Protocol:

Assessing the behavioural and social science curricula components for undergraduate medical students: A BEME Systematic Review

BEME Identification number:

Version 3.1

10th January 2016

Author: Dr E J Hothersall

Lead reviewer:

Dr E Hothersall, University of Dundee e.hothersall@dundee.ac.uk, 01382 383 812

E J Hothersall, MBChB, MPH, MD; Honorary Senior Lecturer in Public Health, Consultant in Public Health Medicine and Systems in Practice Convenor. Dr Hothersall is teaching lead for Population Health and Evidence Based Medicine at the University of Dundee medical school. Her interests are in teaching and assessment of population health topics. She has previously conducted a systematic review of the effects of obesity on elective surgery. She will serve as the PI and lead all aspects of this review.

Other reviewers:

Jeni Harden (University of Edinburgh) has extensive experience in undergraduate teaching and assessment in social sciences. She is also an experienced qualitative researcher, and has previously published systematic reviews.

Evridiki Fioratou (University of Dundee) is the lead for undergraduate teaching and assessment of Behavioural and Social Sciences including Human Factors at University of Dundee medical school. She also has previous involvement in a

Health Technology Assessement systematic review of the management of obesity in men.

Annalisa Manca (University of Dundee) is a medical education technologist, specialising in the use of social media in medical education. Her background is in social sciences and humanities, which ground her skills in the application of theory in teaching and research practices. She also holds an MEd in Communication, Education Technologies.

Morris Gordon (University of Central Lancashire) is active in Cochrane systematic review, with several published and in progress projects. He has published a number of medical education systematic reviews and is the lead for the UCLAN BEME BICC, on the BEME Editorial board and lead for BEME educational developments.

Susie Schofield (University of Dundee) has is a senior lecturer with extensive experience of teaching and assessment in medical education. She is responsible for ongoing research into assessment, staff development and flexible learning.

Assistance is provided by Scott Mcgregor (University of Dundee), subject librarian.

Abstract

Behavioural and social science (BSS) topics such as anthropology, psychology and sociology are crucial for modern medical practice, and in the UK are enshrined in the core curricula of undergraduate and post graduate medical teaching. Medical school assessment, including BSS topics, is increasingly tending towards a multiple choice format. The "best" assessment of BSS topics would be one that was valid, reliable, generalisable, feasible and fair, with demonstrable educational impact, but evidence is that such assessments are used is lacking. There are no reviews of assessment that relate to BSS subjects.

The research questions for this review will be:

- What methods are used to assess BSS curricula components within undergraduate medical education? (Descriptive)
- How and why are these methods selected? (Clarification)
- What are the validity, reliability, generalizability, feasibility and fairness of these methods for assessment of BSS topics in medical education? (Justification).

Searching and data extraction will be done by two researchers, with 10% checked for agreement. The theoretical framework for analysis has not been determined in advance as it will be shaped by the papers that are found.

1.Background to the topic

Behavioural and social science (BSS) curricula components such as anthropology, psychology and sociology are crucial for modern medical practice, and enshrined in the core curricula of undergraduate[1] and post graduate[2,3] medical teaching. However, the underlying ontological alignment of these sciences can be considered a stark contrast to other areas of medical education. Biomedical science is aligned with realism, whilst conversely BSS topics are aligned with a relativist viewpoint. As relativism accepts multiple realities, interpretations and truths, this presents a particular challenge for the development of appropriate assessment tools [4]. In the United Kingdom, the national guidance for the undergraduate medical curriculum, Tomorrow's Doctors, does not specify modes of assessment but notes that 'assessments will be fit for purpose – that is: valid, reliable, generalisable, feasible and fair' [1, p. 48]. Anecdotal evidence suggests that medical school assessment, including BSS topics, is increasingly tending towards a multiple choice format [5,6]. However, such methods, suggesting the best fit of a single 'correct' stem, are not naturally aligned with a relativist paradigm. As such, it may be other forms of assessment that allow students to reflect their understanding of the various and competing underpinning theories and frameworks relevant to a problem are employed.

A comparable topic with similar difficulties is the assessment of professionalism, where summaries are mostly restricted to the descriptive level, including feasibility, without any exploration of validity, reliability, generalizability, or fairness [7,8]. There are no comparable reviews that relate to BSS curricula components.

It is clear that students and faculty use assessment as a proxy for "value" within a curriculum, spending more time and effort on topics which are summatively assessed [9]. Without evidence of robust and reliable assessment of BSS topics, there is a risk that they are overlooked in summative assessment in favour of topics which are judged to be more reliably assessed [10], which in turn perpetuates the diminished role and status of BSS topics. Similarly, where BSS topics might be included in formative assessment, without evidence of

educational impact of formative assessment, this too risks being sidelined by both students and tutors [11].

2.Review topic/question(s), objectives and key words

Question to be investigated

The research questions for this review will be:-

- What methods are used to assess BSS curricula components within undergraduate medical education? (Descriptive)
- How and why are these methods selected? (Clarification)
- What are the validity, reliability, generalizability, feasibility and fairness of these methods for assessment of BSS topics in medical education? (Justification).

Theoretical approach to the review

No single research paradigm will be used to underpin the approach for this review. The authors will embrace both positivism (through justification of assessment methods used) and constructivism (through consideration of underpinning theoretical frameworks that inform assessment choices)

An early stage of the review process will entail determining which forms of validity are described within the relevant literature, enabling identification of the theoretical frameworks used in the analysed studies. Only after this will an approach to describing the findings be able to be agreed. Thus, authors may discuss criterion, content and/or construct validity, each with different theoretical underpinnings.

Aim

To systematically review and synthesize the research on assessment of BSS curricula components within undergraduate medical education.

Key words

assessment, medical education, health professions education, population health, public health, behaviour change, health behaviour, behavioural science, social

science, psychosocial, sociology, anthropology, sociology, psychology, validity, psychometrics

Search sources and strategies

ERIC, Medline, Psychinfo, Pubmed, Google scholar, EMBASE, CINAHL, British Education Index, Applied social sciences index and abstracts, CERUK, Education Abstracts. Hand searching of listed references for full papers will also be performed by the study investigators. We will have assistance in search strategies from a subject expert librarian

4. Study selection criteria

4.1 Inclusion and exclusion criteria

We will use the above search sources and will use both electronic and hand searching of relevant reference. The majority of papers found are likely to fall into the category of "ideographic descriptions" or "case reports", or "justification research" [12]. Inclusion and exclusion criteria are listed in Table 1.

Table 1. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Papers describing summative or formative assessment of any/all of: population health, anthropology, psychology or sociology	Papers describing teaching only
Undergraduate medical curricula (including graduate medical school programs in appropriate countries)	Papers giving opinions or reviews, but not describing the primary use of a particular assessment method for these topics

Elective or core components of curricula	Exclusive discussion of intercalated degree years, or other highly specialised student group
All languages	Exclusive discussion assessment of statistics or biostatistics, epidemiology or medical humanities
	Papers describing assessment of counseling skills
	Papers describing assessment at admission to medical school

4.2 Definition of "behavioural and social science" (BSS) curricula components

For the purposes of this review, the term "BSS curricula components" will be used to mean topics addressing the learning outcomes specified by Tomorrow's Doctors 2009 [1] (Table 2), and/or the CANMEDS 2015 Physician Competency Framework [3] (Table 3).

Table 2. Learning outcomes for behavioural and social science (BSS) topics from Tomorrow's Doctors 2009 [1]

Apply psychological principles, method and knowledge to medical practice

- Explain normal human behaviour at an individual level
- Discuss psychological concepts of health, illness and disease
- Apply theoretical frameworks of psychology to explain the varied responses of individuals, groups and societies to disease
- Explain psychological factors that contribute to illness, the course of the disease and the success of treatment
- Discuss psychological aspects of behavioural change and treatment compliance
- Discuss adaptation to major life changes, such as bereavement; comparing and contrasting the abnormal adjustments that might occur in these situations

- Identify appropriate strategies for managing patients with dependence issues and other demonstrations of self-harm

Apply social science principles, method and knowledge to medical practice

- Explain normal human behaviour at a societal level
- Discuss sociological concepts of health, illness and disease
- Apply theoretical frameworks of sociology to explain the varied responses of individuals, groups and societies to disease
- Explain sociological factors that contribute to illness, the course of the disease and the success of treatment − including issues relating to health inequalities, the links between occupation and health and the effects of poverty and affluence
- Discuss sociological aspects of behavioural change and treatment compliance

Apply to medical practice the principles, method and knowledge of population health and the improvement of health and healthcare

- Discuss basic principles of health improvement, including the wider determinants of health, health inequalities, health risks and disease surveillance
- Assess how health behaviours and outcomes are affected by the diversity of the patient population
- Describe measurement methods relevant to the improvement of clinical effectiveness and care
- Recognise the role of environmental and occupational hazards in ill-health and discuss ways to mitigate their effects
- Discuss the role of nutrition in health
- Discuss from a global perspective the determinants of health and disease and variations in healthcare delivery and medical practice

Table 3. BSS competencies from CANMEDS 2015 Physician Competency Framework [3]

Respond to an individual patient's health needs by advocating with the patient within and beyond the clinical environment

- Work with patients to address determinants of health that affect them and their access to needed health services or resources

- Work with patients and their families to increase opportunities to adopt healthy behaviours
- Incorporate disease prevention, health promotion, and health surveillance into interactions with individual patients

Respond to the needs of the communities or populations they serve by advocating with them for system-level change in a socially accountable manner

- Work with a community or population to identify the determinants of health that affect them
- Improve clinical practice by applying a process of continuous quality improvement to disease prevention, health promotion, and health surveillance activities
- Contribute to a process to improve health in the community or population they serve

5. Procedure for extracting data

We will use two searchers and two extractors throughout. The main review will include articles meeting the above quality criteria. We will use a rating scale for each element with conducting our review of articles.

Process:

1. Search for and obtain papers based on above strategy
2. Find and remove duplicates
3. Code on abstract (definite in, definite out, doubtful by at least two reviewers)
4. Member checking on doubtful abstracts by at least one additional reviewer
5. Full text papers meeting the criteria of "definite in" or "doubtful" will be reviewed using the above criteria by at least two reviewers. Full papers which are then deemed "definite out" will be excluded from data extraction
6. Papers will be rated for quality prior to data extraction

For all articles a rating scale for quality will be used, based on the consensus document on medical education assessment research [12]. The rating scales do not focus on the methodological quality of the primary research paper, as it is

not expected that traditional trials such as randomised controlled trials will be found. Instead, they focus on key items that will inform our research questions. Whilst low quality may reflect a poor primary study or poor reporting, as these two cannot be distinguished, the judgement will be made based on the information presented.

It is not proposed that this assessment will lead to the exclusion of studies, rather it will inform when making conclusions about the strength of evidence, as a lower quality paper may contain bias through missing information. The proposed elements are detailed in table 4.

Table 4. Quality indicators

Quality indicator	Good quality	Unclear quality	Low quality
Underpinning framework	Clear and relevant description of theoretical models or conceptual frameworks that underpin the choice of assessment	Some limited discussion of underpinning, with minimal interpretation in the context of the assessment choice	No mention of underpinning
Assessment method	Clear description of the process and outcomes of the assessment	Some limited description that will not facilitate replication	No mention of assessment method in any detail
Setting	Clear details of the educational context and learner characteristics of the study	Some description, but not significant as to support dissemination	No details of learner characteristics or setting
Psychometrics	Clear description of relevant psychometrics and	Some psychometric information, but not	No details of psychometrics

	how applied to this assessment	enough to fully inform for dissemination	
Content	Provision of detailed materials (or details of access), such as mark sheets, rubrics, etc to allow assessment replication	Some elements of materials presented or summary information	No assessment content presented
Conclusion	Conclusions of the study reflect the findings	Some mismatch between the conclusions and findings	No correlation between the findings and conclusions

The data extraction instrument will include standard PRISMA [13] recommendations for data extraction to include study characteristics (PICOS, funding, study characteristics, results, limitations, and conclusions).

Full papers will be divided amongst group members, coding according to outcomes, 20% cross checking or until sufficient agreement. Coding sheets (incorporating the above criteria elements) will be used.

Inter-rater reliability between authors will be established on 20% of articles and randomly checked for the entire sample, or all articles if <10 are found. Where discrepancies are found, the papers will be reviewed by a third reviewer, and discussed until agreement is met.

6. Synthesis of extracted evidence

The data extraction instrument will be pilot tested and refined on a small sample (10%) of the identified articles. The coders will discuss findings until agreement is met.

Once consensus is reached in the coding, a further 10% of articles will be reviewed and inter-rater reliability will be calculated. Following this process, and any needed revisions, the remainder of the articles will be double coded so that

additional inter-rater calculation can be performed. A narrative synthesis of the data will be built with the construct determined by the theoretical frameworks of the papers contained in the review [14]. Articles will also be categorized by Kirkpatrick category for outcomes, consistent with multiple prior BEME reviews.

Descriptive synthesis

A descriptive analysis will take place, summarising the data from the extraction form, to focus on describing summary data regarding: Study type, assessment type, psychometric details and outcomes of the primary study. Additionally, content related to the quality assessment indices will be extracted, including where relevant any additional content or appendices.

Meta-analysis

If suitably homogenous outcome data are presented (considering educational and methodological heterogeneity as highlighted in the quality assessment), meta-analysis may be employed. This will be assessed by calculating odds risk (OR) to establish the statistical significance of any differences OR for dichotomous variables or the weighted mean difference for continuous variables with 95% confidence intervals.

Thematic analysis

Clarification analysis ("How and why were these methods of assessment selected?") will use the framework described by Bearman et al. [15]. Further thematic analysis will be determined by the theoretical construct of the papers reviewed, as above. As mentioned above, there are no comparable systematic reviews to emulate.

7. Project timetable

1 year, starting February 2016

8. Conflict of interest statement

None to report

9. Plans for updating the review

 At present there are no plans to update this review.

10. Changes to the Protocol

Although no changes to the protocol are forseen, there may be changes following the initial searches, or at other unpredictable times. These will be fully documented and if substantial, submitted to BEME for ratification.

Acknowledgements

Many thanks for Steven Durning, whose "Protocol for a BEME review on assessment of diagnostic clinical reasoning" has formed the basis for several sections of this protocol.

References

1. General Medical Council. (2009). *Tomorrow's Doctors: Outcomes and standards for undergraduate medical education.* London.

2. Joint Royal Colleges of Physicians Training Board (2013) *Specialty training curriculum for core medical training*. London.

3. Frank JR, Snell L, Sherbino J, editors. (2015) *CanMEDS 2015 Physician Competency Framework.* Royal College of Physicians and Surgeons of Canada, Ottawa.

4. MacFarlane, J. (2011) Relativism and Knowledge Attributions, in *Routledge Companion to Epistemology*, ed. S Bernecker and D Pritchard (London: Routledge), 536-544

5. Lyon, A., Hothersall, E. J., Gillam E. (2015) Teaching public health in UK medical schools: 'things have improved: teaching no longer feels like an expensive hobby' *Journal of Public Health* doi: 10.1093/pubmed/fdv127

6. Scambler, G. (2012) Teaching Sociology to Medical Students http://www.grahamscambler.com/teaching-sociology-to-medical-students/ Accessed 29/10/2014

7. Jha, V., Bekker, H. L., Duffy, S. R. G., & Roberts, T. E. (2007) A systematic review of studies assessing and facilitating attitudes towards professionalism in medicine *Medical Education* 41: 822–829

8. Hawkins, R. E., Katsufrakis, P. J., Holtman, M. C., & Clauser, B. E. (2009). Assessment of medical professionalism: who, what, when, where, how, and ... why? Medical Teacher, 31(4), 348–61. doi:10.1080/01421590902887404

9. Newble, D. I., & Jaeger, K. (1983) The effect of assessments and examinations on the learning of medical students. *Medical Education*. 17:165-71

10. Knight, P. T. (2002) Summative Assessment in Higher Education: Practices in disarray, *Studies in Higher Education*. 27:3, 275-286,

11. Nicol DJ, Macfarlane-Dick D. (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*. 31(2):199–218.

12. Schuwirth, L., Colliver, J., Gruppen, L., Kreiter, C., Mennin, S., Onishi, H., Pangaro, L., Ringsted, C., Swanson, D., van der Vleuten, C., & Wagner-Menghin, M. (2011) Research in assessment: Consensus statement and recommendations from the Ottawa 2010 Conference *Medical Teacher* 33:224-233

13. PRISMA. *The PRISMA Statement* http://www.prisma-statement.org/statement.htm Accessed 29th September 2015

14. Thomas, J., and Harden, A. (2008) Methods for the thematic synthesis of qualitative research in systematic reviews *BMC Medical Research Methodology* 8:45

15. Dawson, P., Bearman, M., Boud, D. J., Hall, M., Molloy, E. K., Bennett, S. & Joughin, G. (2013). Assessment might dictate the curriculum, but what dictates assessment?. *Teaching and Learning Inquiry*, 1 (1), 107-111.