W, Holmboe ES, *et al*: The association between physicians' cognitive skills and quality of diabetes care. Acad Med 87:157–163, 2012
8. Choudhry NK, Fletcher RH, Soumerai SB: Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med 142:260–273, 2005
9. Cruft GE, Humphreys JW Jr, Hermann RE, *et al*: Recertification in surgery, 1980. Arch Surg 116:1093–1096, 1981
10. Meskauskas JA, Webster GD: The American Board of Internal Medicine recertification examination: process and results. Ann Intern Med 82:577–581, 1975
11. Norcini JJ, Lipner RS, Benson JA Jr, *et al*: An analysis of the knowledge base of practicing internists as measured by the 1980 recertification examination. Ann Intern Med 102:385–389, 1985
12. Ramsey PG, Carline JD, Inui TS, *et al*: Changes over time in the knowledge base of practicing internists. JAMA 266:1103–1107, 1991
13. Leigh TM, Young PR, Haley JV: Performances of family practice diplomates on successive mandatory recertification examinations. Acad Med 68:912–918, 1993
14. Rhodes RS, Biester TW: Certification and maintenance of certification in surgery. Surg Clin North Am 87:825–836, 2007
15. Lipner R, Song H, Biester T, *et al*: Factors that influence general internists' and surgeons' performance on maintenance of certification exams. Acad Med 86:53–58, 2011
16. Day SC, Norcini JJ, Webster GD, *et al*: The effect of changes in medical knowledge on examination performance at the time of recertification. Res Med Educ 27:139–144, 1988
17. O'Neill TR, Puffer JC: Maintenance of certification and its association with the clinical knowledge of family physicians. Acad Med 88:780–787, 2013
18. American Academy of Psychiatry and the Law: What is forensic psychiatry? Available at http://www.aapl.org/org.htm. Accessed March 19, 2013
19. Zonana H: Forensic psychiatry, in The American Board of Psychiatry and Neurology: Looking Back and Moving Ahead. Edited by Aminoff MJ, Faulkner LR. Arlington, VA: American Psychiatric Publishing, 2012, pp 151–159
20. Faulkner LR, Juul D, Andrade NN, *et al*: Recent trends in American Board of Psychiatry and Neurology psychiatric subspecialties. Acad Psychiatry 35:35–39, 2011
21. Accreditation Council for Graduate Medical Education: data on number of accredited programs for the current academic year (2012–13), United States. Available at http://www.acgme.org. Accessed December 5, 2013