ORIGINAL RESEARCH

Development and validation of a written credentialing examination for overseas educated dietitians

Running head: Credentialing exam for overseas trained dietitians

Date accepted: November 2017

Neville G. CHIAVAROLI,¹ Eleanor J. BECK,² Catherine ITSIOPOULOS,³ Paul WILKINSON,⁴ Kay GIBBONS⁵ and Claire PALERMO⁶

¹*Department of Medical Education, University of Melbourne, Melbourne, VIC, Australia*

²*University of Wollongong, Wollongong, NSW, Australia*

³*La Trobe University, Melbourne, Victoria, Australia*

⁴*Dietitians Association of Australia, Deakin, ACT, Australia*

⁵*Victoria University, Melbourne, VIC, Australia*

⁶*Monash University, Melbourne, VIC, Australia*

N. G. Chiavaroli, Med, Senior Lecturer

E. J. Beck, PhD, AdvAPD, Discipline Leader Nutrition and Dietetics

C. Itsiopoulos, PhD, APD, Head of School of Allied Health

P. Wilkinson, PhD APD, Executive Manager Accreditation, Recognition and Journal Services

K. Gibbons, BAppSc (Dietetics), FDAA, Professor Food & Nutrition

C. Palermo, PhD, AdvAPD, Associate Professor Department of Nutrition, Dietetics and Food

Correspondence: N. G. Chiavaroli, Department of Medical Education, Melbourne Medical School, University of Melbourne, VIC 3010, Australia. Tel: +61 3 8344 8238; Email: n.chiavaroli@unimelb.edu.au.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/1747-0080.12403](https://doi.org/10.1111/1747-0080.12403)

Abstract

Aim: Health professionals seeking employment in foreign countries are commonly required to undertake competency assessment in order to practice. This study aims to outline the development and validation of a written examination for Dietetic Skills Recognition (DSR), to assess the knowledge, skills, capabilities and professional judgement of overseas-educated dietitians against the competency standards applied to dietetic graduates in Australia.

Methods: This study reviews the design, rationale, validation and outcomes of a multiple-choice (MCQ) written examination for overseas-educated dietitians, based on five years of administration. The validity of the exam is evaluated using Messick's validity framework which focuses on five potential sources of validity evidence – content, internal structure, relationships with other variables, response process and consequences. The reference point for the exam cutscore is the minimum standard required for safe practice.

Results: One hundred and fourteen candidates have completed the MCQ examination at least once, with an overall pass rate of 52% on first attempt. Pass rates are higher from countries where dietetic education more closely reflects the Australian model. While the pass rate for each exam tends to vary with each cohort, the pass mark or 'cutscore' has remained relatively stable over the eight administrations outlined in the paper.

Conclusions: The findings provide important data supporting the validity of the MCQ exam. Fuller evaluation of the validity of the exam must be sought within the context of the whole DSR programme of assessment. The DSR written component may serve as a model for use of the MCQ format for dietetic and other professional credentialing organisations.

Key words: competency standards, credentialing, multiple-choice questions, professional competence, standard setting, validity.

Introduction

Credentialing is the processes by which health professional organisations verify the qualifications and experience of practitioners in order to determine their ability to provide safe, high quality health care services within a specific health care setting and role.¹ The credential to practise as a dietitian in Australia is issued and monitored by the Dietitians Association of Australia (DAA). DAA self-regulates the licensing or credentialing of dietitians, governed under the Accredited Practising Dietitian (APD) credential. Dietitians who are educated and qualified in other countries are required to undertake dietetic skills credentialing, managed by DAA on behalf of the Australian Government's Department of Immigration and Border Protection. This credentialing is a complex multi-stage process that includes assessment of knowledge, skills, capabilities and professional judgement aligned with the same competency standards² applied to dietetic graduates in Australia. The Dietetic Skills Recognition (DSR) process for overseas educated dietitians involves three stages: (i) assessment of qualifying degree and relevant experience, including currency of practice and

credentialing in country of origin (desktop review); (ii) completion of a written dietetic examination, and (iii) an oral counselling interview examination. Candidates are given three opportunities to pass each of the two examinations. Candidates who complete all three stages of the DSR process are then eligible to apply for membership of the DAA and the APD program. Passing the DSR written exam is a prerequisite for being able to attempt the oral counselling interview examination.

During 2009 and 2010, the DAA undertook an extensive evaluation of the process for credentialing overseas educated dietitians with particular focus on its validity, ability to assess scope of practice, and cost-efficiency.³ The evaluation determined that while the previous case-based short answer (SA) examination format was theoretically appropriate, there were concerns with elements of the examination process, including the degree of subjectivity in the marking of written responses, the limited scope of practice that could be assessed, and the financial burden of conducting a case-based examination system. In response, DAA considered the multiple choice question (MCQ) format as an alternative format for the DSR written examination.

The use of MCQs for professional credentialing is now relatively common,⁴ including in the health professional context, although some distrust about the format remains.⁵ Theoretically, at least, the MCQ has several advantages as an assessment format, one of the most widely acknowledged of which is its efficiency in terms of response; this not only facilitates the examination of large numbers of candidates, but also allows excellent assessment coverage of a professional domain.⁴ Such efficiency allows a relatively large number of items within the designated testing time compared with other testing formats.

Importantly, despite being commonly referred to as ‘objective tests’, MCQs are actually only objectively *scored*, rather than objectively developed. Just as much thought, discussion and debate must go into their development as for any other assessment instrument used for high stakes purposes.^{4,6} Research papers continue to emerge from various health professional contexts presenting data which support the validity of using MCQs for credentialing purposes, and such validity studies are important in helping shore up public confidence in the decisions made on the basis of the results of MCQ examinations.⁷⁻¹²

Ideas about validity have also evolved, from being initially regarded as an inherent property of a test to an *argument* for the particular interpretations and decisions based on the results of the test.¹³ In earlier conceptualisations, the validity of a test tended to be sought in specific and often isolated claims about the construct, content and criterion validity,¹⁴ with emphasis frequently given to the latter (as ‘predictive validity’) in relation to subsequent real-life performance. More recent approaches to validation of tests regard validity as a unitary phenomenon and classify the *types* of validity evidence, rather than distinguishing between different forms of validity.¹³ Accordingly, decisions made about credentialing test results are more appropriately thought of as inferences requiring supporting evidence, rather than unproblematic measures of professional competence.¹³

Two contemporary frameworks, developed respectively by Messick¹⁵ and Kane,^{16,17} reflect this new thinking about validity. The former framework outlines five potential sources of validity evidence (content, internal structure, relationships with other variables, response process and consequences) while the latter focuses more explicitly on the type of validity inference, namely, scoring, generalisation, extrapolation and implications. One of the effects

of reconceptualising validity in these ways is to play down the emphasis formerly invested in predictive validity. While some relationship between achievement on an examination and subsequent performance in a work context is clearly desirable, credentialing exams are not primarily intended to predict individual performance in a specific job, but rather aim to determine whether a candidate has acquired the (minimum) knowledge and understanding necessary for competent performance,¹⁸ at least at the level represented by the pass mark.

It follows that validity can never be ‘proven’ as such,¹⁷ but rather consists of a deliberate and continuous process of evidence gathering and justification for use and application in a given context. Ultimately, validity discussions help explain how credentialing examinations reflect a profession’s conception of competence.¹⁹ This issue is fundamental to all credentialing approaches (not only MCQs), and so the content and processes of credentialing must be appropriately aligned with the scope of practice and standards of a profession.

This study aimed to describe the development and validity of the MCQ examination administered in Australia to assess overseas educated dietitians for readiness to practice in Australia.

Methods

The test development process, from the competency standards through to results, analysis and candidate communication is represented in the flowchart in Appendix 1. We adopted Messick’s approach for organising and reporting on the development of the DSR MCQ

examination and the analysis of results. Ethics approval for this validation study was Approved by Ethics Committee University of Wollongong Protocol Number: 2017/219.

The content assessed by the DSR MCQ examination was determined by reference to the National Competency Standards.² The MCQ exam blueprint is based on these Standards and focuses on three common areas of dietetic practice – Food Service and Dietetics Management (FSM), Community and Public Health Nutrition (CPH), and Individual Case Management (ICM). The MCQs were developed to sample widely the knowledge, application and clinical reasoning required of dietitians in these areas, equally represented on the exam. It should be noted that while the DSR examination was originally based on the 2009 competency standards, a recent review of the content against the updated 2015 standards determined that few changes were necessary as questions already encompassed areas of professionalism, communication, evidence based practice, improving nutrition outcomes and collaboration critical to the new standards.²

Question writers were recruited from subject matter experts within each practice area. Expressions of interest were sought from APDs with the appropriate dietetic expertise, as well as experience in education and assessment of entry-level practitioners. Twelve APDs were initially selected across the three areas, and experienced APDs were chosen to lead each area and to form part of the Examination Committee (EC). An assessment expert and the DAA Executive Manager for Accreditation also formed part of the EC. Training in this area was conducted by one of the authors (blinded) via a two-day face-to-face workshop and was structured around understanding and implementing key item writing principles, and the need for and approach to peer review of questions.

This process produced an initial pool of 300 items. All drafted questions were reviewed by fellow item writers in each area, as well as the area leads, to confirm the content and question accuracy, and then further reviewed by the EC to confirm accuracy of content, alignment with the competency standards, and appropriately targeting of generalist entry-level dietitians. Ambiguous, contentious, inaccurate or unreasonably difficult questions were similarly identified by the EC and removed from the question pool.

The response format was 120 MCQs to be completed in 150 minutes, comprising 40 questions for each area of practice. The MCQ item formats adopted were single best answer and the extended matching format, as recommended by current handbooks of test development.²⁰⁻²¹ In line with recommendations for professional competency assessments, the majority of items included a 'stem' (also known as a 'scenario' or 'vignette') with relevant clinical or contextual information, which candidates needed to read and understand in order to respond meaningfully to the questions.⁴ The use of such a stem enabled the format to assess higher levels of cognition beyond simply recall of isolated facts, such as application of knowledge in different contexts and more integrated clinical reasoning.²²⁻²⁴ The items were written so that the key should be identifiable as the best or most defensible option by knowledgeable candidates. The incorrect options, referred to as 'distractors', were designed to present plausible alternatives in the context of the material and question, but would be able to be identified as incorrect by candidates with sufficient knowledge and understanding. To further confirm the accuracy and suitability of questions in terms of assessed level of competence and overall technical quality, questions were piloted with final year Australian dietetics students from five Australian universities. Two-hundred and forty questions were

selected and administered to the volunteer students under standardised conditions (timed and invigilated). Two hundred and ten students sat the pilot exam. Data from this pilot was analysed (according to the examination processes and performance indicators described further below) and the vast majority of questions (in excess of 90%) were deemed suitable to form part of the initial exam question bank for assessment through the official DSR assessment process.

The final selection of 120 questions for each examination was assembled by DAA professional staff. Examinations were administered in approved computer exam laboratories twice annually under strict invigilation and standardised conditions. Different versions of the examination were developed for each administration, with some common items to maintain appropriate blueprinting and to provide comparative data between administrations. Sample questions were made available to candidates via the DAA website.²⁵

Following each administration, candidate scores and item statistics were reviewed to ascertain the internal structure of the exam and analyse relationships with other variables. Evidence relating to the internal structure was particularly sought in the reliability index and in item quality indicators, namely the item facility (percentage of candidates correctly answering each question) and the discrimination index (a measure of the consistency of performance on an individual item with respect to the candidates' overall performance).²⁷⁻²⁸ A wide spread of item facilities is both expected, given the broad blueprint and different complexity of clinical scenarios presented, and desirable, as it helps contribute to the discrimination of the test.²⁷ It is important to note that differences in the overall difficulty of tests do not necessarily undermine the fairness of the different exam forms, due to the

standardised standard setting procedure which was applied to ensure an equitable cutscore for each exam (described below). The discrimination index (DI) provides information about the validity of a question in terms of its relationship with other questions and the examination as a whole, and is commonly calculated as a point-biserial correlation,²⁸ which is the method adopted for this study. Good quality items which align with the overall blueprint and appropriate domain are expected to have a positive DI; the threshold for acceptability is commonly stated as 0.20, although slightly lower values are acceptable for tests which assess different areas.²⁸ In addition, distractor analysis, an inspection of candidate selection patterns on the distractors (incorrect options), is useful to help confirm that the distractors are inferior to the designated key. In good quality items, the key should have a higher discrimination value than the distractors, indicating that the high-performing candidates are less likely to choose a distractor than lower performing candidates overall.²⁹

In accordance with our validity framework, we also sought to investigate the association between performance on the DSR MCQ exam and other factors relevant to dietetic competence. Data on subsequent performance of successful candidates in the DSR process was not available to the EC, nor, as noted above, was the exam designed to necessarily predict future performance. However, a key variable in performance, the country and language of dietetic training of applicants, was available through the desktop review; therefore, results were analysed according to regions by language and frequency.

The consequences of any test, according to Messick's validity framework, refers to its impact, beneficial or harmful, intended or unintended.¹⁷ One of the most important of the intended consequences is to set and confirm a defensible passing standard, that reflects fairly

and accurately the knowledge and understanding required by entry-level dietitians in Australia, including overseas-educated. To achieve this, after each test administration, the EC convened to determine the level of performance which represented the minimum acceptable achievement on the exam, a process known in testing theory as 'standard setting'. Traditionally, universities and other educational institutions have relied on a fixed and pre-determined passing score; however, this does not account for potential variations in the overall difficulty of examinations, which obviously affects how many test-takers will pass. Accordingly, most credentialing organisations now set passing scores systematically and empirically using one of many different methods as a way of gathering and comparing expert judgements about acceptable performance on a particular test, and expressing that level of performance as a single score on a test, known as the 'cutscore'.³⁰

For the DSR MCQ examination, a simplified version of the popular Angoff approach was used, commonly known as the 'Yes/No' variant.³¹⁻³² In this method, subject matter experts estimate the likelihood (in terms of 'likely' vs 'not likely') of a correct response from a hypothesised borderline or minimally competent candidate, for each item on the exam, and their collective judgements are converted into an overall percentage score based on the number of 'yes' decisions, which represents the minimum acceptable level of performance. Candidate scores are deemed acceptable if they meet or exceed the cutscore determined in this way. Applying such a standard setting approach means that exams which differ in terms of overall difficulty can be assigned comparable cutscores representing an equivalent level of competence; as a result, candidates have similar chances of success on each exam, regardless of the actual selection of items. Decisions made for any single item are carried over for

subsequent examinations, unless changes to the question (or dietetic practice) require the standard to be re-set.

Following the determination of the cutscore for each test, candidates were notified of their result and consequent eligibility to attempt the oral exam within six weeks of administration. Successful candidates received notification only of their overall Pass result, while unsuccessful candidates who failed based on substandard performance in a particular area were alerted to their poor performance in that area. Feedback was also provided in the form of reference to the National Competency Standards based on their performance. Unsuccessful candidates could resit the MCQ examination on another two occasions.

Further validity evidence was obtained through calculation of a correlation coefficient (r-value) to measure the association in performance between candidate mean performance, the required cutscore to pass the MCQ examination, and the resulting candidate passing rate.

Results

The DSR MCQ exam was first administered in 2012. Up to the time of writing (August 2016), there have been 8 administrations at six-month intervals (Table 1). Candidate numbers have been relatively stable over the eight administrations to date. One hundred and five overseas-educated candidates and nine Australian-educated candidates have completed the examination at least once. (Candidates from Australia were required to complete the examination due to limited currency of practice or disciplinary action, in order to be re-admitted into the APD program). Approximately half (52%) of all candidates passed on first

attempt. Twenty of the first-time fails subsequently passed on second attempt, with two candidates passing on third attempt; three candidates have failed all three attempts (Table 1).

Table 1 Descriptive statistics of all administrations across eight examination rounds and all candidates.^a

Administration Exam Number	1	2	3	4	5	6	7	8
Date of administration	Sep-12	Mar-13	Sep-13	Mar-14	Sep-14	Mar-15	Sep-15	Mar-16
No of questions	120	120	120	120	120	120	120	120
No of candidates ^a	20	18	23	16	10	10	21	19
No of repeat candidates	N/A	1	5	6	4	3	5	6
Mean candidate score (%)	67.2	63.8	70.3	66.7	65.3	70.8	66.9	71.3
Maximum total score (%)	76.7	72.5	80.0	73.3	73.3	81.7	79.2	85.0
Minimum total score (%)	52.5	41.7	51.7	52.5	47.5	60.0	47.5	50.8
Mean item facility (%)	66.2	64.3	70.9	67.0	63.6	70.8	70.6	74.8
Mean discrimination index	0.19	0.20	0.19	0.14	0.18	0.16	0.25	0.21
Reliability index ^b	0.72	0.81	0.77	0.63	0.77	0.69	0.85	0.80
Cutscore (%)	67	68	69	69	72	72	71	75
Pass rate (%)	70	39	61	56	30	63	52	68

^a There are 114 unique candidates but this number includes candidates sitting repeat examinations (to a maximum of 3 attempts).

^b Internal consistency calculated using the Quest programme²⁶

Despite different test forms and a majority of first time candidates for each administration, the overall difficulty of the examination has remained relatively stable, with the mean candidate score between 64% and 71% (Table 1 and Figure 1).

Insert Figure 1 about here

There has been a variable relationship between candidate performance and cutscore ($r = 0.46$) (Figure 1), although the most recent increase in the cutscore was accompanied by a similar increase in candidate pass rate.

The reliability of the test was above 0.70 for most administrations, with lower figures recorded in March 2014 and March 2015. These lower values paralleled those of the mean DI, suggesting that the items on those test forms were less coherent or homogeneous than the other forms. These findings and how the EC dealt with them are discussed below.

The majority of candidates who trained in countries with English as the primary language of instruction passed on first attempt, and achieved higher overall scores (and pass rates) compared to candidates from non-English speaking backgrounds (NESB) (Table 2). Of these candidates, only two passed on first attempt, six on second attempt and two on the third (and final) attempt.

Table 2 Demographics and pass rates of candidates by region of training.

Region of dietetics training ^a	No of Candidates	Total Passes (up to 3 attempts)	Overall pass rate (%)
English as first language of instruction ^b (overseas-trained)	73	60	82.2
Australia ^c	9	5	55.6
Europe (excluding UK/Ireland)	6	3	50.0
Asia/Subcontinent	14	4	28.6
South America	6	3	50.0
Middle East	6	4	66.7

^a Note that individual country data is not reported due to the low numbers from some countries which may make some candidates identifiable.

^b Canada/USA/Ireland/South Africa/United Kingdom/New Zealand

^c Australian candidates sitting for return to practice

Figure 2 shows that while the cutscore has remained relatively stable over the eight administrations, the pass rate for each exam has varied considerably, pointing to a cohort effect. The independence of these two indicators is confirmed by the correlation coefficient of 0.03.

Insert Figure 2 about here

Discussion

The methods and results reported above show that the DSR MCQ exam has been successfully developed, piloted and administered over the period 2012–2016. The development processes and the data collected through administration of the DSR examinations provide important validity evidence consistent with Messick's framework focussing on test content, internal structure, relationship with other variables, response process and consequences.

An initial question bank of 300 items across three key content areas of dietetic practice was developed and judged by an expert panel to reflect appropriate entry-level dietetic competence (with relevant updating and additional writing over time not reported here). Candidate scores and item performance data are consistent with the purpose of the exam to discriminate between different levels of knowledge by effectively assessing the foundational content and application of knowledge of candidates. The growing bank of MCQ items enables wide sampling of the areas of dietetic knowledge and the stem-based response format allows assessment of higher order clinical reasoning. Individual item analysis points to the

majority of questions being of acceptable quality, while also highlighting any questions on the examination which need to be rescored or removed from scoring. New questions for the DSR MCQ examination continue to be developed and reviewed. Because of the impracticality and potential security implications of using local dietetics students for ongoing piloting of each item, this is achieved through the formal examination itself, where a small number of new items are included in each administration and evaluated for quality, and accepted for scoring only if shown to be psychometrically sound. In this way, a sustainable bank of suitable items is continuously developed and benchmarked against the appropriate level of candidature.

In terms of internal structure, the reliability of the examination, shows some variability across administrations. Given the small candidate numbers for some administrations, demonstrated variability of candidate knowledge, and the multi-stage nature of the assessment process, the EC determined that a coefficient alpha index of 0.70 or above was acceptable for the purposes and context of the DSR process, in line with many psychometric guidelines.³³ Although higher reliability indices (eg 0.90 or above) are sometimes advocated,^{27,34} this is debatable when test content may be intentionally heterogeneous for blueprint purposes, as in the DSR MCQ exam, since very high reliabilities may in fact indicate excessive redundancy in test content.³³ However, the low reliability of the March 2014 exam prompted particular review by the EC, and was attributed to a selection of items which had slightly lower discrimination indices in comparison with previous years, as evidenced by the low mean DI for that administration. In order to offset the greater error in measurement indicated by the lower reliability, the EC lowered the cutscore for that

administration by one standard error of measurement (SEM), in accordance with common psychometric practice.³⁵ Nevertheless, the lowered cutscore was still well above the next highest score, and therefore resulted in no change to the pass rate. Subsequent item selection has since taken into consideration item quality indicators more explicitly alongside content requirements, as borne out by subsequent mean discrimination indices and reliability indices.

There appears to be a clear association between performance on the DSR MCQ examination and English language competence, although this is also likely to be influenced by the quality and scope of dietetic education and training available in the country of origin. It would be expected that candidates from countries whose dietetics curricula are more aligned to those in Australia will, in general, perform better in the exam, especially where the language of instruction is also in English. This would seem to explain the lower rate of success for certain groups of overseas-educated dietitians, particularly for candidates who have worked in limited patient care settings in their countries of origin. While candidates have a range of resources available to assist them to prepare for the exam via the DAA website, pass rates will inevitably reflect the nature of contemporary dietetics practice as defined by the Australian competency standards. This commitment to improving nutrition across multiple practice settings in Australia is a strength of the preparation of dietitians and continues to be recognised in a recent review of those standards.² Ongoing review of the relevance of the MCQ items generally, but particularly in relation to new competency standards introduced in 2015,² is a current priority of the EC.

The main result (or ‘consequence’ in Messick’s terms) of the DSR MCQ examination is the setting of an appropriate standard for the minimum knowledge and understanding

necessary to practice as a dietitian in Australia. The cutscore represented by this standard must necessarily vary in accordance with the facilities of the individual items which make up any version of the exam. However, the data shows that the cutscore has remained relatively stable over the eight administrations. The same cannot be said for the pass rate for each exam, which has varied considerably across administrations. This variation in pass rate for each administration is not unexpected, given the difference in candidates who may sit a particular administration. In the context of the relative stability of the cutscores and the standard setting process, the variability in pass rate is most likely explained by candidate variability, rather than differences in overall exam difficulty. The mean facility of the items on each administration has been relatively stable, and the mean candidate score, related to the overall difficulty of the examination, shows similar stability. The current cumulative pass rate of 52% on first attempt reflects a challenging examination which attempts to strike an appropriate balance between protecting the public and professional standards, and fairness to qualified candidates.

While candidates who fail to pass the MCQ examination are provided with feedback regarding their performance and advised to seek mentoring, there are clearly several negative consequences associated with failure in a credentialing exam, including disappointment, frustration and anxiety, as well as the financial cost of having to re-attempt the exam. Such consequences are arguably an inherent part of the credentialing exam process, but DAA attempts to mitigate these to some extent through the possibility of re-attempts, the feedback outlined above, and a formal appeal process for unsuccessful candidates. Nonetheless, the

need to better support and guide prospective candidates through the credentialing process is highlighted in related literature, and acknowledged here.³⁶

It is important to reiterate that the written MCQ examination is only one part of the DSR process of determining competence. While predicting subsequent performance of successful candidates is not a primary purpose of credentialing exams, the authors nevertheless acknowledge that a more complete validation of the DSR MCQ examination would include consideration of workplace impacts of decisions based on the exam performance, including success rates of APD applications. Given the examination has now been in use for over five years, such data collection and analysis may now be feasible.

This study aimed to describe the development and validation of an MCQ examination administered in Australia to assess overseas educated dietitians for readiness to practice in Australia. The findings show that the 120 MCQ examination has thus far shown acceptable reliability and validity across eight separate administrations, and provides a suitable basis for decisions about candidates' competence. The examination is an important milestone in the overall assessment for overseas-educated dietitians or those seeking to return to practice in Australia, and may serve as a model for other professions and the dietetic profession elsewhere.

Funding source

No funding was received for this research.

Conflict of interest

Dr. Wilkinson is a member of the Dietitians Association of Australia staff. The remaining authors have no conflicts of interest to declare.

Authorship

All authors were involved in the conception of the study and contributed to the assessment process which generated the data contained in the study. NC was responsible for the assessment design and data analysis, EB, CP and KG led the development of items and standard setting in their respective domains, and CI and PW oversaw the assessment administration and data collection. The manuscript was written by NC, EB and CP with input from PW, CI and KG. All authors read, edited and approved the final version. The content has not been published elsewhere.

References

- 1 Australian Commission on Safety and Quality in Health Care. Credentialling of clinicians; 2017. (Available from: <https://www.safetyandquality.gov.au/our-work/credentialling>, accessed 15 September 2017).
- 2 Dietitians Association of Australia (DAA). National competency standards for entry-level dietitians, 2015. (Available from: <https://daa.asn.au/wp-content/uploads/2017/01/NCS-Dietitians-Australia-1.0.pdf>, accessed 15 September 2017).
- 3 Dietitians Association of Australia (DAA). Developing a New DAA Dietetic Skills Recognition Examination System – Phase II - Question and Exam Validation. Final report, 2010. (Available from: <http://daa.asn.au/wp-content/uploads/2016/11/Final-Report-October-2010-002.pdf>, accessed 5 July 2017).
- 4 van Der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. *Adv Health Sci Educ Theory Pract* 1996; **1**: 41-67.
- 5 Veloski JJ, Rabinowitz HK, Robeson MR, Young PR. Patients Don't Present With Five Choices: An Alternative to Multiple-Choice Tests in Assessing Physicians' Competence. *Acad Med* 1999; **74**: 539-546.

- 6 Brown S, Knight P. *Assessing Learners in Higher Education*. London: Taylor & Francis, 2012.
- 7 Thellesen L, Bergholt T, Hedegaard M *et al*. Development of a written assessment for a national interprofessional cardiocography education program. *BMC Medical Education* 2017; **17**: 88.
- 8 Bissell V, Chamberlain S, Davenport E, Dawson L, Jenkins S, Murphy R. The Overseas Registration Examination of the General Dental Council. *Br Dent J* 2016; **221**: 257-61.
- 9 Norcini JJ, Boulet JR, Opalek A, Dauphinee WD. The Relationship Between Licensing Examination Performance and the Outcomes of Care by International Medical School Graduates. *Acad Med* 2014; **89**: 1157-62
- 10 Keough VA, Stevenson A, Martinovich Z, Young R, Tanabe P. Nurse practitioner certification and practice settings: Implications for education and practice. *J Nurs Scholarsh* 2011; **43**: 202.
- 11 Newton DW. The NAPLEX: Evolution, purpose, scope, and educational implications. *Am J Pharm Educ* 2008; **72**: 33.
- 12 Bowhay AR, Watmough SD. An evaluation of the performance in the UK Royal College of Anaesthetists primary examination by UK medical school and gender. *BMC Medical Education* 2009; **9**: 38.
- 13 Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ* 2015; **49**: 560-575.
- 14 Artino ARJ, Durning SJ, Creel AH. AM Last Page: Reliability and Validity in Educational Measurement. *Acad Med* 2010; **85**: 1545.
- 15 Messick S. Validity. In: Linn RL, editor. *Educational measurement*. 3rd ed. New York: American Council on Education, 1989; 13-103.
- 16 Kane MT. Validation. In: Brennan RL, editor. *Educational measurement*. 4th ed. Westport: Praeger, 2006; 17-64.
- 17 Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Advances in Simulation*. 2016; **1**: 31.
- 18 American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association; 2014
- 19 D'Costa AG. The Validity of Credentialing Examinations. *Eval Health Prof* 1986; **9**: 137-169.
- 20 Haladyna TM, Rodriguez MC. *Developing and Validating Test Items*. London: Taylor & Francis, 2013.
- 21 Case SM, Swanson, DB. *Constructing Written Test Questions for the Basic and Clinical Sciences*. 3rd Ed (rev.). Philadelphia, PA: National Board of Medical Examiners, 2002.
- 22 Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? *BMC Medical Education* 2007; **7**: 49.

- 23 Tractenberg RE, Gushta MM, Mulroney SE, Weissinger PA. Multiple choice questions can be designed or revised to challenge learners' critical thinking. *Adv Health Sci Educ Theory Pract* 2013; **18**: 945-61.
- 24 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; **65**: S63-7.
- 25 Dietitians Association of Australia (DAA). Overseas Educated Dietitians – Examination Stage, 2016. (Available from <https://daa.asn.au/becoming-a-dietitian-in-australia/recognition-of-dietetic-qualifications/overseas-educated-dietitians-assessment-stage/>, accessed 15 September 2017).
- 26 Adams RJ, Khoo S. *Quest: The Interactive Test Analysis System* (computer programme, version 2.1), Melbourne: ACER Press, 1996.
- 27 McAlpine M. *A Summary of Methods of Item Analysis*. CAA Centre, 2002. (Available from: <http://caacentre.lboro.ac.uk/dldocs/BP2final.pdf>, accessed 15 September 2017).
- 28 Chiavaroli N, Familiari M. When Majority Doesn't Rule: The Use of Discrimination Indices to Improve the Quality of MCQs. *Bioscience Education* 2011; **17**: 1-7.
- 29 Tavakol M, Dennick R. Postexamination Analysis: A Means of Improving the Exam Cycle. *Acad Med* 2016; **91**: 1324.
- 30 McKinley DW, Norcini JJ. How to set standards on performance-based examinations: AMEE Guide No. 85. *Med Teach* 2014; **36**: 97-110.
- 31 Impara JC, Plake BS. Standard Setting: An Alternative Approach. *J Educ Meas* 1997; **34**: 14.
- 32 Yudkowsky R, Downing SM, Wirth S. Simpler standards for local performance examinations: the Yes/No Angoff and whole-test Ebel. *Teach Learn Med* 2008; **20**: 212-217.
- 33 Tavakol M, Dennick R. Post-examination analysis of objective tests. *Medl Teach* 2011; **33**: 447-58
- 34 Downing SM, Yudkowsky R. *Assessment in Health Professions Education*. New York: Routledge, 2009.
- 35 Hays R, Gupta TS, Veitch J. The practical value of the standard error of measurement in borderline pass/fail decisions. *Med Educ* 2008; **42**: 810-815.
- 36 Balasubramanian M, Brennan DS, Spencer AJ, Watkins K, Short SD. Overseas-qualified dentists' experiences and perceptions of the Australian Dental Council assessment and examination process: the importance of support structures. *Aust Health Rev* 2014; **38**: 412-9.

NCS = National Competency Standards

*Examination committee made up of Dietitians Association of Australia Accreditation and Recognition Manager & Recognition and Journal Services Dietitian, three item writers and MCQ expert academic consultant.

Figure 1.png

Figure 2.png

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Chiavaroli, NG;Beck, EJ;Itsiopoulos, C;Wilkinson, P;Gibbons, K;Palermo, C

Title:

Development and validation of a written credentialing examination for overseas-educated dietitians

Date:

2018-04-01

Citation:

Chiavaroli, N. G., Beck, E. J., Itsiopoulos, C., Wilkinson, P., Gibbons, K. & Palermo, C. (2018). Development and validation of a written credentialing examination for overseas-educated dietitians. NUTRITION & DIETETICS, 75 (2), pp.235-243. <https://doi.org/10.1111/1747-0080.12403>.

Persistent Link:

<http://hdl.handle.net/11343/283490>