Manchester Clinical Placement Index (MCPI). Conditions for medical students’ learning in hospital and community placements

Tim Dornan • Arno Muijtjens • Jennifer Graham • Albert Scherpbier • Henny Boshuizen

Abstract The drive to quality-manage medical education has created a need for valid measurement instruments. Validity evidence includes the theoretical and contextual origin of items, choice of response processes, internal structure, and interrelationship of a measure’s variables. This research set out to explore the validity and potential utility of an 11-item measurement instrument, whose theoretical and empirical origins were in an Experience Based Learning model of how medical students learn in communities of practice (COPs), and whose contextual origins were in a community-oriented, horizontally integrated, undergraduate medical programme. The objectives were to examine the psychometric properties of the scale in both hospital and community COPs and provide validity evidence to support using it to measure the quality of placements. The instrument was administered twice to students learning in both hospital and community placements and analysed using exploratory factor analysis and a generalizability analysis. 754 of a possible 902 questionnaires were returned (84% response rate), representing 168 placements. Eight items loaded onto two factors, which accounted for 78% of variance in the hospital data and 82% of variance in the community data. One factor was the placement learning environment, whose five constituent items were how learners were received at the start of the placement, people’s supportiveness, and the quality of organisation, leadership, and facilities. The other factor represented the quality of training—instruction in skills, observing students performing skills, and providing students with feedback. Alpha coefficients ranged between 0.89 and 0.93 and there were no redundant or ambiguous items. Generalisability analysis showed that between 7 and 11 raters would be needed to achieve acceptable...
reliability. There is validity evidence to support using the simple 8-item, mixed methods Manchester Clinical Placement Index to measure key conditions for undergraduate medical students’ experience based learning: the quality of the learning environment and the training provided within it. Its conceptual orientation is towards Communities of Practice, which is a dominant contemporary theory in undergraduate medical education.

**Keywords** Workplace learning · Learning environment · Evaluation · Communities of practice · Experience based learning

**Introduction**

‘Placements’ are those parts of undergraduate medical programmes where students learn by being ‘placed’ in practice settings. Different terms like ‘rotations’, ‘firms’, ‘clerkships’, and ‘GP attachments’ are used, each carrying rather different assumptions about what students experience during placements. The present move towards greater curriculum integration, however, makes it interesting to explore what is common to different settings for practice-based learning. During placements, students meet practitioners and patients, observe practice, contribute to patient care, and are taught. The term ‘clinical teaching’, which emphasises the contribution of teaching to students’ learning, is widely used to describe what goes on in placements. Researchers have developed reliable and valid ways of evaluating and improving clinical teaching (Beckman 2010; Beckman et al. 2005; Dolmans et al. 2010; Fluit et al. 2010; Stalmeijer et al. 2010). The concept of clinical teaching, however, has several limitations. It is, self-evidently, a teacher-centred perspective but there is a shift to student-centred and patient-centred perspectives on clinical learning (Bleakley and Bligh 2008). Clinical teaching has a strong conceptual link to what Sfard (1998) termed the ‘acquisition metaphor’ (competence as a commodity that is passed from teachers to learners) whilst contemporary learning theory emphasises the ‘participation metaphor’ (learning as a social process). The final limitation of clinical teaching is the sheer number of constructs it embraces and the lack of consensus about which of them should be measured (Fluit et al. 2010).

Social phenomena can, themselves, be viewed from different theoretical perspectives (Mann et al. 2010). There are psychological and sociocultural social learning theories, the former having a more individualistic focus and the latter having a more communal focus. Sociocultural theory—particularly Communities of Practice (COP) Theory (Lave and Wenger 1991; Wenger 1998)—is very influential, judged by how many of an international panel of authors recently quoted it as an informative conceptual orientation (Dornan et al. 2010). COP theory emphasises learning rather than teaching. Our own Experience Based Learning model (eXBL) of how medical students learn in practice settings, (Dornan et al. 2007) which contextualizes COP theory to clinical education, holds that supported participation in practice is the central condition for medical students’ learning. Instruction is one important type of support but it is only one way in which practitioners contribute to students’ learning. Affective and organisational aspects of learning environments, as well as teaching, make important contributions (Dornan et al. 2007).

The theory-driven shift from teaching as the primary condition for learning to learning as something that results from being placed in practice settings calls for new instruments to measure educational quality. Learning environments (Isba and Boor 2010) provide social, organisational, and instructional support to students’ learning from real patients at
curriculum, placement, and individual interactional level (Dornan et al. 2012). Different measures are generally used to evaluate learning environments for undergraduate medical students (Soemantri et al. 2010) and interns/residents (Schonrock-Adema et al. 2009; van Hell et al. 2009) because medical students cannot be full participants in practice until they become interns. A recent systematic review found 12 published learning environment measurements instruments applicable to undergraduate medical education, three of which were judged to have acceptable psychometric properties (Soemantri et al. 2010). Two were able to distinguish between ‘traditional’ and ‘innovative’ learning environments. The authors recommended the 50-item Dundee Ready Education Environment Measure (DREEM) (Roff et al. 1997) because it is usable in different cultural settings and correlates with measures of academic achievement. No new instruments have been published since then but there have been many new studies using DREEM to evaluate, for example, different provider sites, (McKendree 2009; Veerapen and McAleer 2010) staff and student perspectives within a programme, (Miles and Leinster 2009) and different programmes in different countries (Dimoliatis et al. 2010).

The originators of DREEM did not ground its development deeply in learning theory or provide robust psychometric validity evidence (Roff 2005; Roff et al. 1997) and researchers have not been able to confirm its purported five subscale structure (Isba and Boor 2010). Regarding the reported correlation between DREEM scores and academic performance, ten Cate has pointed out that students placed in poor learning environments assiduously ‘learn to the test’, (ten Cate 2001) which confounds the relationship between a learning environment and academic performance. Our empirical research, in which we found little shared variance between a psychometrically validated learning environment measure and students’ performance in summative assessments, (Dornan et al. 2006) tends to support ten Cate’s view (ten Cate 2001). We conclude that there is a need to develop new instruments.

The aim of this research was to validate a simple measurement instrument for use in both hospital and non-hospital clinical learning environments. To achieve the aim, we tested an 11-item scale based on our own previously published development work (Dornan et al. 2004, 2006, 2003). Objectives were to examine the psychometric properties of the scale when applied to (1) hospital and (2) community COPs, and (3) recommend how it could be used to measure the quality of placements. The research was guided by Beckman et al.’s (2005) application of the American Education Research Association approach to validity evidence (American Education Research Association and American Psychological Association 1999).

Methods

Conceptual orientation

A premise of the study was that it is appropriate to evaluate hospital firms and general practices with the same set of items because, although they are different contexts of care, they do not differ as learning environments in socioculturally important respects. eXBL research (Dornan et al. 2007, 2012) assumes a ‘realist’ epistemology of causality (Wong et al. 2012) according to which certain conditions favour certain mechanisms, which favour certain outcomes. The ‘unit of analysis’ was groups of doctors from single clinical disciplines and their allied staff, working together in or out of hospital to support students’
participatory learning whilst giving patients primary, secondary, or tertiary care. In other words, communities of practice.

Research ethics approval

The University of Manchester (UK) Research Ethics Committee approved the study.

Programme

The research was conducted in year 3 of Manchester Medical School’s undergraduate medical programme, which uses problem-based learning (PBL) as its main educational method. The programme has 3 phases. In Phase 1 (years 1 and 2), students are based physically in the University and have little clinical contact. In Phase 2 (years 3 and 4), students gain clinical experience in one of four ‘sectors’, each with an academic hospital and affiliated district hospitals and general practices. In phase 3 (year 5), students have short clinical placements (we use the relatively neutral term ‘placement’ for what would be called ‘rotations’ in North America, ‘firms’ and ‘community attachments’ in Britain, and ‘stages’ in the Netherlands). After a final summative assessment, students have a period of clinical immersion in preparation for practice. Phase 2 has four system-based modules, during each of which groups of up to 12 students (usually fewer) rotate through 7-week hospital placements. They receive clinical instruction, have access to real patients with disorders relevant to the subject matter of their current curriculum module, attend PBL tutorials, receive instruction in their hospital’s clinical skills laboratory, and attend seminars open to students from all clinical unit participating in the same curriculum model. Whenever possible, a clinician from the unit to which a student is attached is their PBL tutor. Short scenarios, which students work through using ‘8-step’ study skills, (O’Neill et al. 2002) are the ‘trigger material’ for PBL. Throughout this Phase, students spend 1 day per week in general practice (GP [family medicine]) placements, whose intended learning outcomes and educational processes are similar to hospital placements. The programme is horizontally integrated in two senses: one student might learn on a surgical placement the subject matter that another student learns on a medical placement; and all students have hospital and GP placements running in parallel with one another to achieve a shared set of intended learning outcomes.

Study design

Twice per year, each student was asked to complete an online questionnaire evaluating their most recent hospital placement, GP placement, PBL tutoring experience, and the hospital as a whole. This analysis is restricted to evaluation of the hospital and GP placements. The questionnaire was delivered through the programme’s virtual learning environment (VLE). Its first page presented the argument that students have a responsibility to evaluate learning environments for the benefit of subsequent students and warned that failure to do so might result in students being debarred from using the IT system, although that sanction was never applied. The questionnaire also guaranteed that data fed back to teachers would be anonymized. We had permission to export data from the system using unique numerical identifiers to conceal students’ identities and obtain the placement identifiers that would allow us to cluster students by placement from administrators.
Administrators’ responses were very incomplete despite repeated requests but there was no evidence of systematic bias.

**Subjects**

All students in year 3 during the academic year 2006–2007 (after which the Medical School stopped giving us information that could link student and placement identifiers) were eligible for inclusion.

**Scale**

Three items—‘There was leadership of this placement’, ‘There was an appropriate reception to this placement’, and ‘I was supported by the people I met on this placement’—describe behaviours by practitioners that make learners less peripheral and therefore better able to participate in the activities of communities of practice (Dornan et al. 2012; Lave and Wenger 1991). Those behaviours create conditions for ‘mutual engagement’ between students and practitioners which, according to Wenger, contributes to their development of professional identity (Wenger 1998). A fourth one—‘This placement provided an appropriate learning environment’ is accompanied in the questionnaire by a rubric explaining ‘Your learning environment may include such things as space for students (to write notes, read, and be taught) and resources (books, computers or other materials) that support your real patient learning.’ Sociocultural theory sees physical artefacts and spaces as being important mediators of learning (Tsui et al. 2009) and ‘reification’—the crystallisation of practice in material objects—is said by Wenger to be an important counterpart to participation (Wenger 1998). The preceding four items were copied directly from our previously validated instrument. A fifth item—‘This placement was appropriately organised’—describes an attribute that, according to our recent research, makes COPs more effective learning environments (Dornan et al. 2012). A sixth item—‘I was inspired by my teachers’—was derived from other instruments used in our medical school and retained because it reflects the importance placed by eXBL on affects both as conditions for learning, and outcomes of learning (Dornan et al. 2012). Although we could have chosen other affective attributes, the ability to inspire is often quoted by learners as an important property of practitioners. A seventh item—‘This placement gave me access to appropriate real patients’—was also copied from our previous instrument and included because horizontally integrated, outcome-based education expects learners to achieve specific learning outcomes. This item measures the match of the content of practice to learning need. Another item from our previous instrument—‘I received appropriate clinical teaching’—might seem out of place in a learning environment measure but our research has consistently found it to load strongly onto the same factor as the more clearly sociocultural items listed above. The term ‘clinical teaching’ is often used rather loosely to describe interactions between learners and practitioners and we infer that the social role fulfilled by teaching is an important part of a learning environment. We included three new items—‘I was instructed in how to perform clinical skills on real patients’, ‘I was observed performing clinical tasks on real patients’, and ‘I received feedback on how I performed clinical tasks on real patients’, derived from training theory, (Patrick 1992) to elicit more focused feedback on teaching behaviours for formative purposes.

The format was identical to our previously reported scale (Dornan et al. 2004, 2006). The response format was 0–6 Likert scales (disagree-agree) with an option to enter free text. Items were in the same form as our previous measure, naming a construct, defining
the construct in terms of how it might affect a student’s learning, presenting a positively worded statement for respondents to (dis)agree with, providing the scale for them to enter their ratings, and inviting them to comment on the strengths and weaknesses of their placement in terms of that construct. We chose only to use positively worded items because the alternating positively and negatively worded format of our first generation scale (Dornan et al. 2004) was found by respondents to be very confusing. A specimen item is shown in Box 1, the full set of items used in this study are summarised briefly in Table 1, and the questionnaire we recommend for future use is included as “Appendix”.

Analysis

Data from two consecutive evaluation episodes—January and June—within a single academic year were included in analytical procedures, which used SPSS, version 15.0 (SPSS Inc, Chicago, IL, USA). We regarded it as legitimate to enter each student and each placement into the analysis twice (once for each episode) because the student-placement permutations were different on those two occasions. To avoid imprecision caused by small respondent numbers, we only included hospital placements that had been rated by three or more students (two or more for community placements because students were typically attached there individually or in very small groups). Exploratory factor analysis (EFA) using principal components analysis and varimax rotation was conducted separately for hospital and community placements, selecting factors with eigenvalues >1. To minimise ambiguity, items were only included in the final recommended version of the scale if their loadings on the two factors differed by more than 0.2 in both hospital and community. That reduced the number of items from eleven to eight, representing two distinct constructs. We decided a priori to use the factor loadings of the final EFA as evidence of validity.

Inter-rater reliability was estimated in a generalizability analysis. For the data at rater-level the variance components for placement (the variance of interest) and for rater-nested-within-placement (the error-variance) were obtained in an ANOVA. Based on the obtained variance components, the generalizability coefficient $G$ (inter-rater reliability) was obtained for a varying number of raters according to the expression

$$G = \frac{\sigma_p^2}{\sigma_p^2 + \sigma_{rp}^2/N_r},$$

where $\sigma_p^2$, $\sigma_{rp}^2$, and $N_r$, are the placement-variance, the rater-nested-within-placement-variance, and the (hypothetical) number of raters, respectively (Brennan 2001).
Results

Since each student was eligible to complete the questionnaire twice in the study period and there were 451 students in the cohort, the 754 completed questionnaires represented an 84% response rate. Hospital placement details were known for 615 respondents (68%). Five hundred and ninety two responses (66%) rated 90 of 101 hospital placements (89%) the three or more times that qualified a placement to be included in the analysis. The number of eligible responses concerning community placements was predictably smaller; 253 valid responses (28%) of 902 possible ones evaluated 78 placements.

Hospital placements

Table 2 shows the results of the principal component analysis which loaded onto two factors, together accounting for 76% of the variance in the data. Factor loadings of three of the eleven items—This placement gave me access to appropriate real patients; I received appropriate clinical teaching; I was inspired by my teachers—did not differ by >0.2.

Community placement

Table 2 shows that the data loaded onto the same two factors, together accounting for 79% of the variance. Factor loadings of two of the three items named above—I received appropriate clinical teaching; I was inspired by my teachers—did not differ by >0.2.

Final scale

Because our goal was to provide a single measurement instrument that was applicable to both hospital and community learning environments and whose structure was
unambiguous, we repeated the factor analysis after removing the three ambiguous items, which reduced the scale from eleven to eight items. Results are shown in Table 3. A two factor solution explained 78% of the variance in hospital data and 82% of the variance in community data. We termed the two factors ‘learning environment’—because the five items that loaded onto the first factor referred to social and material aspects of the learning environment—and ‘training’, because instruction and extrinsic feedback based on observation are key components of training. Patrick (1992, pp. 34, 306) Alpha coefficients were 0.89 for learning environment in hospital and 0.93 for learning environment in community; neither coefficient increased when any item was deleted. Alpha coefficients for training were 0.92 in hospital and 0.93 in community; neither coefficient increased when any item was deleted. Pearson correlation coefficients were 0.68 for the relationship between the two scales in hospital data and 0.71 between the two scales in community data, showing there was significant interdependence between them.

Table 2 Results of exploratory factor analysis

| Item labels          | Hospital | Community |
|----------------------|----------|-----------|
|                      | Factor 1 | Factor 2  | Factor 1 | Factor 2  |
| Reception            | .809     | .227      | .806     | .365      |
| People               | .796     | .276      | .829     | .321      |
| Leadership           | .771     | .362      | .742     | .415      |
| Organisation         | .756     | .440      | .811     | .388      |
| Facilities           | .742     | .276      | .821     | .209      |
| Real patient access  | .500     | .613      | .822     | .347      |
| Inspiration          | .694     | .526      | .637     | .543      |
| Observation          | .263     | .914      | .322     | .880      |
| Feedback             | .262     | .854      | .316     | .881      |
| Instruction          | .428     | .825      | .369     | .847      |
| Clinical teaching    | .627     | .652      | .665     | .594      |

Values are shown in bold when the loading on factors 1 and 2 did not differ by >0.2. Items are marked in italics to indicate the factor onto which they loaded more heavily. The item previously named ‘learning environment’ has been renamed ‘facilities’ to avoid confusion.

Table 3 Factor loadings for the final scale

| Item labels     | Hospital | Community |
|-----------------|----------|-----------|
|                 | Learning environment | Training | Learning environment | Training |
| Reception       | .814     | .249      | .826     | .376      |
| People          | .790     | .252      | .827     | .318      |
| Organisation    | .776     | .434      | .812     | .400      |
| Leadership      | .772     | .372      | .745     | .443      |
| Facilities      | .771     | .248      | .834     | .191      |
| Observation     | .286     | .919      | .322     | .896      |
| Feedback        | .280     | .886      | .317     | .877      |
| Instruction     | .438     | .799      | .366     | .843      |

Items are marked in italics to indicate the factor onto which they loaded more heavily. The two factors are now named ‘learning environment’ and ‘training’. 
Generalisability analysis showed that 7 raters would be needed to evaluate a hospital learning environment reliably (generalisability coefficient $\geq 0.7$) and 9 raters to evaluate a community learning environment. Eleven raters would be needed to evaluate training in hospital and nine in community.

Discussion

Principal findings and meaning

Our simple 8-item instrument reliably measured the social/material quality of workplace learning environments for undergraduate medical students and the training provided within them. Following the American Education Research Association approach to providing validity evidence (American Education Research Association and American Psychological Association 1999) as applied to medical education (Beckman et al. 2005) the instrument’s content validity of items rests on Communities of Practice (Wenger 1998) and eXBL (Dornan et al. 2007, 2009) theories, so it reads theories as well as the research and practitioner engagement that went into developing the instrument (Dornan et al. 2004, 2006, 2003). We have previously published a theoretical and empirical justification for the way we use learner self-report response processes in this study (Dornan et al. 2004). This paper reports the internal structure and interrelationship of the measure’s variables. In addition to the psychometric data reported here, we have previously reported how the free text items in the same instrument can give rich information about students’ learning (Bell et al. 2009). We do not yet have data on the consequences of the measure. We propose, however, that there is sufficient validity evidence to prompt further testing of this measure to support quality improvement of integrated clinical education.

Strengths and limitations

Strengths of this research were its grounding in education theory and prior empirical research. Data were provided by large numbers of medical students on large numbers of placements in four academic hospitals and a range of general practices. Further, students evaluated learning environments additional to their hospital firms and general practices (e.g. the whole hospital as a learning environment) so the measure was validated within a more comprehensive test battery. An important potential limitation is that the two factors—learning environment and training—correlated with one another, so they were not fully independent variables. That has been a consistent finding in our research. Whilst providing good training and providing a supportive learning environment may seem to be separate constructs, students have told us that the best training emerges from warm social environments, so some interdependence between these two factors may be inescapable.

Relationship to other publications

As explained in ‘introduction’, DREEM is widely used to evaluate clinical placements and was recently suggested to be the best validated measure (Soemantri et al. 2010). It calls on students to rate 50 items, compared with the 8 reported here, and does not include free text items. Textual comments are at least as useful as numerical ones because they can define how quality improvement effort should be invested. DREEM asks students to rate their
perceptions of themselves as well as of their learning environment, for which the rationale is unclear. That lack of clarity is supported by lack of evidence that the various constructs it evaluates are psychometrically independent of one another (Isba and Boor 2010) so we argue that this ‘Manchester Clinical Placement Index’ is a simpler, validated, mixed methods instrument that is worth considering as an alternative to DREEM.

Implications

Until further validation research has been carried out in the context of educational practice, we caution against uncritical adoption of this measure, although we are optimistic it will prove useful. Future research should include a head-to-head comparison with DREEM, which Soemantri et al. (2010) have proposed as the standard for evaluating learning environments. Finally, we agree with Beckman et al. conclusion (Beckman et al. 2005) that what is now needed is evidence of the consequential validity of this and other placement quality measures. The 8-item measurement instrument we recommend for use is contained in “Appendix”.

Learning environment is calculated as a percentage according to the formula:

\[
\frac{\text{Leadership} + \text{Reception} + \text{People} + \text{Facilities} + \text{Organisation}}{30} \times 100
\]

Training is calculated as a percentage according to the formula:

\[
\frac{\text{Instruction} + \text{Observation} + \text{Feedback}}{18} \times 100
\]

Reliable measurements can be made by averaging the ratings of over 10 respondents. Free text responses can be used to explain numerical differences and identify strengths and weaknesses for quality improvement purposes.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix: Manchester Clinical Placement Index (MCPI)

Leadership

There is leadership if one or more senior doctors (consultant, GP, registrar) take responsibility for your education

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

There was leadership of this placement

Please add comments to either or both of the next two boxes

Strengths of leadership were … (Free text box)

Weaknesses or ways leadership could be improved … (Free text box)

Reception/induction

An appropriate reception is a welcome that includes an explanation of how the placement can contribute to your real patient learning
Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

**There was an appropriate reception to this placement**

Please add comments to either or both of the next two boxes

- **Strengths of the reception were…** (Free text box)
- **Weaknesses or ways the reception could be improved …** (Free text box)

**People**

The support to your real patient learning from people (like doctors, secretaries, receptionists, nurses, and others) you met on the placement

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

I was supported by the people I met on this placement

Please add comments to either or both of the next two boxes:

- **Strengths of any or all of the groups of people listed above were …** (Free text box)
- **Weaknesses of any of the groups of people listed above or ways they could contribute more …** (Free text box)

**Instruction**

Clinical teaching may include instruction in how to perform clinical skills (like history taking, examination, practical procedures etc.) on real patients

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

I was instructed in how to perform clinical skills on real patients

Please add comments to either or both of the next two boxes:

- **Strengths of instruction were …** (Free text box)
- **Weaknesses or ways instruction could be improved …** (Free text box)

**Observation**

Clinical teaching may include teachers observing you perform clinical tasks on real patients

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

I was observed performing clinical tasks on real patients

Please add comments to either or both of the next two boxes:

- **Strengths of observation were …** (Free text box)
- **Weaknesses or ways observation could be improved …** (Free text box)
Feedback

Clinical teaching may include teachers giving you feedback on how you performed clinical tasks on real patients

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

I received feedback on how I performed clinical tasks on real patients

Please add comments to either or both of the next two boxes:

Strengths of feedback were… (Free text box)
Weaknesses or ways feedback could be improved … (Free text box)

Facilities

Your learning environment may include such things as space for students (to write notes, read, and be taught) and resources (books, computers or other materials) that support your real patient learning

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

This placement provided appropriate facilities

Please add comments to either or both of the next two boxes:

Strengths of the facilities were … (Free text box)
Weaknesses or ways the facilities could be improved … (Free text box)

Organisation of the placement

An appropriately organized placement is one whose teaching and learning activities are organized in a way that supports your real patient learning

Please rate your agreement (0 = strongly disagree; 3 = neither agree nor disagree; 6 = strongly agree) with this statement:

This placement was appropriately organized

Please add comments to either or both of the next two boxes:

Strengths of organization were… (Free text box)
Weaknesses or ways organisation could be improved … (Free text box)

References

American Education Research Association and American Psychological Association. (1999). Standards for educational and psychological testing. Washington, DC: American Education Research Association.

Beckman, T. J. (2010). Understanding clinical teachers: Looking beyond assessment scores. Journal of General Internal Medicine, 25(12), 1268–1269. doi:10.1007/s11606-010-1479-6.

Beckman, T., Cook, D. A., & Mandrekar, J. N. (2005). What is the validity evidence for assessments of clinical teaching? The Journal of General Internal Medicine, 20, 1159–1164.
for the evaluation of clinical teachers. *Academic Medicine, 85*(11), 1732–1738. doi:10.1097/ACM.0b013e3181f554d6.

ten Cate, O. (2001). What happens to the student? The neglected variable in educational outcome research. *Advances in Health Sciences Education, 6*, 81–88.

Tsui, A. B. M., Lopez-Real, F., & Edwards, G. (2009). Sociocultural perspectives of learning. In A. B. M. Tsui, G. Edwards, & F. Lopez-Real (Eds.), *Learning in school-university partnership. Sociocultural perspectives*. New York: Routledge.

van Hell, E. A., Kuks, J. B. M., & Cohen-Schotanus, J. (2009). Time spent on clerkship activities by students in relation to their perceptions of learning environment quality. *Medical Education, 43*, 674–679.

Veerapen, K., & McAleer, S. (2010). Students’ perceptions of the learning environment in a distributed medical programme. *Medical Education, 15*. doi:10.3402/meo.v15i0.5168.

Wenger, E. (1998). *Communities of practice. Learning, meaning and identity*. Cambridge: Cambridge University Press.

Wong, G., Greenhalgh, T., Westhorp, G., & Pawson, R. (2012). Realist methods in medical education research: What are they and what can they contribute? *Medical Education, 46*, 89–96.