THE CLINICAL LEARNING CLIMATE

Klarke Boor

Amsterdam, Vrije Universiteit Amsterdam, Faculteit Geneeskunde
Thesis with summary in Dutch
Proefschrift met een samenvatting in het Nederlands
ISBN: 978-90-865-9370-5
Lay-out en Print: Optima BV, Rotterdam

Financial support of this thesis is gratefully acknowledged and was provided by the St Lucas Andreas Hospital and the Free University.

VRIJE UNIVERSITEIT

THE CLINICAL LEARNING CLIMATE

ACADEMISCH PROEFSCHRIFT

ter verkrijging van
de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op vrijdag 23 oktober 2009 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

Door

Klarke-Berendina Boor

Geboren te Kirinda, Rwanda

Promotoren: Prof.dr. F. Scheele
Prof.dr. A.J.J.A. Scherpbier
Prof.dr. C.P.M. van der Vleuten

aan mijn ouders

TABLE OF CONTENTS

1

Introduction

INTRODUCTION

The topic of this thesis is the learning climate of medical students and residents working in hospitals, in other words the clinical learning climate. Depending on their curriculum, medical students are introduced to the clinical workplace at different points in their training: some do not enter the clinical workplace until the final two years of training; others are offered earlier clinical experiences. Residents (doctors with a basic medical degree undertaking specialist training) receive their training predominantly in the clinical workplace. Both students and residents perceive a certain learning climate while working in their departments. Their impression of this climate may be 'good', 'inspiring', or even 'optimal', but it is as likely to be 'bad', 'discouraging', or even 'deplorable', with every possible shade of appreciation in between. What is important is that the local learning climate crucially shapes the learning experiences in that particular hospital ward.

Both medical students and residents learn and work in a constant state of flux. This is first and foremost inherent in working within the complex reality of hospital departments, but it is also related to changes in curriculum and governance. The last decade, for instance, competency-based residency programmes have been introduced in many countries.[1;2] These competency-based curricula are, amongst others, a reaction to public demand for residents who are not 'just' medical experts, but also communicators, collaborators, managers, health advocates, scholars, and professionals.[3]
Changing a curriculum is costly: governments, accreditation bodies, and the people organising and running the new training programme invest time, effort, and/or money in this project. It should therefore not come as a surprise that curriculum change can lead to calls for more transparency through evaluating curricula by measuring educational outcomes and processes.[4;5] One aspect of process evaluation is evaluation of the educational climate. Considering that the quality of the educational climate influences learning, satisfaction, and career choices of learners[6], better understanding of learning climates can afford insight into the effectiveness of (changed) curricula. Although measurement of educational climates (predominantly in medical schools) has been used extensively to evaluate curricula, educational climate as a concept has remained ambiguous and hard to define.[6-9] So, although it is generally considered relevant to measure learning climates in order to evaluate educational quality, the concept itself remains to be elucidated by further studies. This will be the subject of this thesis.

This introduction presents an overview of earlier research into learning climates in order to provide a background to the central research questions of this thesis. The chapter starts with an overview of the literature on the environments in which people work and learn, beginning with organisational climate studies and how these are related to research into organisational

culture. Next, the role of the learning climate in medical education studies is summarised. This leads up to the central research questions and an overview of the studies that constitute the body of this thesis.

ORGANISATIONAL SCIENCE: CLIMATE AND CULTURE RESEARCH

Organisational climate research

There are many definitions of organisational climate, ranging from "a summary perception which people have of an organisation"[10] to the more elaborate definition of: "Organisational climate is a relatively enduring characteristic of an organisation which distinguishes it from other organisations: and (a) embodies members collective perceptions about their organisation with respect to such dimensions as autonomy, trust, cohesiveness, support, recognition, innovation, and fairness; (b) is produced by member interaction; (c) serves as a basis for interpreting the situation; (d) reflects the prevalent norm values and attitudes of the organisation's culture; and (e) acts as a source of influence for shaping behaviour".[11] There is, however, not one, generally accepted definition.

The first publication on 'climates' was an article by Lewin et al. in 1939. The authors described how artificially created social climates influenced the aggressive behaviour of young boys attending various club activities.[12] The steep rise in the number of articles on organisational climate in the 1960s and 1970s reflected the popularity of climate measurements as a tool to get a grasp of organisational functioning.[13-15] Organisational climate researchers did not devote a great deal of time to defining the concept, but rather concentrated on gathering data and developing instruments.[15] They focused on individual perceptions of observable procedures and practices close to the surface of 'organisational life'.[14;16] Mainly quantitative methods (surveys) were used to measure these perceptions with individual outcomes being aggregated at departmental or organisational level. Empirical findings showed a slight relationship between organisational climate and organisational performance[17;18] as well as a manifest influence of organisational climate on motivation[17;19] and job satisfaction.[17;18] In a commentary in 1973, Guion claimed that organisational climate was actually the same as job satisfaction,[16] but others disagreed.[10;18;20] Schneider and Snyder, for instance, showed that people working in positively rated climates were not always the most satisfied employees.[10]

Dozens of questionnaires enabled managers to measure, and potentially influence, their organisational climates (for an overview of climate instruments see Schneider's book on organisational climate and culture (p. 10-13)[15]).

Organisational culture research

In the early 1980s, however, many felt that testing the organisational climate might not be the best way to determine how organisations really functioned: climate measurements lacked detail and resulted in unsatisfactory feedback.[13] The subsequent increase in studies of organisational culture can be seen as a sign of an 'anti-climate movement'. This was clearly expressed by Schwartz and Davis who stated in their 1981 paper that whatever culture might be, it definitely was not climate.[21]

Although undisputed definitions are lacking, it is generally assumed that organisational culture is linked to 'deep', taken-for-granted assumptions that emerge through more 'superficial', observable behaviours and artefacts. Schein described three cultural levels: the basic (implicit) assumptions, (more explicit) beliefs and values, and (manifest) patterns of behaviours and artefacts.[11;22;23] Cultures should be seen as historically constituted and socially constructed entities.[11;22;23]

Many efforts aimed at accomplishing a change in organisational culture do not move beyond the artefact level, leaving underlying assumptions unexamined. This may explain the relative lack of success of interventions aimed at producing cultural change.[24] The question whether (interventions aimed at a change of) organisational culture influences organisational performance or success remains to be answered.[24] Scott et al. reviewed the potential influence of organisational culture on (the quality of) health care performance and concluded that such an influence is hard to demonstrate. Different operationalisations of culture and performance together with the multi-factorial nature of both concepts may explain these findings.[25]

Differences and similarities

There has been much debate about the differences and similarities between research efforts relating to climate and culture.[13;14;26;27] Denison has clearly summarised the differences between these two approaches (table 1).[14] A noticeable difference, for instance, relates to epistemology and methodology. Generally, climate researchers judge knowledge of climates in relation to other climates: they aim to compare, whereas culture researchers are primarily concerned with valuing the contextualised knowledge of the culture under study. As for methodology, climate researchers use mostly quantitative techniques, whereas culture researchers focus on the unique and situated characteristics of social action, using (qualitative) ethnographic and anthropological methods.

In the last two decades, however, the emphasis has gradually shifted to the similarities between the two approaches.[13;14;20;26-28] An example is the interactionist approach to the formation of organisational climates.[11;20] Based on symbolic interactionism- a theoretical foundation of organisational culture research - this approach characterises climate ".. as a joint property of both the organisation and the individual."(p. 838)[20] Another illustration is

Table 1: Characteristics of organisational climate and culture research

	Organisational climate	Organisational culture
Epistemology	Comparative	Contextualised
Methodology	Quantitative survey data	Qualitative field observation
Level of analysis	Surface-level manifestations	Underlying values / assumptions
Temporal orientation	A -historical snapshot	Historical evolution
Theoretical foundation	Lewinian field theory	Symbolic interaction& Social construction
Discipline	Psychology	Sociology & anthropology

(after Denison[14])

that culture researchers have also started to use questionnaires to compare different organisational cultures (a method that used to be climate researches' prerogative).[14;22;29] Moreover, there is a growing interest in the combined use of qualitative and quantitative methods among researchers of organisational cultures and climates.[28-30] Denison concludes his article by saying that "..these two research traditions should be viewed as differences in interpretation rather than differences in phenomenon."[14](p. 645).

In sum, the organisational science literature started out with climate research characterised by quantitative surveys to chart how organisations work. In reaction to this - predominantly positivist - approach, an 'anti-climate movement' arose, which favoured (qualitative) anthropological methods to investigate organisational culture. Most recently, several authors have advocated for an integrative take on climate and cultural research to guide efforts to understand the working of organisations. However, none of these developments have managed to produce an undisputed definition, neither of organisational climate nor of organisational culture.

MEDICAL EDUCATION: LEARNING CLIMATE RESEARCH

Genn defined the learning climate as "...the soul and spirit of the medical school educational environment (..) Put more prosaically, "the educational climate is a manifestation or operationalisation of the educational environment and of the curriculum".[6](p.446) Rotem described it as "..a subtle concept, encompassing physical, interpersonal, and organisational elements".[9](p706) What these descriptions highlight is that it is not easy to capture in a definition the factors that constitute and contribute to the learning climate.[7;8] What medical education authors agree on, however, is that it is important to create a 'good' learning climate and to make this the subject of further research.[6;8;31-34] Some authors have stressed the

importance of developing instruments for specific situations (i.e., the clinical setting),[6;31;32] or for specific groups of students (i.e., women, clinical students, people from different cultural backgrounds).[8] Another author promoted qualitative analysis of clinical environments to further understand the impact of the clinical context.[33]

The next paragraphs describe the literature on the development of surveys, their validity and reliability, and on qualitative research. With respect to content, the focus is on studies concerned with non-clinical as well as clinical settings. Although this thesis is limited to the clinical setting, non-clinical studies are of interest too because they provide a historical perspective and some of them have laid the foundation for clinical studies.

Surveys

When organisational climate studies were flourishing, some studies were published about the medical education climate. Most authors assumed that insight into educational climate would yield insight into the functioning of (changed) curricula. In 1961 Hutchins reported a study in which he examined medical school graduates' perceptions of the educational climate using a questionnaire developed for measuring university college climates (the College Characteristics Index[35]).[36] Rothman and Ayoade developed the Learning Environment Questionnaire (LEQ)[37], which was then used by Levy et al. to assess curricular change in an non-clinical curriculum.[38] Marshall developed a questionnaire based on these two surveys, which he used to measure elements of the learning environment relevant to student stress.[39] The validity of Marshall's questionnaire was subsequently measured by Feletti and Clarke.[40]
After a period of relative silence around this subject, the 1990s saw an upsurge in the interest in measuring learning climates. The group of Roff and McAleer developed the Dundee Ready Educational Environment Measure (DREEM) as a 'culture-free' inventory to measure the educational environment for the health professions.[8;41] This instrument includes items on subjects such as teaching, teachers, and atmosphere. Experts from multiple countries participated in the development and the questionnaire was administered to a reasonably large group of international students. The article describing the development of DREEM has some flaws, however: (a) an incomplete, and imprecise, description of the methodologies (the term 'Delphi procedure' was used with reference to independent expert meetings and not to consecutive rounds[42]; exploratory factor analysis was mentioned without data to back up the results); (b) there was no discussion section.[41] Despite these shortcomings, the instrument has been used in multiple countries in different medical schools and nursing colleges.[43-48]
Within the context of this thesis, the Postgraduate Hospital Educational Environment Measure (PHEEM) is of interest because it is tailored to the learning climate in the clinical setting. Roff, McAleer, and Skinner developed PHEEM to gather information on subjects such as feedback, facilities, and atmosphere.[49] After (unspecified) literature research, educational experts and

post registration house officers reduced the initial 180 items to 90, followed by a focus group session to decide on the definitive 40 items. The authors identified three subscales without empirical testing.[49] Several other 'EEM –instruments[7;50-55]and additional educational climate surveys[56;57] have been published.

Validity and reliability

Validity and reliability are important aspects of survey quality. The American Psychological and Education Research Associations published standards identifying five sources of validity evidence: (1) Content, (2) Response process, (3) Internal Structure, (4) Relations to other variables, and (5) Consequences (see table 2 for further explanation of these five sources).[58;59] Reliability coefficients estimate measurement error in assessment and quantify the measurement's consistency.[60] Most used estimates are Cronbach's α (based on the test-retest concept and indicating internal consistency), the Kappa statistic (a correlation coefficient, indicating inter-rater reliability), ANOVA (also indicating inter-rater reliability) or generalisability theory (a more 'elegant' estimate of the concurrent effect of multiple sources on reliability).[60;61]
Table 3 summarises the properties of the questionnaires described above and shows which validity evidence sources and reliability estimates have (not) been used.
In summary, several survey instruments have been developed and are widely used to measure educational climates. Unfortunately, the psychometric qualities tend to be rather weak. Many instruments report minimal data on validity evidence and focus predominantly on the categories 'Content' and 'Internal structure' (this resembles findings from other studies within

Table 2: Five sources of validity evidence

Validity evidence source	Definition
Content	The relationship between a test's content and the construct it is intended to measure. Refers to themes and wording of items. Includes experts' input. Also includes development strategies to ensure appropriate content representation.
Response process	Analyses of responses, including the strategies and thought processes of individual respondents. Differences in response processes may reveal sources of variance that are irrelevant to the construct being measured. Also includes instrument security, scoring, and reporting of results
Internal structure	The degree to which items fit the underlying construct. Most often reported as measures of internal consistency and factor analysis.
Relation to other variables	The relationship between scores and other variables relevant to the construct being measured. Relationships may be positive (convergent or predictive) or negative (divergent or discriminant).
Consequences	Surveys are intended to have some desired effect, but they also have unintended effects. Evaluating such consequences can support or challenge the validity or score interpretations.

(after Beckman[54])

assessment[58]). Reliability estimates are either absent or limited to Cronbach's α (which can be deceptive, as it may be high due to large numbers of items and thus convey little information about reliability[62]).

Qualitative studies

Studies using (a combination of quantitative and) qualitative methodologies to examine the medical learning climate are not as abundant as the above-described research and most of them were only published in the past decade. In a thorough, ethnographic study, Seabrook

Table 3: Properties of surveys measuring the medical education learning climate

Name	Authors	Nr items, nr subscales	Study population	Validity Evidence sources*		Reliability Estimates
College Characteristics Index (CCI)†	Pace & Stern (1958)	300 items, 30 subscales	423 students, 71 faculty members	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal Minimal No	Inter-rater reliability item discrimination test
Medical School Environment Inventory	Hutchins (1961)	180 items, 18 subscales	1901 med school graduates	Content Response process Internal structure Relat. other variables Consequences	Minimal No No No No	-
Learning Environment Questionnaire	Rothman & Ayoade (1971)	65 items, 7 subscales	145 1st yr medical students	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Internal consistency
Medical School Learning Environment	Marshall, (1978); Feletti & Clarke (1981)	50 items, 7 subscales	93 1st yr medical students	Content Response process Internal structure Relat. other variables Consequences	Yes Minimal Minimal Minimal No	Cronbach's α Other test-retest measures
	Rotem et al. (1995)	46 items, 8 subscales (1 other section: nr of items not described)	209 1st yr residents	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α
Dundee Ready Educational Environment Measure (DREEM)	Roff et al. (1997)	50 items, 5 subscales	490 medical students, 256 nursing students	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α
	Pololi & Price, 2000	31 (or 15) items, 3 subscales	619 non clinical medical students	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α

* 'No' indicates no data regarding this category; 'Minimal' indicates minimal data regarding this category; 'Yes' indicates sound and multiple data sources supporting the category.
† This is not a medical educational climate survey, but since it serves as a basis for many later developed surveys it is included in this list.

Table 3: Properties of surveys measuring the medical education learning climate (continued)

Name	Authors	Nr items, nr subscales	Study population	Validity Evidence sources*		Reliability Estimates
Anaesthetic Theatre Educational Environment Measure (ATEEM)	Holt & Roff (2004)	40 items, 5 subscales	218 anaesthesiology residents	Content Response process Internal structure Relat. other variables Consequences	Minimal No No No No	-
Surgical Theatre Educational Environment Measure (STEEM)	Cassar (2004)	40 items, 4 subscales	25 surgical residents	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α
Practice-based Educational Environment Measure	Mulroone y (2005)	37 items, 4 subscales	48 GP's in training	Content Response process Internal structure Relat. other variables Consequences	Minimal No No No No	-
Postgraduate Hospital Educational Environment Measure (PHEEM)	Roff et al. (2005)	40 items, 3 subscales	No study population	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	-
DREEM for residents	De Oliveia Filho et al. (2005)	50 items, 5 subscales	97 residents	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal Minimal No	Cronbach's α Other test-retest measures
Operating Room Educational Environment Measure (OREEM)	Kanashir o et al. (2006)	40 items, 4 subscales	22 surgical residents	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α
Mini-STEEM	Nagraj et al.(2007)	14 items	83 final yr medical students	Content Response process Internal structure Relat. other variables Consequences	Minimal No Minimal No No	Cronbach's α

* 'No' indicates no data regarding this category; 'Minimal' indicates minimal data regarding this category; 'Yes' indicates sound and multiple data sources supporting the category.

elicited medical students' experiences immediately after their first clinical encounters.[63] She paints a picture of students striving 'to fit in' but at the same time feeling that they are 'in the way'. Early clinical experiences, proper induction meetings, and greater awareness among staff of students' needs could perhaps ease the transition to clinical learning. In an extensive qualitative study in a large Canadian medical school, Cleave-Hogg and Rothman illuminated the effects of learning environment characteristics on students' approaches to learning.[64] Their interviews with both teachers and medical students revealed that a good teacher-student relationship could cancel out factors hindering learning (like the highly competitive environment).[64;65] This was confirmed by a study among PRHOs where "feeling

fully supported" was the factor most frequently mentioned as enhancing the learning environment.[66] Other studies highlighted the chaotic reality facing students and residents in combining working and learning in clinical contexts.[67;68]
Frye et al. listed qualitative methods that could be used to study clinical learning environments. They recommended multiple techniques (triangulation) and also soliciting feedback from respondents after data analysis.[69]
In short, these studies offer some insights into (non) clinical learning climates. The clinical workplace can be a daunting and unstructured environment. Teachers and staff members have a strong impact on how medical students and residents perceive the learning climate.

CENTRAL RESEARCH QUESTIONS

As described above, climate research by organisational scientists shifted from predominantly quantitative survey studies to ethnographic situated case studies of unique work places. In addition, in recent years, many authors have advocated a combination of these two approaches to study workers' perceptions.
Within the medical education field, climate studies using surveys have dominated the picture. Many instruments, most of them developed with limited attention to psychometric properties, have been described and used. The last decade, a few qualitative studies have attempted to shed light on the complex characteristics of (clinical) learning environments. Both quantitative and qualitative studies have stressed the importance of the role of teachers and staff members in determining how students' and residents' perceive the learning climate.

Still, a clear definition of the learning climate, let alone the clinical learning climate, remains elusive and much remains to be explored about the factors and concepts that are involved. What we can say, however, based on the literature summarised above, is that research into learning climates probably requires:

- qualitative and quantitative methodologies,
- an interconnection between the literatures on organisational climates and cultures: for evaluating educational climates it is important to compare different organisations but the literature also stresses the significance of historical and social influences,
- psychometrically sound instruments for evaluating clinical learning climates.

The main research questions of this thesis are:

Which elements play a part in the clinical learning climate? How can they be measured? What is the clinical learning climate? How can it be defined?

Overview of the studies in this thesis

Several studies were conducted to address the research questions. In order to arrive at a description or definition of the clinical learning climate and the factors that play a role, three studies were conducted. Chapter 3 focuses on the clinical teacher as a key player in the clinical learning climate. This qualitative study describes which clinical teacher characteristics are favoured by obstetric-gynaecological residents. Chapters 4 and 5 highlight medical students' and residents' perceptions of important factors for clinical learning climates. Chapter 4 employs a combined quantitative and qualitative approach, using survey results and semi-structured interviews to construct a model of medical students' clinical learning climate. The study in Chapter 5 uses a two-step approach: semi-structured interviews, resulting in an initial schematic representation of residents' learning climate, followed by several focus group sessions to finalise the scheme. This resulted in a final model of residents' learning climate.
The thesis starts, however, with a study of an existing questionnaire that is highly relevant for the clinical phase: the Postgraduate Hospital Educational Environment Measure (PHEEM) (Chapter 2). PHEEM's validity and reliability were tested to determine its ability to measure the clinical learning climate. Chapter 6 describes a similar study concerning the validity and reliability of an instrument measuring residents' learning climate but this chapter also describes how the questionnaire (the Dutch Residents Educational Climate Test, D-RECT) was developed based both on the judgements of an expert panel and the theoretical model described in Chapter 5. This study gives an account of how this thorough development process resulted in a valid and reliable questionnaire to measure the clinical learning climate. Finally, Chapter 7 summarises the main findings, the strengths and weaknesses, the theoretical significance, and the practical implications of this thesis.
Because the studies in most chapters have been published as journal articles, some repetition across chapters was unavoidable.

REFERENCES

1. Bannon M. What's happening in postgraduate medical education? *Arch.Dis.Child* 2006;**91**:68-70.
2. Ludmerer KM, Johns MM. Reforming graduate medical education. *JAMA* 2005;**294**:1083-7.
3. Societal Needs Working Group. Skills for the new millennium: CanMEDS 2000 project. *Ann R Coll Phys Surg Can* 1996;**29**:206-16.
4. Wilkes M, Bligh J. Evaluating educational interventions. *BMJ* 1999;**318**:1269-72.
5. Hayden J, McKinlay D. Monitoring standards of training. *Med.Educ.* 2001;**35**:68-72.
6. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. *Med.Teach.* 2001;**23**:445-54.
7. Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. *Med Teach.* 2004;**26**:260-4.
8. Roff S, McAleer S. What is educational climate? *Med Teach* 2001;**23**:333-4.
9. Rotem A, Bloomfield L, Southon G. The clinical learning environment. *Isr.J.Med.Sci.* 1996;**32**:705-10.
10. Schneider B, Snyder RA. Some relationships between job satisfaction and organizational climate. *J Appl Psychol* 1975;**60**:318-28.
11. Moran ET, Volkwein JF. The cultural approach to the formation of organizational climate. *Hum Relat* 1992;**45**:19-47.
12. Lewin K, Lippitt R, White RK. Patterns of aggressive behavior in experimentally created "social climates". *J Soc Psychol* 1939;**10**:271-99.
13. Ashkanasy NM, Wilderom CPM, Peterson MF. Introduction. In Ashkanasy NM, Wilderom CPM, Peterson MF, eds. *Handbook of organizational culture & climate*, pp 1-18. Thousand Oaks CA: Sage Publications, 2000.
14. Denison DR. What *is* the difference between organizational culture and organizational climate? A native's point of view on a decade of paradigms wars. *Acad Manage Rev* 1996;**21**:619-54.
15. Reichers AE, Schneider B. Climate and culture: an evolution of constructs. In Schneider B, ed. *Organizational climate and culture*, pp 5-40. San Francisco: Jossey-Bass Inc., 1990.
16. Guion RM. A note on organizational climate. *Organ Behav Hum Perform* 1973;**9**:120-5.
17. DeCotiis TA, Summers TP. A path analysis of a model of the antecedents and consequences of organizational commitment. *Hum Relat* 1987;**40**:445-70.
18. Pritchard RD, Karasick BW. The effects of organizational climate on managerial job performance and job satisfaction. *Organ Behav Hum Perform* 1973;**9**:126-46.
19. Litwin GH, Stringer RA. Motivation and organizational climate. Boston: Harvard University Press, 1968.
20. Ashforth BE. Climate formation: issues and extensions. *Acad Manage Rev* 1985;**10**:837-47.
21. Schwartz H, Davis S. Matching corporate culture and business strategy. *Organ Dyn* 1981;**10**:30-8.
22. Ashkanasy NM, Broadfoot LE, Falkus S. Questionnaire measures of organizational culture. In Ashkanasy NM, Wilderom CPM, Peterson MF, eds. *Handbook of organizational culture & climate*, pp 131-46. Tousand Oaks CA: Sage Publications, 2000.
23. Schein EH. Organizational culture and leadership: a dynamic view. San Francisco: Jossey-Bass, 1985.
24. Davies HT, Nutley SM, Mannion R. Organisational culture and quality of health care. *Qual.Health Care* 2000;**9**:111-9.
25. Scott T, Mannion R, Marshall M, Davies H. Does organisational culture influence health care performance? A review of the evidence. *J.Health Serv.Res.Policy* 2003;**8**:105-17.

26. Schneider B. Organizational climate and culture. San Francisco: Jossey-Bass Inc., 1990.
27. Payne RL. Climate and culture: how close can they get? In Ashkanasy NM, Wilderom CPM, Peterson MF, eds. *Handbook of organizational culture & climate*, pp 163-76. Thousand Oaks CA: Sage Publications, 2000.
28. Rentsch JR. Climate and culture: interaction and qualitative differences in organizational meanings. *J Appl Psychol* 1990;**75**:668-81.
29. Scott T, Mannion R, Davies H, Marshall M. The quantitative measurement of organizational culture in health care: a review of the available instruments. *Health Serv.Res* 2003;**38**:923-45.
30. Siehl C, Martin J. Measuring organizational culture. Mixing qualitative and quantitative methods. In Jones MO, Moore MD, Snyder RC, eds. *Inside Organizations: Understanding the Human Dimension*, pp 79-103. Newbury Park, CA: Sage Publications, 1988.
31. Genn JM. AMEE Medical Education Guide No. 23 (Part 1): Curriculum, environment, climate, quality and change in medical education-a unifying perspective. *Med.Teach.* 2001;**23**:337-44.
32. Roff S. New resources for measuring educational environment. *Med.Teach.* 2005;**27**:291-3.
33. Snadden D. Clinical education: context is everything. *Med Educ* 2006;**40**:97-8.
34. Hutchinson L. Educational environment. *BMJ* 2003;**326**:810-2.
35. Pace CR, Stern GG. An approach to the measurement of the psychological characteristics of college environments. *J Educ Psychol* 1958;**49 (5)**:269-77.
36. Hutchins EB. The 1960 medical school graduate: his perception of his faculty, peers, and environment. *J Med Educ* 1961;**36**:322-9.
37. Rothman AI, Ayoade F. The development of a learning environment: a questionnaire for use in curriculum evaluation. *J.Med.Educ.* 1970;**45**:754-9.
38. Levy M, Morse PK, Liebelt RA, Dallman JJ, McDonald TF. Use of the learning environment questionnaire to assess curricular change. *J Med Educ* 1973;**48**:840-5.
39. Marshall RE. Measuring the medical school learning environment. *J Med Educ* 1978;**53**:98-104.
40. Feletti GI, Clarke RM. Review of psychometric features of the Medical School Learning Environment Survey. *Med Educ.* 1981;**15**:92-6.
41. Roff S, McAleer S, Harden RM, Al-Qahtani M, Ahmed AU, Deza H *et al*. Development and validation of the Dundee Ready Education Environment Measure (DREEM). *Med Teach* 1997;**19**:295-9.
42. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995;**311**:376-80.
43. Bassaw B, Roff S, McAleer S, Roopnarinesingh S, De LJ, Teelucksingh S *et al*. Students' perspectives on the educational environment, Faculty of Medical Sciences, Trinidad. *Med Teach.* 2003;**25**:522-6.
44. Pimparyon P, Roff S, McAleer S, Poonchai B, Pemba S. Educational environment, student approaches to learning and academic achievement in a Thai nursing school. *Med Teach* 2000;**22**:359-65.
45. Roff S, McAleer S, Ifere OS, Bhattacharya S. A global diagnostic tool for measuring educational environment: comparing Nigeria and Nepal. *Med Teach* 2001;**23**:378-82.
46. Roff S. The Dundee Ready Educational Environment Measure (DREEM)-a generic instrument for measuring students' perceptions of undergraduate health professions curricula. *Med.Teach.* 2005;**27**:322-5.
47. Varma R, Tiyagi E, Gupta JK. Determining the quality of educational climate across multiple undergraduate teaching sites using the DREEM inventory. *BMC.Med Educ* 2005;**5**:8.
48. Thrush CR, Hicks EK, Tariq SG, Johnson AM, Clardy JA, O'Sullivan PS *et al*. Optimal learning environments from the perspective of resident physicians and associations with accreditation length. *Acad.Med* 2007;**82**:S121-S125.

49. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Med.Teach.* 2005;**27**:326-31.
50. Afrin LB, Arana GW, Medio FJ, Ybarra AF, Clarke HS, Jr. Improving oversight of the graduate medical education enterprise: one institution's strategies and tools. *Acad.Med* 2006;**81**:419-25.
51. Holt MC, Roff S. Development and validation of the Anaesthetic Theatre Educational Environment Measure (ATEEM). *Med Teach.* 2004;**26**:553-8.
52. Kanashiro J, McAleer S, Roff S. Assessing the educational environment in the operating room-a measure of resident perception at one Canadian institution. *Surgery* 2006;**139**:150-8.
53. Nagraj S, Wall D, Jones E. The development and validation of the mini-surgical theatre educational environment measure. *Med Teach.* 2007;1-6.
54. Mulrooney A. Development of an instrument to measure the Practice Vocational Training Environment in Ireland. *Med Teach.* 2005;**27**:338-42.
55. Oliveira Filho GR, Vieira JE, Schonhorst L. Psychometric properties of the Dundee Ready Educational Environment Measure (DREEM) applied to medical residents. *Med.Teach.* 2005;**27**:343-7.
56. Pololi L, Price J. Validation and use of an instrument to measure the learning environment as perceived by medical students. *Teach.Learn Med* 2000;**12**:201-7.
57. Rotem A, Godwin P, Du J. Learning in hospital settings. *Teach Learn Med* 1995;**7**:211-7.
58. Beckman TJ, Cook DA, Mandrekar JN. What is the validity evidence for assessments of clinical teaching? *J Gen.Intern.Med* 2005;**20**:1159-64.
59. American Education Research Association and American Psychological Association. Standards for Educational and Psychological Testing. Washington DC: American Education Research Association, 1999.
60. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ.* 2004;**38**:1006-12.
61. Crossley J, Davies H, Humphris G, Jolly B. Generalisability: a key to unlock professional assessment. *Med Educ* 2002;**36**:972-8.
62. Field A. Exploratory Factor Analysis. *Discovering statistics using SPSS*, pp 619-80. London: Sage, 2005.
63. Seabrook MA. Clinical students' initial reports of the educational climate in a single medical school. *Med.Educ.* 2004;**38**:659-69.
64. Cleave-Hogg D, Rothman AI. Discerning views. Medical students' perceptions of their learning environment. *Eval Health Prof* 1990;**14**:456-74.
65. Rothman AI, Cleave-Hogg DM. A medical school learning environment: views from the trenches. *Teach Learn Med* 1990;**2**:12-9.
66. Kendall ML, Hesketh EA, Macpherson SG. The learning environment for junior doctor training--what hinders, what helps. *Med Teach* 2005;**27**:619-24.
67. Cross V, Hicks C, Parle J, Field S. Perceptions of the learning environment in higher specialist training of doctors: implications for recruitment and retention. *Med Educ* 2006;**40**:121-8.
68. Hoffman KG, Donaldson JF. Contextual tensions of the clinical environment and their influence on teaching and learning. *Med.Educ.* 2004;**38**:448-54.
69. Frye AW, Hoban JD, Richards BF. Capturing the complexity of clinical learning environments with multiple qualitative mehods. *Eval Health Prof* 1993;**16**:44-60.

2

Psychometric properties of an instrument to measure the clinical learning environment

Klarke Boor
Fedde Scheele
Cees van der Vleuten
Albert Scherpbier
Pim Teunissen
Klaas Sijtsma

Medical Education 2007; 41(1):92-99

ABSTRACT

Objectives

The clinical learning environment is an influential factor in work-based learning. Evaluation of this environment gives insight into the educational functioning of clinical departments. The Postgraduate Hospital Educational Environment Measure (PHEEM) is an evaluation tool consisting of a validated questionnaire with 3 subscales. In this paper we further investigate the psychometric properties of the PHEEM. We set out to validate the 3 subscales and test the reliability of the PHEEM for both clerks (clinical medical students) and registrars (specialists in training).

Methods

Clerks and registrars from different hospitals and specialities filled out the PHEEM. To investigate the construct validity of the 3 subscales, we used an exploratory factor analysis followed by varimax rotation, and a cluster analysis known as Mokken scale analysis. We estimated the reliability of the questionnaire by means of variance components according to generalisability theory.

Results

A total of 256 clerks and 339 registrars filled out the questionnaire. The exploratory factor analysis plus varimax rotation suggested a 1-dimensional scale. The Mokken scale analysis confirmed this result. The reliability analysis showed a reliable outcome for 1 department with 14 clerks or 11 registrars. For multiple departments 3 respondents combined with 10 departments provide a reliable outcome for both groups.

Discussion

The PHEEM is a questionnaire measuring 1 dimension instead of the hypothesised 3 dimensions. The sample size required to achieve a reliable outcome is feasible. The instrument can be used to evaluate both single and multiple departments for both clerks and registrars.

INTRODUCTION

Working and learning in the clinical environment represents a challenging phase for doctors in training. According to Daugherty et al. they "... must learn to balance such diverse demands as responsibility for patient care, economic hardships, on-call schedules, patient death, the need for constant learning, the task of teaching, the requirements of attending physicians and senior residents, along with the necessities of family and personal life."[1] This phase is further complicated by recent changes in legislation for working hours in Western Europe and the USA; the clinical workload has grown, whereas the time available for educational activities has diminished.[2-4] Meanwhile, the quality of health care attracts greater public attention.[5,6]

One important component of the educational experience is the clinical learning environment. This environment encompasses many important aspects, such as the quality of supervision[7,8], the quality of teachers[9,10], and facilities and atmosphere.[11,12] The Standing Committee on Postgraduate Medical Education (SCOPME) stated that "...a working environment that is conducive to learning is critically important to successful training."[13,14] The extent to which this is the case should be subject to evaluation. Such evaluation would allow us, for example, to assess the educational functioning of a single department. Evaluation of the learning environments in multiple hospitals is also valuable, as some studies suggest differences between types of hospitals (e.g. university-based versus non-university-based hospitals).[14,15]

Only a few instruments specifically assess the quality of the clinical learning environment. Roff et al. constructed and validated the Postgraduate Hospital Educational Environment Measure (PHEEM).[16] The developers of the questionnaire used a form of grounded theory involving focus groups, nominal groups and a Delphi panel drawn from the target population to validate the items of the PHEEM.[16,17] The 40-item questionnaire consists of items about the quality of teaching and content of work, but also takes into account social and emotional factors, such as being part of the team, quality of supervision, and working in a no-blame culture. The original authors identified 3 subscales which measured perceptions of role autonomy, perceptions of social support, and perceptions of teaching.[16,17] The items and their subscales are shown in table 1. The mean item score on the 40 items from the PHEEM represents an overall indicator of the quality of the learning environment. The mean item scores on the 3 subscales indicate strengths and weaknesses on 3 domains: autonomy, social support, and teaching. The investigated department or hospital may use these scores to stimulate improvements.

In this article we investigate 2 psychometric properties of the PHEEM. The first psychometric property is the construct validity of the 3 subscales. To our knowledge, no validation of these subscales has been published previously. The second property is the reliability of the ques-

Table 1: Items, subscales and descriptive statistics of the PHEEM for clerks and registrars

			Clerks (n= 256)			Registrars (n= 339)		
	Item	**Sub-scale***	**RR† (%)**	**M†**	**SD†**	**RR†**	**M†**	**SD†**
1	I have a contract of employment that provides information about hours of work	Aut	100	4.25	1.01	99.7	4.39	0.81
2	My clinical teachers set clear expectations	Teach	100	3.65	0.90	99.4	3.66	0.90
3	I have protected time at this post	Teach	100	3.94	1.20	99.4	3.52	1.13
4	I had an informative introduction programme	Aut	100	3.83	1.20	100	3.09	1.24
5	I have the appropriate level of responsibility in this post	Aut	100	4.00	0.89	100	4.13	0.73
6	I have good clinical supervision at all times	Teach	100	3.97	0.93	99.4	3.75	0.99
7	*There is racism in this post*	SocS	*98.9*	*4.66*	*0.83*	*99.4*	*4.80*	*0.62*
8	*I have to perform inappropriate tasks*	Aut	*100*	*3.92*	*1.01*	*99.1*	*3.96*	*1.13*
9	There is an informative Junior Doctors‡ hand-book	Aut	100	3.39	1.11	99.4	3.02	1.01
10	My clinical teachers have good communication skills	Teach	99.6	4.05	0.91	100	3.74	0.77
11	*I am bleeped inappropriately*	Aut	*95.7*	*4.04*	*1.02*	*99.4*	*3.35*	*1.12*
12	I am able to participate actively in educational events	Teach	97.7	4.01	1.01	99.1	4.10	0.76
13	*There is sex discrimination in this post*	SocS	*98.8*	*4.66*	*0.79*	*99.7*	*4.61*	*0.82*
14	There are clear clinical protocols in this post	Aut	99.6	3.53	1.12	100	3.81	1.00
15	My clinical teachers are enthusiastic	Teach	99.6	4.27	0.78	99.4	4.06	0.79
16	I have good collaboration with other doctors in my grade	SocS	96.5	4.36	0.73	99.7	4.42	0.66
17	My hours conform to the New Deal	Aut	97.7	3.46	1.22	99.7	3.55	1.20
18	I have the opportunity to provide continuity of care	Aut	98.9	3.49	1.03	100	3.49	1.13
19	I have suitability access to careers advice	SocS	98.4	3.18	1.04	100	3.27	0.95
20	This hospital has good quality accommodation for junior doctors, especially when on call	SocS	85.5	3.75	1.23	99.1	3.34	1.30
21	**There is access to an educational programme relevant to my needs**	**Teach**	**98.9**	**3.66**	**1.10**	**99.4**	**3.39**	**0.95**
22	I get regular feedback from my seniors	Teach	99.6	3.21	1.16	99.4	3.35	1.00

tionnaire, defined as reproducibility of data or scores, independent of time and occasion.[18] Variability and inconsistency among raters' personal opinions may, hence, negatively affect the instrument's reproducibility.[18,19] Our research goal is therefore to examine such influences on the PHEEM's reliability. The PHEEM can be used to measure clerks' and registrars' perceptions of their clinical learning environment. In our study clerks represent medical students who, after 4 years of pre-clinical medical education, enter 2 years of clinical rotations in all the major clinical disciplines. Registrars are specialists in training. For both groups we investigated the reliability of the PHEEM using 2 different analyses, each associated with a different

Table 1: Items, subscales and descriptive statistics of the PHEEM for clerks and registrars (continued)

	Item	Sub-scale*	RR† (%)	M†	SD†	RR†	M†	SD†
23	My clinical teachers are well organised	Teach	99.2	3.71	0.93	99.4	3.43	0.95
24	I feel physically safe within the hospital environment	SocS	99.6	4.04	1.17	99.7	3.73	1.28
25	There is a no-blame culture in this post	SocS	100	3.99	0.95	99.4	3.73	0.98
26	There are adequate catering facilities when I am on call	SocS	93.4	3.99	1.25	99.1	2.53	1.42
27	I have enough clinical learning opportunities for my needs	Teach	99.6	4.22	0.86	99.1	4.04	0.82
28	My clinical teachers have good teaching skills	Teach	98.8	4.07	0.91	99.4	3.67	0.79
29	I feel part of a team working here	Aut	100	3.85	1.00	99.4	4.08	0.84
30	I have opportunities to acquire the appropriate practical procedures for my grade	Aut	100	4.13	0.90	99.4	4.04	0.86
31	My clinical teachers are accessible	Teach	100	4.09	0.87	99.4	4.31	0.66
32	My workload in this job is fine	Aut	100	3.89	0.86	99.4	3.69	0.93
33	Senior staff utilise learning opportunities effectively	Teach	99.6	3.68	0.93	99.1	3.4	0.88
34	The training in this post makes me feel ready to be a SpR/ Consultant	Aut	100	4.05	0.838	99.4	3.93	0.77
35	My clinical teachers have good mentoring skills	SocS	100	3.79	0.96	99.1	3.58	0.87
36	I get a lot of enjoyment out of my present job	SocS	100	4.29	0.86	99.4	4.32	0.71
37	My clinical teachers encourage me to be an independent learner	Teach	100	3.72	0.95	99.4	3.58	0.91
38	There are good counselling opportunities for junior doctors‡ who fail to complete their training satisfactorily	SocS	94.1	2.92	0.74	97.6	2.72	0.92
39	The clinical teachers provide me with good feedback on my strengths and weaknesses	Teach	99.2	3.18	1.07	99.1	3.21	0.97
40	My clinical teachers promote an atmosphere of mutual respect	Aut	100	3.87	0.98	99.4	3.74	0.98

The *Italic* items have recoded scores: they are inverted.

*Three subscales: perceptions of autonomy (Aut), perceptions of social support (SocS) and perceptions of teaching (Teach)

† Response rate (RR), Mean item score (M) and Standard Deviation (SD) for every item.

‡ In the questionnaire we used the appropriate word (so either clerk or registrar)

use of the PHEEM. Firstly, we used the PHEEM to evaluate a single department. Secondly, we used the PHEEM to evaluate a group of departments for the purposes of, for example, comparison across hospitals.

This process referred to the following research questions:

- What is the construct validity of the 3 subscales of the PHEEM (i.e. perceptions of autonomy, social support, and teaching)?

- How many ratings by different *clerks* are necessary to achieve a reliable score representing the learning environment of an individual department?
- How many ratings by different *registrars* are necessary to achieve a reliable score representing the learning environment of an individual department?
- How many *clerk* ratings and departments are needed to achieve a reliable score representing the learning environment of a group of different departments or hospitals?
- How many *registrar* ratings and departments are needed to achieve a reliable score representing the learning environment of a group of different departments or hospitals?

METHODS

Instrument

With the authors' permission, we translated the PHEEM into Dutch. A professional translator then translated this version back into English. The original authors considered this version equivalent to the original questionnaire. Each subject (clerks and registrars) scored the 40 items on a 5-point Likert scale, where 1 = totally disagree and 5 = totally agree. (The original questionnaire used a 5-point Likert scale of 0-4, which we replaced with the more conventional 1-5 range). Because 4 items contained negative statements (items 7, 8, 11, and 13), we inverted the score on the scale. Clerks and registrars received the exact same questionnaire, except for the use of specific words such as 'clerk' and 'registrar'.

Subjects and Procedure

Clerks from 14 different departments (including internal medicine, surgery, obstetrics and gynaecology, paediatrics, neurology, and psychiatry) in 6 different hospitals filled out the PHEEM between April 2003 and May 2005. As clerks had to be able to assess the clinical learning environment, we evaluated their perceptions of this environment in the second half of their clerkship.
Paediatrics registrars from 25 hospitals and obstetrics and gynaecology registrars from 44 hospitals completed the questionnaire during March-April 2005.

Statistical analysis

After checking the normality of the distribution of PHEEM scores we assumed an interval level of the data and further used parametric statistical methods.

Exploratory factor analysis

To evaluate the construct validity of the 3 subscales of the PHEEM, we used an exploratory factor analysis (specifically, principal components analysis) followed by varimax rotation. Exploratory factor analysis enables us to determine whether the observed variables (i.e. the items) can be explained by a considerably smaller number of factors.[20] Principal components analysis calculates zero-correlating factors (called orthogonal components) to maximise explained variance from the items and thus summarises the statistical information in the items as efficiently as possible. Next, we performed a varimax rotation on these selected factors to obtain a solution that had optimal interpretation in terms of the correlations (in this context known as 'loadings') of each of the items with each of the rotated factors. We interpreted the results with a scree-plot of the eigenvalues.

We checked the results of the exploratory factor analysis by means of a successive clustering method, which is known in psychometrics as Mokken scale analysis.[21,22] This method selects items that measure the same construct into clusters and thus can be used to determine the dimensionality of the PHEEM data. A careful comparison of exploratory factor analysis and Mokken scale analysis revealed that these methods provide different perspectives on the dimensionality in data. For example, exploratory factor analysis considers all items simultaneously, whereas Mokken scale analysis selects items one after another. Likewise, exploratory factor analysis aims at maximising explained variance, whereas Mokken scale analysis optimises a psychometric scalability criterion. However, despite their differences, these methods lead to the same conclusions when a dimensionality structure is clearly present.[23]

Generalisability theory

We used generalisability theory to address the research questions about reliability. This theory allows estimation of the size of the relevant influences that affect the measurement. The subsequent estimation of the reliability of the instrument is based on a variety of reliability indices. Here reliability is expressed as the standard deviation (SD) of the 'noise in the measurement', i.e. the SD of all influences that have a random or noisy effect on the measurement (noisy as in signal-versus-noise). We considered items to be a fixed facet and used the PHEEM total (subscale) score as the unit for analysis. We carried out a random-effects ANOVA model with 2 factors for clerks and registrars separately. The factors were departments (d) and subjects (s). In generalisability theory terms, we carried out a single-facet analysis with subjects nested within departments, separately for clerks and registrars. An unbalanced design using the UrGenova program estimated variance components.[24] Following variance component estimation, we estimated the standard error of measurement (SEM), again separately for clerks and registrars. The formula used to provide information on a single department was:

$$SEM = \sqrt{\frac{\sigma^2_{s:d}}{N_s} + \frac{\sigma^2_{si:d}}{N_s \times N_i}}$$

in which $\sigma^2_{s:d}$ is the variance associated with subjects within departments and $\sigma^2_{si:d}$ represents the interaction between subjects and items within departments. Both variance components are divided by the sample size associated with the component.
The SEM can be interpreted on the original scoring scale and helps to define a maximum acceptable noise in the measurement. In this study we wanted a difference of at least half a unit on the scale to be interpretable. We therefore used an SEM < 0.13 (1.96 x 0.13 x 2 ≈ 0.5) as the smallest admissible value for a 95% confidence interval interpretation.

To use the PHEEM across a group of departments, we estimated the root mean squared error (RMSE) which can be interpreted in the same way as the SEM but now at the group level:

$$RMSE = \sqrt{\frac{\sigma^2_d}{N_d} + \frac{\sigma^2_{s:d}}{N_s \times N_d} + \frac{\sigma^2_{si:d}}{N_s \times N_i \times N_d}}.$$

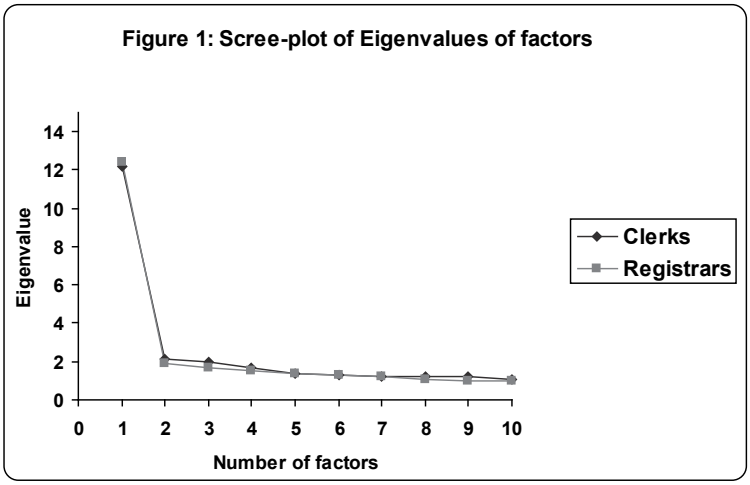
We carried out these reliability estimation procedures for the mean item score of the PHEEM and for each of the subscales.

RESULTS

The PHEEM was completed by a total of 256 clerks, of whom 80 (31%) were male. They came from 14 departments; the number of clerk ratings within departments ranged from 2 to 26. The questionnaire was also filled out by a total of 339 registrars, of whom 83 (25%) were male. They came from 45 departments; the number of registrars within departments ranged from 2 to 24. Table 1 shows the response rate, descriptive statistics and mean item score for both groups. We found no significant difference between the answers of men and women.

Construct validity of the three subscales

Exploratory factor analysis followed by varimax rotation of the clerk group resulted in 10 factors with an eigenvalue > 1. The first factor had an eigenvalue of 12.2 (accounting for 30.6 % of variance), and the next 9 factors had eigenvalues < 2.1 (scree plot in figure 1). The analysis of the registrar group showed 9 factors with eigenvalues > 1. The first factor had an eigenvalue of 12.4 (accounting for 31.1% of variance), and the following 8 had eigenvalues < 1.9 (scree plot in figure 1).



These findings are not consistent with a questionnaire measuring 3 distinct factors. In such a case, the results would show 3 factors with relatively high eigenvalues (which would preferably together account for a sizable percentage of the variance). The results, however, suggest 1 factor and thus a 1-dimensional scale. Next, we performed a Mokken scale analysis on both data sets. The results confirmed the factor analysis results: 1 large item cluster was found indicating a 1-dimensional scale.

As 2 independent statistical analysis methods supported a unidimensional data structure and we found no support of the existence of 3 subscales, we present only the results of the reliability analysis with the mean item score.

Reliability analysis

Clerks

The mean item score was 3.87. The score varied from 2.92 (item 38: 'There are good counselling opportunities for junior doctors who fail to complete their training satisfactorily') to 4.66 (inverted score [originally 1.34] for both item 7: 'There is racism in this post' and item 13: 'There is sex discrimination in this post'). Response rate varied from 85.8% (item 20: 'This hospital has good quality accommodation for junior doctors, especially when on call') to 100% (table 1).

Table 2 presents our estimated SEMs and RMSEs for clerks. The upper part of table 2 presents SEMs for the evaluation of one department. The SEM reached a reliable level < 0.13 when ≥14 respondents completed the PHEEM.

The reliability of an evaluation of multiple departments (lower part of table 2) depends on the number of respondents and departments. An RMSE < 0.13 could be established with 15 departments and 2 respondents. Ten departments and 3 respondents also give a reliable

Table 2: Clerks: Standard error of measurement (SEM) for individual scoring of and Root Mean Square Error (RMSE) for group evaluation

	N respondents											
	1	**2**	**3**	**4**	**5**	**10**	**11**	**12**	**13**	**14**	**15**	**20**
SEM	0.47	0.33	0.27	0.23	0.21	0.15	0.14	0.13	0.13	**0.12***	0.12	0.10
N departments	**RMSE**											
1	0.53	0.41	0.36	0.34	0.32	0.29	0.28	0.28	0.28	0.28	0.27	0.27
2	0.37	0.29	0.26	0.24	0.23	0.20	0.20	0.20	0.20	0.19	0.19	0.19
3	0.30	0.24	0.21	0.20	0.19	0.17	0.16	0.16	0.16	0.16	0.16	0.15
4	0.26	0.21	0.18	0.17	0.16	0.13	0.13	0.13	**0.12***	0.12	0.12	0.12
5	0.24	0.18	0.16	0.15	0.14	0.13	0.13	0.13	**0.12***	0.12	0.12	0.12
10	0.17	0.13	**0.12***	0.11	0.10	0.09	0.09	0.09	0.09	0.09	0.09	0.08
15	0.14	**0.11***	0.09	0.09	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.07
20	**0.12***	0.09	0.08	0.08	0.07	0.06	0.06	0.06	0.06	0.06	0.06	0.06

* Value < 0.13 is considered reliable

result. By contrast, 1 department cannot achieve a reliable outcome unless the number of respondents is unfeasibly high. Clearly, when evaluating a group of departments, it is more efficient to enlarge the number of departments than the number of respondents.

Registrars

The mean item score was 3.71. The score varied from 2.53 (item 26: 'There are adequate catering facilities when I am on call') to 4.80 (inverted score [originally 1.20] item 7: 'There is racism on this post'). Response rate varied between 97.6% (item 38: 'There are good counselling opportunities for junior doctors who fail to complete their training satisfactorily') and 100%. Table 3 shows our estimated SEMs and RMSEs for registrars. A reliable evaluation of the clinical learning environment of 1 department could be achieved with ≥ 11 respondents. For a reliable outcome of group evaluation of multiple departments the easiest option is to increase the number of departments rather than the number of respondents. Three respondents and 10 departments give a reliable result.

DISCUSSION

This study investigated the construct validity of 3 subscales and the reliability of an instrument to measure the clinical learning environment, known as the PHEEM. Clerks and registrars filled out the questionnaire. The first research question addressed the construct validity of 3 subscales, as hypothesised by the original designers of the PHEEM. The statistical analysis of

Table 3: Registrars: Standard error of measurement (SEM) for individual scoring of and Root Mean Square Error (RMSE) for group evaluation

	N respondents											
	1	**2**	**3**	**4**	**5**	**10**	**11**	**12**	**13**	**14**	**15**	**20**
SEM	0.40	0.28	0.23	0.20	0.18	0.13	**0.12***	0.11	0.11	0.11	0.10	0.09
N departments	**RMSE**											
1	0.50	0.41	0.38	0.36	0.35	0.32	0.32	0.32	0.32	0.32	0.32	0.31
2	0.35	0.29	0.27	0.25	0.24	0.23	0.23	0.23	0.22	0.22	0.22	0.22
3	0.29	0.24	0.22	0.21	0.20	0.19	0.18	0.18	0.18	0.18	0.18	0.18
4	0.25	0.20	0.19	0.18	0.17	0.16	0.16	0.16	0.16	0.16	0.16	0.16
5	0.22	0.18	0.17	0.16	0.15	0.14	0.14	0.14	0.14	0.14	0.14	0.14
10	0.16	0.13	**0.12***	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.10	0.10
15	0.13	**011***	0.10	0.09	0.09	0.08	0.08	0.08	0.08	0.08	0.08	0.08
20	**0.11***	0.09	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.07

* Value < 0.13 is considered reliable

these subscales did not support the 3-dimensional structure hypothesised earlier.[16] Instead, our analysis suggested a 1-dimensional scale. Apparently the content analyses of the PHEEM as performed by the original authors cannot be replicated empirically.

The second research question focused on the number of respondents necessary to achieve a reliable evaluation of the clinical learning environment. Clerks can establish a reliable score with 14 completed questionnaires. Registrars need 11 evaluations to get a reliable result.

The third research question assesses the number of respondents and departments needed to obtain a reliable outcome for a group of departments or hospitals. The number is the same for both clerks and registrars: for 10 departments, 3 questionnaires per department are needed. For both groups it is more efficient to improve the reliability by increasing the number of departments rather than the group of respondents.

We used 256 and 339 filled-out questionnaires, respectively, for this study. These numbers are high enough to perform a reliable exploratory factor analysis and a Mokken scale analysis. Thus, our finding of a 1-dimensional construct as measured by the PHEEM seems plausible. The number of questionnaires is also large enough to give a good estimation of the PHEEM's reliability. By contrast, the different specialties and hospitals are not represented equally. Among the 45 different hospitals included in our study, we investigated only paediatrics, and obstetrics and gynaecology registrars. Clerks were mainly derived from 1 hospital and 2 specialties (obstetrics and gynaecology, and internal medicine). For widespread application of the PHEEM, further research among other specialties in different countries is necessary.

The statistical boundaries we used were rather strict. We chose a standard error < 0.13 as cut-off point, whereas some other studies settled for 0.24.[25,26] Thus, the reliability of this instrument is high.

This study is part of an ongoing effort to understand and possibly influence the clinical learning environment. We consider this research into the reliability and construct validity of the PHEEM to represent a starting point for further research. Because we found only 1 construct underlying the PHEEM, it would be of interest to investigate what exactly constitutes the clinical learning environment: in other words, what is the content validity of the PHEEM? Further research should focus on this psychometric property, as well as on evaluation of clinical learning environments within different hospitals and departments.

The PHEEM is a 1-dimensional, reliable questionnaire for measuring the clinical learning environment for both clerks and registrars. Reliable findings can be accomplished with feasible sample sizes. It is remarkable how stable the findings are, given the high turnover of clerks and, to a lesser extent, registrars. Results offer insight into the existing clinical learning environment created by 1 or multiple departments.

Acknowledgements

The authors thank Thomer Gil for his continuous efforts in revising the original text to fit the demands of English grammar and style.

REFERENCES

1. Daugherty SR, Baldwin DC, Jr., Rowley BD. Learning, satisfaction, and mistreatment during medical internship: a national survey of working conditions. JAMA 1998; 279(15):1194-1199.
2. Hoff TJ, Pohl H, Bartfield J. Creating a learning environment to produce competent residents: the roles of culture and context. Acad Med 2004; 79(6):532-539.
3. Ko CY, Escarce JJ, Baker L, Sharp J, Guarino C. Predictors of surgery resident satisfaction with teaching by attendings: a national survey. Ann Surg 2005; 241(2):373-380.
4. Baldwin PJ, Newton RW, Buckley G, Roberts MA, Dodd M. Senior house officers in medicine: postal survey of training and work experience. BMJ 1997; 314(7082):740-743.
5. Mitchell M, Srinivasan M, West DC, Franks P, Keenan C, Henderson M et al. Factors affecting resident performance: development of a theoretical model and a focused literature review. Acad Med 2005; 80(4):376-389.
6. Robinson AR, Hohmann KB, Rifkin JI, Topp D, Gilroy CM, Pickard JA et al. Physician and public opinions on quality of health care and the problem of medical errors. Arch Intern Med 2002; 162(19):2186-2190.
7. Kilminster SM, Jolly BC. Effective supervision in clinical practice settings: a literature review. Med Educ 2000; 34(10):827-840.
8. Cottrell D, Kilminster S, Jolly B, Grant J. What is effective supervision and how does it happen? A critical incident study. Med Educ 2002; 36(11):1042-1049.
9. Parsell G, Bligh J. Recent perspectives on clinical teaching. Med Educ 2001; 35(4):409-414.
10. Irby DM. Clinical teacher effectiveness in medicine. J Med Educ 1978; 53(10):808-815.
11. Rotem A, Bloomfield L, Southon G. The clinical learning environment. Isr J Med Sci 1996; 32(9):705-710.
12. Bleakley A. Pre-registration house officers and ward-based learning: a 'new apprenticeship' model. Med Educ 2002; 36(1):9-15.
13. Standing Committee on Postgraduate Medical Education. Good Practice in SHO Training. London: SCOPME; 1991.
14. Parry J, Mathers J, Al Fares A, Mohammad M, Nandakumar M, Tsivos D. Hostile teaching hospitals and friendly district general hospitals: final year students' views on clinical attachment locations. Med Educ 2002; 36(12):1131-1141.
15. Rotem A, Godwin P, Du J. Learning in hospital settings. Teach Learn Med 1995; 7:211-217.
16. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. Med Teach 2005; 27(4):326-331.
17. Skinner AM. Development, validation and application of a questionnaire to study the perceived learning environment of hospital based junior doctors in training. Masters of Medical Education Thesis, Centre for Medical Education, University of Dundee; 2002.
18. Downing SM. Reliability: on the reproducibility of assessment data. Med Educ 2004; 38(9):1006-1012.
19. van der Vleuten CP. The assessment of professional competence: developments, research and practical implications. Adv Health Sci Educ 1996; 1:41-67.
20. Norman GR, Streiner DL. Exploratory factor analysis. PDQ. Third ed. London: BC Decker inc; 2003: 144-155.
21. Mokken RJ. A theory and procedure of scale analysis. Berlin: De Gruyter; 1971.

22. Sijtsma K, Molenaar IW. Introduction to nonparametric item response theory. Thousand Oaks, CA: Sage; 2002.
23. Scheirs JGM, Sijtsma K. The study of crying: Some methodological considerations and a comparison of methods for analyzing questionnaires. In: Vingerhoets AJJM, Cornelius RR, editors. Adult crying. A biopsychosocial approach. Hove, UK: Brunner-Routledge; 2001: 277-298.
24. Brennan RL. Manual for UrGenova. Iowa Testing Programs Occasional Paper Number 46. Version 1.4 ed. Iowa City, IA: The University of Iowa; 1999.
25. van der Hem-Stokroos HH, van der Vleuten CP, Daelmans HE, Haarman HJ, Scherpbier AJ. Reliability of the clinical teaching effectiveness instrument. Med Educ 2005; 39(9):904-910
26. De Grave WS, Dolmans DH, van der Vleuten CP. Tutor intervention profile: reliability and validity. Med Educ 1998; 32(3):262-268.

3

Residents' perceptions of the ideal clinical teacher- a qualitative study

Klarke Boor
Pim Teunissen
Albert Scherpbier
Cees van der Vleuten
Jonas van de Lande
Fedde Scheele

European Journal of Obstetrics & Gynecology and Reproductive Biology 2008; 140(2):152-157

ABSTRACT

Objective

The object of this study was to establish what residents in 1994 and 2003 characterised as an ideal clinical teacher and whether differences existed between residents' views in 1994 and 2003.

Study design

Setting: postgraduate medical education in the Netherlands.
Subjects: 207 obstetric-gynaecologic residents.
Intervention: open-ended questionnaire.
Analysis: qualitative data analysis with two coding dictionaries based on current literature. Differences between 1994 and 2003 were estimated using the chi-square test.

Results

Residents preferred the 'person' role both in 1994 (42%) and in 2003 (48%). The 'physician' role was significantly more important in 1994 than in 2003; the 'supervisor' role was significantly more important in 2003 than in 1994 ($p < 0.01$). Seventy percent of the comments related to 'direct interaction' (i.e. between residents and clinical teachers), 30% to 'indirect interaction' (i.e. clinical teachers' behaviour affecting residents indirectly).

Conclusion

The data showed that almost half of residents' comments described 'person' role characteristics. There was a significant shift in the role ranked second, from the 'physician' role in 1994 to the 'supervisor' role in 2003. The findings highlighted that teachers, in order to be perceived as ideal, should adapt their behaviour to residents' learning needs.

INTRODUCTION

Residency typically involves apprenticeships in university and non-university based hospitals where a resident delivers patient care under progressively diminishing supervision of a clinical teacher.[1] The relationship between residents and faculty members is of great importance for the clinical learning environment. According to Paice "...the quality of that relationship can make the difference between a post that is rewarding and one that is demoralising."[2]

Most research on the concept of the ideal clinical teacher relates to personal traits he or she should have.[3-11] In these studies students, residents, and sometimes clinical teachers themselves filled out questionnaires or took part in interviews to define the ideal clinical teacher. Some studies were concerned with general practitioners, anaesthetists, or surgeons, and one older study described characteristics of gynaecologists, rated by medical students.[3-7,10] A compelling paper was written by Ullian et al.[9] They describe four roles based on the literature and on research of residents' perceptions of their clinical teachers. These roles are: "..'physician' (models knowledge and skills in performing medical duties), 'supervisor' (provides opportunities for performance, observes, gives feedback), 'teacher' (selects, organises, and delivers information), and 'person' (exhibits certain interpersonal and intrapersonal characteristics)".[9] A study among medical students on their opinion of clinical teacher characteristics confirms this pattern of roles.[8]

In our view, the aforementioned literature on teachers is predominantly written from a cognitivistic perspective.[12] Cognitivism focuses on individuals and their mental activities. Knowledge can be seen as consisting of schemata or symbolic mental constructions.[13] Teachers play a central role by, for instance, organising knowledge. In recent years, however, there has been growing interest in social cultural theories of learning.[12;14-16] Social cultural frameworks emphasise the importance of the social context and relations in which learning takes place.[17-19] Swanwick states "..it takes two to tango", stressing the importance of reciprocal interaction between learner (i.e. resident) and workplace environment (i.e. residency).[12] Similarly, people and personalities are constructed in relation to others; as Burr states "..one way of looking at this is to think of personality (..) as existing not within people but between them. Take some of the personality-type words we use to describe people: for example, friendly, caring, shy, self conscious, charming, bad-tempered, thoughtless (..) words which would completely lose their meaning if the person described lived on a deserted island. The point is we use these words as if they refer to entities within the person they describe but once the person is removed from their relationship with others the words become meaningless."[20] We hypothesised that whether or not a clinical teacher is perceived as ideal depends on the interaction of a resident with his/her teacher in a certain context.

The clinical learning environment, i.e. the context, has changed over the past 15 years. Shifts in the organisation and delivery of health care make residency an ever-changing endeavour, apart from the challenges posed by evolution of medicine itself. A significant change, for example, are limits on the number of hours residents are allowed to work.[21;22] This leads to additional pressure on quality of patient care and creates fewer possibilities for training.[1;12] Moreover, residents spent fewer hours with clinical teachers. This may have led to changes in residents' requirements of an ideal clinical teacher.

In this study we asked obstetric-gynaecologic residents in 1994 and 2003 which characteristics they value most in clinical teachers. The first objective of this study is to take the current literature as a point of departure and describe characteristics of clinical teachers favoured by residents. The second objective is to complement this point of view using a socio-cultural approach as described above. Our third objective is to establish whether there are differences in residents' perceptions of ideal clinical teachers between 1994 and 2003.

MATERIAL AND METHODS

Multiple quantitative questionnaires measuring clinical teachers' effectiveness or quality exist. However, to answer our research questions a qualitative research method is most appropriate. This kind of research enables participants to express their personal views without forcing them to choose predefined answers. In 1994 and 2003 we asked all obstetric-gynaecologic residents in the Netherlands to fill out a questionnaire containing three open-ended questions. In this study we report the responses to the first question on the questionnaire: "Which three characteristics should an ideal clinical teacher have?" The other two questions were: "What do you miss in your clinical teachers?" and "What suggestions do you have for improving residency training?". The second and the third question were aimed at local quality improvement and they did not add to the current research question. Therefore, we did not include the answers to the last two questions in our analysis. The form did not include a section on resident's year of training or sex.

Participants

In 1994, 74 obstetric-gynaecologic residents (62% of a total of 120 obstetric-gynaecologic residents in the Netherlands at the time) filled out the questionnaire. In 2003, 133 residents (55% of a total of 240 obstetric-gynaecologic residents) filled it out. In both cases, we sent a questionnaire to the home address of every resident and we requested them to return the filled out form. Since we included every obstetric-gynaecologic resident in the Netherlands, residents with varying experience levels received the questionnaire. All completed forms in

1994 were readable; they contained 248 'units of analysis'. A unit of analysis is every word, line or phrase that expresses a single characteristic of a clinical teacher. Out of 133 questionnaires in 2003, 13 were not readable due to photocopying problems and therefore, excluded from analysis, making a total of 120 questionnaires. These questionnaires contained 316 units of analysis.

Analysis of qualitative data

According to Miles and Huberman the analysis of qualitative data roughly consists of three concurrent 'flows of activity': data reduction (i.e. classifying unstructured data into coding categories for retrieval and organising purposes), data display (i.e. using matrices, charts, etc., both for data reduction as for explaining and seeing 'the bigger picture'), and conclusion drawing and verification (for instance making conceptual coherence and checking for researcher effect).[23]

As a means of data reduction we constructed two coding dictionaries. The first author (KB) created the first coding dictionary using the four roles described in Ullian's study ('physician', 'supervisor', 'teacher', and 'person').[9] She coded the 564 units of analysis into 64 codes that were each placed in one of 18 subcategories within Ullian's four main roles. After revising the coding dictionary by removing code duplicates, the final coding dictionary contained 59 codes within 15 subcategories (for data display, see table 1). We used the chi-square test to estimate whether significant differences existed for the four roles' frequency between 1994 and 2003. The second coding dictionary covered the content analysis from a socio-cultural point of view. Social interactions play an important role within this epistemology. Thus, we performed a content analysis of the presence of resident-clinical teacher interaction. We divided the content analysis into two categories: first 'direct interaction' (i.e. between residents and clinical teachers), second 'indirect interaction' (i.e. clinical teachers' behaviour affecting residents indirectly).

The first author (KB) constructed the coding dictionaries. To test the accuracy of both coding dictionaries (verification testing) the second author (PT) coded a random sample of 100 responses using the dictionaries. KB and PT compared their codes to check for possible differences. The kappa inter-rater agreement coefficient was 0.81 for the first coding dictionary and 0.82 for the second. Different codes were discussed until they reached consent. Subsequently KB coded the rest of the responses twice, using both the first and the second coding dictionary.

RESULTS

We present our results in accordance with both steps described in the material and methods section. First we describe the analysis according to Ullian et al.'s categories[9], then we present our content analysis of resident-clinical teacher interaction.

Ullian et al.'s categories

In this section we describe our analysis of the data using the four roles described by Ullian et al.[9] as template. Within these four roles we have recognised 15 subcategories (table 1).

According to residents, the most important role of a gynaecologist as a clinical teacher was 'person'. In 1994, 41.9% of residents' answers related to the 'person' role. This was not significantly different from 2003, when 47.9% of the responses related to the 'person' role. In 1994, the 'physician' role was the second most important role for obstetric-gynaecologic residents

Table 1: Number of phrases within every role and subcategory

	1994	2003	Type of interaction
Physician	**51*(20.5 %)**	**31* (9.8 %)**	
Expert	33	23	Indirect
Role model	5	5	Indirect
Up-to-date	13	3	Indirect
Person	**104 (41.9 %)**	**151 (47.9 %)**	
Commitment	23	40	Direct
Supportive	25	26	Direct
Trustworthy	10	35	Direct
Organiser	15	15	Indirect
Dialogue	10	14	Direct
Personality	21	21	Direct
Supervisor	**43*(17.3 %)**	**86* (27.2 %)**	
Approachable	17	31	Direct
Stimulates	15	17	Direct
Coaches	11	38	Direct
Teacher	**50 (20.1 %)**	**48 (15.2 %)**	
Didacticism	26	15	Direct
Vision on training	7	8	Indirect
Facilitates training	17	25	Indirect

* p< 0.01, using the Chi-Square test

(20.5%), closely followed by the 'teacher' role (20.1%), and the 'supervisor' role (17.3%). In 2003 the situation was notably different: the second most important role was the 'supervisor' role (27.2%), followed by the 'teacher' role (15.2 %), and the 'physician' role (9.8%). The number of units relating to the 'physician' role differed significantly between 1994 and 2003, as did the number of units concerning the 'supervisor' role ($p < 0.01$). Table 2 shows the 15 subcategories with illustrative quotes.

Table 2: Subcategories with illustrative quotes

Person	
Commitment	"should be committed to and interested in the residents"
Support	"expresses empathy"
Trustworthiness	"is open and honest, does not talk behind people's back"
Organising qualities	"being available rather than attending international conferences"
Open dialogue	"is receptive to criticism and new ideas"
Personality	"charisma!"
Physician	
Expertise	"is knowledgeable and skilful"
Role modelling	"inspires through their interaction with patients"
Being up-to-date	"is aware of the current trends and treatments"
Supervisor	
Approachable	"is approachable (both figurative and literally)"
Stimulates	"stimulates development: both in patient care as in research activities"
Coaches	"gives appropriate feedback and does not say: 'you are doing it completely wrong, idiot!'"
Teacher	
Didactic skills	"knows how to teach"
Vision on training	"has distinctive ideas about the education of doctors in training"
Facilitates training	"prioritises education over production"

Resident-clinical teacher interaction

In this section we describe the analysis of our data from an socio-cultural point of view. We recognised two major types of interaction in our data. Seventy percent of the units of analysis related to direct interaction (see table 1). Interestingly, we observed substantial variations between the comments of different residents, although all remarks related to direct interaction. One resident, for example, wanted a clinical teacher who allowed a lot of autonomy,

while another wanted more guidance. One resident preferred a personal relationship with a clinical teacher, another favoured a more distant relationship.
Thirty percent of the comments concerned indirect interaction. One aspect of indirect interaction was related to organising an inspiring learning environment. For example, residents wanted their clinical teacher involved in their work schedule, in the attitude of other staff members towards teaching, and in prioritising education vs. production. Another vital aspect was the effect of a clinical teacher interacting with his/her patients in an inspiring way.

COMMENT

What characterises an ideal clinical teacher? We used two different approaches for categorising our data. First, we followed Ullian et al.'s categorisation of the ideal clinical teacher.[9] Both in 1994 and in 2003 residents named the 'person' role as the most important role of an ideal clinical teacher. In 1994 residents valued the 'physician' role significantly more than their colleagues of 2003. Residents of 2003 valued the 'supervisor' role significantly more than residents in 1994. The ranking of the 'teacher' role did not change.
We also interpreted the data from a socio-cultural point of view. Seventy percent of all comments related to direct resident-clinical teacher interaction. Thirty percent of residents' comments related to indirect interaction; a clinical teacher could express this through organising skills, for example.

Our study has several strengths. First, we highlighted a different view on the concept of the ideal clinical teacher. Besides the categorisation described by Ullian et al., we introduced an analysis based on socio-cultural theories. The results of both analyses complemented each other: many high-valued 'person' characteristics were in addition necessary traits for teachers' meaningful interaction with residents and their learning needs. Second, this study is based on qualitative research. This resulted in information-rich data, with refreshing views on the concept of the ideal clinical teacher. Third, we used two datasets over time. The fact, for instance, that the importance of the 'person' role was repeated after an interval of ten years, adds to the stability and validity of our outcomes. Finally, this study focuses specifically on obstetric-gynaecologic residents. To our knowledge no other qualitative study on characteristics of ideal gynaecologists-as-teachers has been published.
Some limitations need to be considered when interpreting this study. First, generalisability might be compromised due to the response rate of around 60%. Second, the study was done in the Netherlands only. However, we included all teaching hospitals in the Netherlands so the responses came from many different departments. A third weakness is the lack of teacher responses. For a complete view of ideal clinical teachers, their opinion should be included. Fourth, since we did not include a section in our questionnaire asking for level of experi-

ence or sex, we cannot explain our findings from that point of view. This also weakens the comparability of the two resident cohorts.

Ullian et al. have described four roles based on residents' perceptions of clinical teachers. In their study the 'teacher' role was mentioned most often.[8;9] In our study, however, the 'person' role stood out as most important. The other three roles occurred in a changing distribution scheme over time. In 1994 the residents described aspects relating to the 'physician' role more frequently than their colleagues in 2003. Remarks relating to the 'supervisor' role were mentioned more often in 2003 than in 1994. One possible explanation might be found in changing perspectives towards teaching and learning. There is a shift from teacher-centred to learner-centred education.[24] The 'physician' role is associated with a teacher-centred approach: focus is on the (medical) expertise of the clinical teacher as a role model for residents. The 'supervisor' role, however, is more associated with a learner-centred approach: the clinical teacher coaches and guides the resident. Another possible explanation for this statistically significant difference is that residents work less due to limits on working hours, but still have to do the same amount of clinical work. This reduces time spent on observing role-models and increases their need for supervisional activities. A recent study underlines residents' current need for supervision: the odds ratio for resident burn-out increased by 2.1 (95% CI 1.3-2.9) in the group with a greater workload combined with unsupervised practice.[25] Because the results give only limited insight in what caused this change of preference, further research is necessary.

The 'person' role, like the other three roles, related strongly to direct interaction, which seemed omnipresent in this context. Literature on supervision[26] and learning in the clinical workplace [12;14;15;18] underlines the importance of interaction. In literature on ideal clinical teachers attention for the interactive nature of clinical teaching is sparse. The significance of interaction is further stressed by the diverse and sometimes even contradictory responses about valued characteristics residents gave in our study. The differences may be caused by individual differences between residents as well as varieties in residents' years of experience. In any case, an ideal clinical teacher should master a diverse pallet of strategies to guide residents. Personal qualities, a flexible attitude, and organising a stimulating learning environment seem desired traits of the clinical teacher. However, in staff development programs the predominant focus is on teaching and role-modelling skills and less so on the vital interaction between resident and clinical teacher. We argue for more awareness of these important traits. Clinical teachers need to invest significantly in their own teaching skills. Ericsson describes this investment in expertise development as part of 'deliberate practice theory'.[1;27] Central to this theory is the notion that the acquisition of expertise requires a continuous drive to improve, participation in challenging activities, and receptiveness to immediate and informative feedback.[1;27] Ericssons' theory stresses the vital role of interaction, for the clinical teacher

should be receptive to residents' feedback in order to reach a state of proficiency. In sum, interaction with, and personal interest in, residents can lead to more content residents on the one hand, and to more proficient clinical teachers on the other.

Acknowledgements

The authors thank all residents for filling out the questionnaire. We thank Andreas Thurkow and Marion Heres for critically reviewing the manuscript. We also thank Thomer M. Gil for his invaluable support in English grammar and style.

REFERENCES

1. Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: a tradition worth defending? Acad Med 2005; 80(10 Suppl):S106-S111.
2. Paice E, Moss F, Heard S, Winder B, McManus IC. The relationship between pre-registration house officers and their consultants. Med Educ 2002; 36(1):26-34.
3. Boendermaker PM, Schuling J, Meyboom-de Jong BM, Zwierstra RP, Metz JC. What are the characteristics of the competent general practitioner trainer? Fam Pract 2000; 17(6):547-553.
4. Buchel TL, Edwards FD. Characteristics of effective clinical teachers. Fam Med 2005; 37(1):30-35.
5. Cleave-Hogg D, Benedict C. Characteristics of good anaesthesia teachers. Can J Anaesth 1997; 44(6):587-591.
6. Cox SS, Swanson MS. Identification of teaching excellence in operating room and clinic settings. Am J Surg 2002; 183(3):251-255.
7. Gjerde CL, Coble RJ. Resident and faculty perceptions of effective clinical teaching in family practice. J Fam Pract 1982; 14(2):323-327.
8. Paukert JL, Richards BF. How medical students and residents describe the roles and characteristics of their influential clinical teachers. Acad Med 2000; 75(8):843-845.
9. Ullian JA, Bland CJ, Simpson DE. An alternative approach to defining the role of the clinical teacher. Acad Med 1994; 69(10):832-838.
10. Irby D, Rakestraw P. Evaluating clinical teaching in medicine. J Med Educ 1981; 56(3):181-186.
11. Sloan DA, Donnelly MB, Schwartz RW. The surgical clerkship: characteristics of the effective teacher. Med Educ 1996; 30(1):18-23.
12. Swanwick T. Informal learning in postgraduate medical education: from cognitivism to 'culturism'. Med Educ 2005; 39(8):859-865.
13. Illeris K. Learning and emotions. The three dimensions of learning. Frederiksberg: Roskilde University Press, 2002: 63-85.
14. Billett S. Knowing in practice: re-conceptualising vocational expertise. Learning and Instruction 2001; 11:431-452.
15. Bleakley A. Pre-registration house officers and ward-based learning: a 'new apprenticeship' model. Med Educ 2002; 36(1):9-15.
16. Daley BJ. Learning and professional practice: a study of four professions. Adult Educ Q 2001; 52(1):39-54.
17. Engestrom Y, Miettinen R. Introduction. In: Engestrom Y, Miettinen R, Namaki R, editors. Perspectives on activity theory. Cambridge: Cambridge University Press, 1999: 1-16.
18. Eraut M. Deconstructing Apprenticeship Learning: what factors affect its quality? In: Mulder RH, Sloane PFE, editors. New Approaches to Vocational Education in Europe: the construction of complex learning-teaching arrangements. Oxford: Oxford University Press, 2004: 45-57.
19. Lave J, Wenger E. Situated Learning: Legitimate peripheral participation. Cambridge: Cambridge University Press, 1991.
20. Burr V. An Introduction to Social Constructionism. London: Routledge; 1995
21. Landrigan CP, Barger LK, Cade BE, Ayas NT, Czeisler CA. Interns' compliance with accreditation council for graduate medical education work-hour limits. JAMA 2006; 296(9):1063-1070.
22. Dola C, Nelson L, Lauterbach J, Degefu S, Pridjian G. Eighty hour work reform: faculty and resident perceptions. Am J Obstet Gynecol 2006; 195(5):1450-1456.
23. Robson C. Real world research. 2nd ed. Blackwell Publishing, 2002.

24. Tiberius RG, Sinai J, Flak EA. The Role of Teacher-Learner Relationships in Medical Education. In: Norman GR, van der Vleuten CP, Newble DI, editors. International Handbook of Research in Medical Education. Dordrecht/ Boston/ London: Kluwer Academic Publishers, 2002: 462-498.
25. Castelo-Branco C, Figueras F, Eixarch E, Quereda F, Cancelo MJ, Gonzalez S et al. Stress symptoms and burnout in obstetric and gynaecology residents. BJOG 2007; 114(1):94-98.
26. Kilminster S, Jolly B, van der Vleuten CP. A framework for effective training for supervisors. Med Teach 2002; 24(4):385-389.
27. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004; 79(10 Suppl):S70-S81.

4

How undergraduate clinical learning climates differ: a multi method case study

Klarke Boor
Fedde Scheele
Cees van der Vleuten
Pim Teunissen
Elvira den Breejen
Albert Scherpbier

Medical Education 2008; 42(10):1029-1036

ABSTRACT

Introduction

The clinical learning environment affects undergraduate medical students' behaviour, satisfaction, and success. Most studies predominantly describe aspects of the clinical learning environment using quantitative methodologies, like questionnaires. This study aimed to illuminate medical students' perceptions of the clinical learning climate, and which factors and their interactions explain differences in clinical learning climates.

Methods

We carried out a multi-method case study. Twelve departments of obstetrics and gynaecology distributed the Postgraduate Hospital Educational Environment Measure (PHEEM), a reliable questionnaire measuring the clinical learning environment, among medical students. After analysis (using ANOVA and post hoc tests), 14 medical students from the highest- and lowest-scoring departments participated in semi-structured interviews. We analysed the transcribed recordings using a content analysis approach. Researchers agreed on coding and an expert group reached consensus on the themes of the analysis.

Results

We found a significant difference between departments in PHEEM scores. The interviews indicated that department and medical student characteristics determine the clinical learning climate. For departments, 'legitimacy', 'clerkship arrangements', and 'focus on personal development' were the main themes. For medical students, 'initial initiatives', 'continuing development', and 'clerkship fatigue' were the principal themes. The amount and nature of participation played a central role in all themes.

Conclusions

Differences between clinical learning environments appear to be related to differing approaches to participation among departments. Participation depends on characteristics of both departments and students, and the interactions among them. The outcomes give valuable clues to how a favourable clinical learning climate is shaped.

INTRODUCTION

The clinical learning climate is believed to influence learners' behaviour and predict medical students' satisfaction and success.[1-3] The terms 'clinical learning climate' and 'learning environment' are used interchangeably to describe the perceived "..soul and spirit of the medical school".[1;2] This concept has been the subject of literature reviews[1;2;4] and of studies of clinical learners' experiences[3;5;6] and departmental characteristics fostering a 'good' clinical learning climate.[7] However, no published studies have elucidated *how* certain factors interact to determine the quality of the clinical learning climate.[1-3;5]

With a few exceptions,[3;8;9] studies of learning climates have been quantitative.[1;2;10] Although quantitative measurements are valuable, a qualitative approach was expected to offer understanding on a different level, illuminating how clinical learning climates differ and yielding insights to underpin improvement of clinical learning climates. After quantitatively evaluating the learning climates of a number of departments, we qualitatively explored medical students experiences in the highest- and lowest-scoring departments. Not only did we look for factors that might explain the differences in learning climate, we also sought to elucidate the effects of interactions between these factors.

METHODS

Context

Students at all eight medical schools in the Netherlands undertake their clinical training in academic and affiliated hospitals during clerkship rotations in the final 2-3 years of a 6-year curriculum. We studied the clinical learning climate by comparing students' perceptions of their experiences during clerkships in obstetrics and gynaecology. We limited our study to one specialty to prevent confounding effects of specialty-specific factors. An obstetrics and gynaecology placement is mandatory in all schools. Duration varies from 6 to 8 weeks. This rotation is never the first or last clinical experience for students. Supervision is provided by gynaecologists and specialist registrars (SpRs), sometimes assisted by other members of the educational team, such as midwives.

Six medical schools were selected, because the clerkships in the other two schools differed too much, mainly because students did not stay in the same hospital department for the entire rotation. All six schools agreed to participate. We examined six rotations of similar duration and curriculum in academic and affiliated hospitals using a two-stage, multi-method case study design.

Stage 1. PHEEM questionnaire

Instrument

We used the Postgraduate Hospital Educational Environment Measure (PHEEM), which is a validated, 40-item questionnaire that measures perceptions of the clinical learning environment.[10] It contains items (using a 5-point scale) about feedback, facilities, and atmosphere.[10] The PHEEM has been found to measure one construct and to be reliable with a feasible number of respondents.[11]

Participants

Six obstetrics and gynaecology departments in academic hospitals and six in affiliated hospitals randomly invited 14 students each (the number needed for a reliable outcome)[11] to complete the PHEEM in the period between April and August 2003 (six departments) and between January and June 2005 (six departments). The data collection in 2003 was carried out in connection with another study.[11] Anonymity was guaranteed to all participants. One affiliated hospital was excluded because of an insufficient number of completed questionnaires.

Analysis

Differences between departments were tested by analysis of variance (ANOVA); post hoc tests determined which departments differed from one another. An α-level of 0.05 was considered statistically significant. In order to measure the inter-rater agreement the intra-class correlation coefficient (ICC) per department was calculated.

Stage 2. Semi-structured interviews

Stage 2 involved semi-structured interviews to "…explore meaning and perceptions to gain a better understanding"[12] of the differences found in stage 1.

Participants

It can be useful to look at both ends of a spectrum in order to clarify concepts[13] and thus we interviewed a purposive sample of students from the departments with the highest and lowest PHEEM scores. The departments were not notified of their ranking, but were asked to cooperate in a general study of the clinical learning climate. Consent was obtained from the departments. All students were approached by KB personally and all agreed to be interviewed. Starting in October 2006, we continued recruiting and interviewing students until theoretical saturation was reached in April 2007. Theoretical saturation is reached when no additional themes are found.[14]

Semi-structured interviews

A question schedule was provided but the interviewers were free to follow up on themes that emerged during interviews. Interview questions were based on themes described by Moran and Volkwein.[15] Underlying the notion of the 'interactive approach' is the suggestion that departments and individuals interact to give rise to the organisational - in our case educational - climate.[15] The importance of interaction is also acknowledged in the literature about learning in the clinical environment.[16;17] Students were interviewed by one of three researchers (KB, EdB and PWT) at a time and venue of the student's choice. All participants consented to audiotaping of the interview. Interviews lasted for 45-90 minutes. At most interviews an observer (KB, EdB, or PWT) was present to ask additional questions at the end and discuss the main outcomes with the interviewer afterwards. A summary of each interview was mailed to the respective student within a fortnight. Students were asked to report whether the summary gave an accurate representation of the interview ('member checking').[14] Two students suggested minor revisions.

Analysis

Every interview was transcribed verbatim. Analysis was performed using the method suggested by Miles and Huberman for case study analysis, based on three concurrent 'flows of activity': *data reduction*, *data display*, and *conclusion drawing/verification*.[14;18] To *reduce* and *display the data* we wrote the interview summaries and constructed matrices and figures throughout the analysis process. For additional reduction, we assigned short descriptors to text fragments (coding). The first two interviews were coded by two researchers independently, who resolved any differences by discussion afterwards. The codes were then categorised into themes, which were discussed by two medical doctors and two medical education experts until consensus was reached. These discussions and the member checking of the interviewees constituted the *conclusion drawing/verification* process. The accepted coding and themes were used to analyse the remaining interviews.

RESULTS

Stage 1. The PHEEM questionnaire

Completed questionnaires were returned by 154 medical students from 11 departments. Department ICCs were between 0.73 and 0.85, except for department 7, which had an ICC of 0.54. Mean overall scores varied from 3.53 to 4.30 on a 5-point scale with a significant department effect ($F[10,144] = 4.95$, $P<0.001$). Because the Levene test showed that the dependent variable was not homogeneous, we used Tamhane's T2 post hoc tests. The difference between the highest (department 1) and the lowest (department 11) scoring departments

was significant. The former was an affiliated hospital and the latter an academic hospital. The remaining nine results reflected high- and low-scoring departments in both academic and affiliated hospitals. Thus, there was no evident effect of hospital type.

Stage 2. Semi-structured interviews

We reached theoretical saturation after interviewing seven students from the highest- and lowest-scoring departments each. The highest scoring department had nine staff members and a normal patient population. The lowest scoring department had 28 staff members and a relatively more complicated patient population.

The main themes that emerged from the qualitative analysis referred to departments and medical students. We present the results for various categories within the themes and illustrate our findings with citations from the interviews.

Departments

Three inter-related categories were identified within this theme: legitimacy, clerkship organisation, and focus on students' professional development.

Legitimacy

In department 1 it was fairly easy for medical students to participate in a wide range of activities. All students delivered babies, performed pelvic exams regularly, and assisted in theatre. Most students spoke highly of the department's tradition of allowing students to participate in activities. It was generally accepted that, in order to learn, medical students needed to take part. Gynaecologists, junior doctors, and nurses all tried to involve students in day-to-day practice whenever possible:

"[There is] a climate in which it is taken for granted that you will do things. And that you do not have to ask for that all the time." (D1C2)

"There really was a place for us students. The nurses too (…) they also invite you like, come on, come along. And you just really felt that you were part of the team, that you really belonged." (D1C7)

In department 11, it was difficult for students to obtain any hands-on experience. Very few interviewees had obtained any experience in deliveries or pelvic examinations. The students understood that some patients refused to be examined because of the intimate nature of the pelvic examination. However, some doctors did not even ask patients if students could ex-

amine them, thus strengthening the barrier against participation. One interviewee described examples of frequent responses to requests to carry out a pelvic examination as:

"… "The next patient" or "Well this lady has been so often", eh…"You are of the opposite sex", "This woman is from a different ethnic group", "This woman is young", "This woman is old", "This woman is ill", "This woman .. has already seen many other students", ehh … "This one has been lying here for so long", "This lady is very ill indeed", ehh … " (D11C3)

Clerkship organisation

Both departments organised an introduction and sent students a booklet containing their weekly schedule and explaining what the department expected from them. However, in department 11 there was some confusion: some students expected ample and others few opportunities for performing skills.

In department 1, students were assigned a personal mentor (a gynaecologist) who gave formative feedback at least once and assessed the student at the end of the rotation. Every week students were evaluated by the whole team. In department 11, students had no personal mentor, supervision was often ad hoc and only sometimes satisfactory. Moreover, no-one knew what students' daily activities were and a student's absence could go unnoticed for days. Some doctors barely responded to questions and students were sometimes urged not to speak at all:

"No-one knew exactly what students were doing.… I might as well have stayed at home for days." (D11C5)

"But I also was with a doctor… eh … who said like "Well, a good student sits in the corner, does not get in the way, does not ask questions and eh …" So, there you are, eh … in the outpatient clinic from 9 a.m. to 2 p.m. eh… sitting on your stool."' (D11C6)

Both departments offered formal education once a week: department 1 scheduled a lecture on an obstetrics and gynaecology topic delivered by one of the department specialists, and department 11 organised a discussion led by a retired gynaecologist of case presentations by students.

In department 1, doctors used interesting events that occurred during the day as teaching opportunities, for instance in theatre or at morning report:

"… the registrar explains, "This is a CTG of that lady, at these bradycardias we decided to do an emergency section" … the specialist said "Show the CTG again, yes, show it again to the students .. so they can take a really good look."' (D1C3)

Focus on students' development

In department 1, students felt taken seriously. Staff listened to their judgements when discussing patients in the outpatient clinic or at morning report and repeatedly acknowledged, both explicitly and implicitly, the students' role as learners. Responsibility for attaining their learning goals was partly delegated to the students:

"They said you are not supposed to say at the end of the rotation "I only saw this or this", that you should bring that up early and that they just will… see eh what we can arrange." (D1C1)

In department 11 there was considerable variation in the level of attention paid by staff to students' personal development. Some doctors regarded teaching and coaching students as part of their job but others seemed barely aware of their teaching tasks:

"Eh … I think that registrars for instance should be more prepared to take time for students, to explain things or to examine a patient. Sometimes they really seem to think that students would do things completely wrong, while you have seen it so often and have practised it too." (D11C1)

Medical students

Students influenced the clinical learning climate considerably at the beginning of the rotation and also later in it.

Initial initiatives

Most students started with a proactive attitude: they were aware of the potential benefits of taking the initiative in asking questions, establishing a personal bond with supervisors, showing enthusiasm and studying hard. They felt responsible for their performance in the department and for becoming a good doctor in general:

"Yes, well as a student you have to look a bit further ahead and not like "once this placement is over, it will be alright". You have to eh… take things with you in your basic package as a basic doctor." (D11C2)

Students spent the first days getting to know the departmental culture; they believed that proposals should come 'at the right moment', because they might otherwise have a counterproductive effect. From personal or vicarious experiences, most students were able to cite examples of 'wrong' questions, remarks, or actions that had led to the awarding of lower grades and a generally less enjoyable clinical rotation.

Clerkship fatigue

Some students had a less proactive attitude, for which they blamed 'clerkship fatigue' caused by too many previous clerkship experiences. They were disinclined to be enthusiastic, ask questions, and learn the local customs:

'Actually I wasn't all that motivated. I thought eh … another clerkship where I have to get to know a whole new department… I always find it very difficult that every time you have to start all over again and new and it's really the same story with, with that when you don't know people, it just takes so much time to sort of get settled socially." (D1C6)

Continuing development

The students in department 1 discovered that proactive behaviour was rewarded with increased levels of participation. This strongly motivated them to initiate opportunities to obtain more responsibility and play more important roles within the team:

"… and when you showed initiative then you were also allowed to do things independently: exam, physical examination, IUDs and things like that … In fact you grew, as the rotation continued you grew. You became more independent and confident;, you were then allowed to do some more things." (D1C3)

In department 1, even the students with clerkship fatigue became motivated and involved in clinical work. Students in department 11 fell into two groups. One group, often 'tired clerks', gave up trying to participate after a number of rejections:

"Well, that eh … then you get demotivated. Then you think like, "Well, okay, then I don't care either" eh.. " (D11C4)

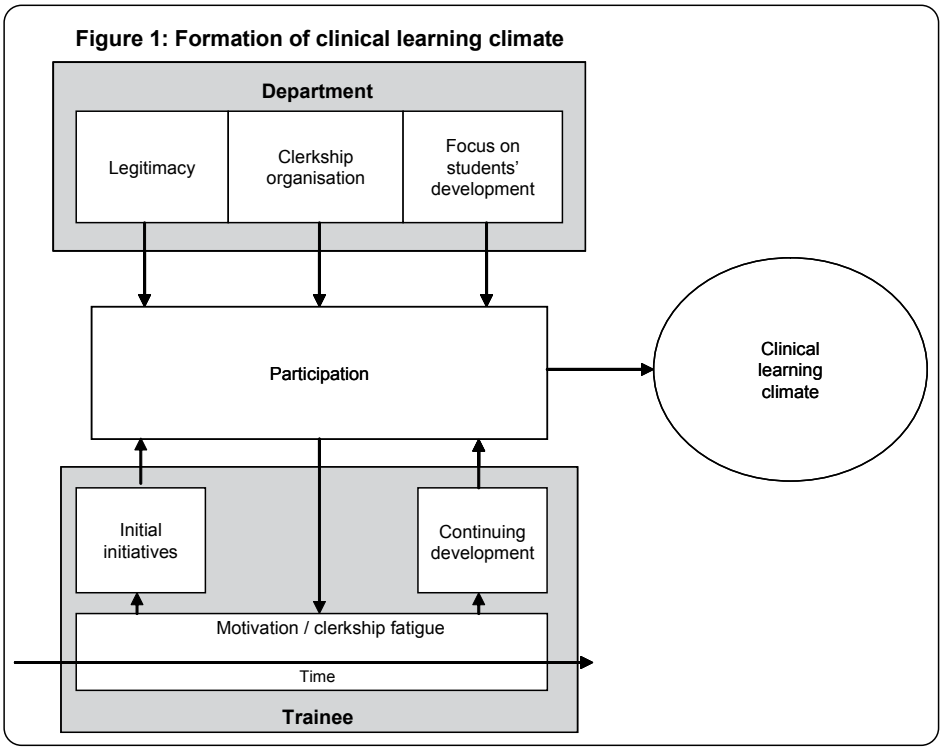
Looking back on the experience, these students felt slightly guilty and blamed themselves for their low level of involvement in the team. The other group persevered and continued their efforts to participate. Eventually, they succeeded and, given the obstacles they had overcome, were proud to have reached some level of involvement.

Figure 1 presents students' perceptions of the departments' and their own roles.

DISCUSSION

The qualitative comparison of two departments with significantly different scores on clinical climate suggest that the differences are associated with department-related and student-

related factors and interactions between them. Figure 1 shows the interaction between overlapping themes that emerged from the results. In department 1, staff encouraged student participation and showed concern for students' personal development. This enhanced the involvement of students, who felt 'rewarded' for their proactive behaviour. In department 11, students experienced less encouragement to participate and some team members appeared not to take teaching responsibilities very seriously. Some students stopped taking initiatives after a while, whereas others persevered and eventually managed to achieve some degree of participation. Participation was also affected by clerkship fatigue among students in both departments.



Participation

Our analysis of student narratives suggests that participation is the key construct that determines the clinical learning climate. Many studies underscore the central importance of participation.[8;19-24] A model of experience-based learning developed by Dornan and colleagues, shows that 'supported participation' is pivotal in clinical workplace learning.[19] Dornan et al. also conclude that student participation is shaped by department-related 'curriculum factors'

and by students' 'human interactions'.[19] Our results show similarities to this model (figure 1). Another study shows that student involvement is a predictor of quality of teaching.[25]

The expansive-restrictive continuum

Our analysis revealed three department-related themes - legitimacy, clerkship organisation and focus on development – all of which are closely related to participation. They resemble approaches described by Fuller and Unwin.[20] Using Lave and Wenger's work on 'situated learning'[21] as starting point, Fuller and Unwin[20] identified characteristics of different approaches to apprenticeship learning and proposed an 'expansive-restrictive continuum'. Expansive approaches promote participation. They are characterised by a 'participative tradition', with explicit institutional recognition of, and support for, the status of learners and with structured support for apprentices. Restrictive approaches are recognisable by narrow access to learning, ambivalence towards the status of learners, no dedicated individual support and restrictive participation.[20] Our results suggest that department 1 is nearer the expansive end of the continuum and department 11 closer to the restrictive end.

Inter-subjectivity

All students were aware that it was important to fit into the departmental culture. They pointed to the need to carefully plan the tone, timing and content of a request or proposal. This is in line with work by Sheehan et al.[22] about 'inter-subjectivity'. Inter-subjectivity occurs "when all members of a health care team understand each other's preferences and idiosyncrasies and where working together can occur without the need for constant negotiation".[22] Medical students who are not sensitive to departmental customs may experience barriers to learning. Departments, in turn, can impede student development by denying students access to critical information or by not including them in the departmental team.

Strengths and weaknesses

One of our study's strengths lies in its use of quantitative and qualitative methods. The PHEEM is a well-researched questionnaire which yields reliable outcomes. The qualitative findings of our study were strengthened by the use of several strategies in analysing the semi-structured interviews, including independent coding, member checking, use of multiple interviewers, and discussion of results in a larger group. The fact that these findings revealed clear discrepancies between the learning climates of the departments with the highest and the lowest PHEEM scores appears to add to the construct validity of the PHEEM.

One of the weaknesses of the study refers to its limitation to one specialty and one group of stakeholders. This means that outcomes cannot be extrapolated to other specialties. It would

also have been of interest to see whether the described 'clerkship fatigue' was reflected in PHEEM outcomes. However, the PHEEM did not include a question on the number of previous rotations.

Conclusions and recommendations

The outcomes of this study give valuable clues as to how a favourable clinical learning climate is shaped. It appears that an expansive approach to participation plays a central role in this, and a favourable departmental orientation towards inter-subjectivity also promotes a good learning climate. This suggests that departments should be aware of the importance of student participation and ways of fostering it. Departments might determine their location on the expansive-restrictive continuum, decide whether they consider it a good place to be and take steps accordingly. Secondly, in recognition of the role of inter-subjectivity in facilitating participation, departments might inform students more explicitly of their customs and rules to help students feel more at ease.[22] Students might find it easier to engage with a departmental team if they knew what was expected (both formally and informally) of them. It is important, however, to realise that inter-subjectivity has a downside. Student and departmental efforts to help students fit into the clinical environment might encourage imitation of ineffective or even undesirable behaviour.[26] Therefore, medical schools should prepare students before and during clerkships to respond to any social pressures they might encounter.

In terms of research, our results can be used to modify questionnaires about the clinical learning climate to better investigate and evaluate clinical learning environments. The different factors identified in the qualitative part of our study may, for instance, serve as input for a multi-dimensional questionnaire. The place of learner participation within the concept of the clinical learning climate deserves further study. Future studies should also look at the clinical learning climate in other specialties and from the perspectives of other members of the clinical team. Residents' perspectives, in particular, might also contribute to understanding of the postgraduate clinical learning climate.

Acknowledgements

The authors would particularly like to thank the students who participated in this study. They also express their gratitude to Mereke Gorsira for her contribution to the English write-up of this paper.

REFERENCES

1. Genn JM. AMEE Medical Education Guide No. 23 (Part 1): Curriculum, environment, climate, quality and change in medical education-a unifying perspective. *Med.Teach.* 2001;**23**:337-44.
2. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. *Med.Teach.* 2001;**23**:445-54.
3. Seabrook MA. Clinical students' initial reports of the educational climate in a single medical school. *Med.Educ.* 2004;**38**:659-69.
4. Roff S. Education environment: a bibliography. *Med.Teach.* 2005;**27**:353-7.
5. Rotem A, Bloomfield L, Southon G. The clinical learning environment. *Isr.J.Med.Sci.* 1996;**32**:705-10.
6. Kendall ML, Hesketh EA, Macpherson SG. The learning environment for junior doctor training--what hinders, what helps. *Med Teach* 2005;**27**:619-24.
7. Hoff TJ, Pohl H, Bartfield J. Creating a learning environment to produce competent residents: the roles of culture and context. *Acad.Med* 2004;**79**:532-9.
8. Deketelaere A, Kelchtermans G, Struyf E, De Leyn P. Disentangling clinical learning experiences: an exploratory study on the dynamic tensions in internship. *Med Educ* 2006;**40**:908-15.
9. Frye AW, Hoban JD, Richards BF. Capturing the complexity of clinical learning environments with multiple qualitative mehods. *Eval Health Prof* 1993;**16**:44-60.
10. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Med.Teach.* 2005;**27**:326-31.
11. Boor K, Scheele F, van der Vleuten CP, Scherpbier AJ, Teunissen PW, Sijtsma K. Psychometric properties of an instrument to measure the clinical learning environment. *Med Educ* 2007;**41**:92-9.
12. Dicicco-Bloom B, Crabtree BF. The qualitative research interview. *Med Educ* 2006;**40**:314-21.
13. Rogoff B. Observing sociocultural activity on three planes: participatory appropriation, guided participation, and apprenticeship. In Wertsch JV, Del Río P, Alvarez A, eds. *Sociocultural studies of mind*, pp 139-64. New York: Cambridge University Press, 1995.
14. Robson C. Real world research. Blackwell Publishing, 2002.
15. Moran ET, Volkwein JF. The cultural approach tp the formation of organizational climate. *Human Relations* 1992;**45**:19-47.
16. Bleakley A. Pre-registration house officers and ward-based learning: a 'new apprenticeship' model. *Med.Educ.* 2002;**36**:9-15.
17. Lyon P. A model of teaching and learning in the operating theatre. *Med.Educ* 2004;**38**:1278-87.
18. Miles MB, Huberman AM. Qualitative data analysis: an expanded sourcebook. Thousand Oaks, California: Sage, 1994.
19. Dornan T, Boshuizen H, King N, Scherpbier A. Experience-based learning: a model linking the processes and outcomes of medical students' workplace learning. *Med Educ* 2007;**41**:84-91.
20. Fuller A, Unwin L. Learning as apprentices in the contemporary UK workplace: creating and managing expansive and restrictive participation. *J Educ Work* 2003;**16**:407-26.
21. Lave J, Wenger E. Situated Learning: Legitimate peripheral participation. Cambridge: Cambridge University Press, 1991.
22. Sheehan D, Wilkinson TJ, Billett S. Interns' participation and learning in clinical environments in a New Zealand hospital. *Acad.Med* 2005;**80**:302-8.
23. Wenger E. Communities of Practice. Learning, Meaning, and Identy. Cambridge: Cambridge University Press, 1998.

24. Teunissen PW, Scheele F, Scherpbier AJJA, van der Vleuten CPM, Boor K, Van Luijk SJ *et al*. How residents learn: qualitative evidence for the pivotal role of clinical activities. *Med Educ* 2007;**41**:763-70.
25. Hoellein AR, Feddock CA, Wilson JF, Griffith CH, III, Rudy DW, Caudill TS. Student involvement on teaching rounds. *Acad.Med* 2007;**82**:S19-S21.
26. Lingard L, Reznick R, Espin S, Regehr G, DeVito I. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Acad.Med* 2002;**77**:232-7.

5

The postgraduate learning climate: integration of work, training, and residents' needs

Klarke Boor
Albert Scherpbier
Pim Teunissen
Cees van der Vleuten
Fedde Scheele

Submitted

ABSTRACT

Introduction

The aim of this study is to gain insight into the (optimal) learning climate in specialist training.

Methods

Semi-structured interviews and focus groups with residents in various phases of training in different specialties.

Results

We conducted 14 interviews and 3 focus groups (consisting of 26 participants). Residents described an optimal learning climate combines work and training, tailored to residents' specific needs. This climate is characterised by focused discussions of specific patients, a safe environment, or a situation where supervisors know the person they are educating and vice versa. Residents thought they learned most from seeing patients and valued participation in patient care. Nevertheless, the learning climate was strongly affected by the uneasy balance between service provision and education. There was also an impact from (lack of) co-operation with other team members. Both residents and supervisors were considered capable of initiating steps towards a better learning climate, e.g., by asking and giving feedback or by studying and teaching.

Conclusions

This study gives insight how optimal postgraduate learning climates can be created. An optimal climate combines learning by participating in the clinical workplace and learning in the more traditional sense, such as self-study and tuition. Other beneficial factors are opportunities for residents to take ownership of their education, and flexible programs. Moreover, ward teams should be aware of residents' desire for more reciprocal relationships. Encouragingly, residents suggested that substantial improvement can be achieved with relatively little effort.

INTRODUCTION

Much has been written about the clinical learning climate in medical education.[1-4] One approach to study the clinical learning climate is by means of instruments measuring this concept. The PHEEM, for instance, is a questionnaire developed to study the learning climate in postgraduate medical education.[5] It consists of relevant questions on, for instance, feedback or atmosphere, but lacks a theoretical basis or thorough description of what defines the learning climate.[5]

Another approach to unravel the clinical learning climate is by using qualitative methods. A growing number of studies uses this approach to focus on specific aspects of the learning climate such as the learning climate in specific places (like operation rooms[6]), specialty fields (for instance anaesthesiology[7] or nursing home medicine[8]), or certain behaviours influencing the learning climate (such as harassment[9]). This research offers valuable insights, but qualitative studies of *general features* of clinical learning climates are scanty, although doing so might afford important insights.[4;10;11] In this study, we aimed to distill generic characteristics of residents' clinical learning climate.

Lave and Wenger's work on 'situated learning'[12] served as a point of departure for our research. This theory of learning states that all learning is situated in a certain context and culture. Social interaction and collaboration are pivotal components: through participating in a 'community of practice' a learner gradually becomes a full participant.[12] An earlier qualitative study in the *undergraduate* setting describing medical students' perceptions of their learning climate confirmed the importance of how departments approached participation.[4]

We conducted a two-part qualitative study in the *postgraduate* setting to examine how residents from different specialties view their clinical learning climate. We purposefully included a diverse range of residents in order to transcend specialty related factors and get a grip on the concept of the general clinical learning climate. Firstly, we conducted interviews in which residents were asked to recount experiences which they felt had enhanced or impaired their learning climate. Next, focus groups were convened to supplement and evaluate the results of the interviews.

METHODS

Context

Undergraduate and postgraduate medical education is provided in eight academic medical centres and numerous affiliated general hospitals in the Netherlands. After six years of undergraduate education, medical graduates apply for a place in a postgraduate training

program in the specialty of their preference. Specialty training takes from four to six years, depending on the specialty. Clinical staff in hospital departments where trainees undertake their training are expected to contribute to specialist education. Working conditions vary per specialty, but generally residents see patients in clinical and outpatient settings with progressive independence. Working hours are restricted to an average of 48 hours per week.

Ethical approval

Because no patients were included, the study was exempt from formal ethical board review. Care was taken to preclude any study-related negative effects to participants. Anonymity was guaranteed and participation was voluntary.

Semi-structured interviews

Participants

We recruited trainees from four major specialties (internal medicine, surgery, obstetrics & gynaecology, and paediatrics). The participants had been in training between several months and six years and were on rotations in academic and general hospitals. Interviewing was continued until 'theoretical saturation', i.e. cessation of the emergence of new themes, was reached: we held 14 interviews.[13]

Semi-structured interviews

The interviews were semi-structured, with interviewers being free to ask the predetermined questions in the order they deemed appropriate.[13] In January and February 2008, the first (KB) and third (PWT) author conducted the interviews, which lasted 45-90 minutes. PWT conducted one interview, observed by KB, who conducted the other interviews. Consent was obtained from all participants for audio taping and verbatim transcribing of the interviews. Insights from the situated learning perspective- such as learning is in the relationship with others, people learn through becoming participants in communities of practice, and knowledge and activities are intimately connected[12;14]- influenced the question route, which was also used in an earlier study in the undergraduate setting.[4]

Within a fortnight of the interview, the interviewees were asked to check the accuracy of a summary of their interview. This 'member checking' resulted in minor revisions suggested by one interviewee, and addition of relevant experiences supplied by two participants, who had omitted to share these during the interview.

Based on the first analysis of the interviews with residents from four specialties only, we constructed a preliminary framework of themes. Because we strived to get an impression of an optimal learning climate for residents from *all* specialties we held focus groups with

residents from additional specialties in order to include their opinions. In addition, the use of another method permitted triangulation which strengthens the results.[13]

Focus groups

Participants

Convenience sampling of trainees across different specialties (Internal Medicine, Cardiology, Anaesthetics, Surgery, Obstetrics & Gynaecology, Paediatrics, Neurology, Clinical Chemistry, ENT, Dermatology and Orthopaedics) included female and male participants and a range of experience years. Each group consisted of trainees from one hospital. After three sessions with a total of 26 participants from one academic hospital and two general hospitals 'theoretical saturation' had been reached.

Focus groups

Focus group interviews are group discussions facilitated by a moderator who closely observes the group interaction.[15] The sessions were held in March 2008 and lasted approximately 90 minutes (excluding a break). Before the actual discussion, the participants were asked to write down some factors they had perceived as facilitating or impeding their learning climate. The moderator (PWT) then invited the participants to share these factors with the group. The observer (KB) recorded every (new) factor on the whiteboard. When no new factors were proposed, the group took a short break during which the two researchers investigated the similarity of the proposed factors to the factors in the preliminary framework derived from the interviews. After the break the moderator presented the preliminary framework to the group and invited comments. By discussing how to deal with themes that did not fit well within the framework, the focus groups contributed to the development of a framework of themes that impact on the learning climate in specialist training. 'Member checking' within a fortnight after the sessions yielded no suggestions to revise the summaries.

Analysis

We analysed the interviews and the focus groups using a strategy described by Miles and Huberman,[16] which involves, concurrently, *data reduction* by simplifying and abstracting, *data display*, and *drawing conclusions and verification*. KB first used an open coding strategy to *reduce the data*, which was followed by multiple interpretative rounds to reach a deeper understanding of the data (one round, for example, consisted of discussing preliminary results with a group of educational experts and clinicians in order to achieve intersubjective consensus); this resulted in thematic codes. PWT independently checked one interview and any differences were resolved through discussion. For *data display* we used different methods, such as networks and graphs. For *drawing conclusions and verification* we employed an

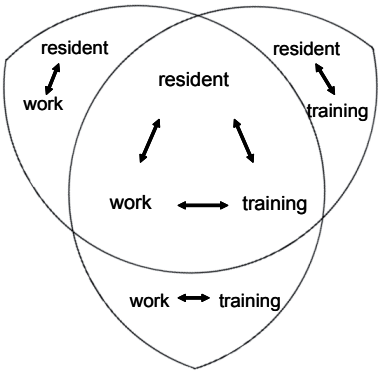
inductive, cyclical approach to direct further data collection (e.g., the use of interviews and focus groups for triangulation, or the use of 'member checking').

RESULTS

Three interrelated themes are important in residents' eyes (and underlined in the following phrase): an optimal learning climate combines work and training, tailored to residents' specific attributes and needs.

When residents talk about work in relation to their learning climate, they mainly discuss their daily efforts to deliver good quality patient care. This is influenced by organisational issues (like working hours or protocols around patient care and safety) and by (un)existing 'collaborative' structures (i.e., (un)noticeable conflicts between consultants). Elements associated with training and its relationship to learning climate were: feedback, i.e., supervisors' comments on trainees' functioning during (semi-)annual appraisal interviews and in ad hoc situations; formal education, with consultants sharing their knowledge and skills on specific topics, frequently combined with preparatory self-study by trainees; and the role of the program director, who was expected to monitor trainees' progress and engage in continuous improvement of training.

Figure 1: Interaction of the three main themes



Learning climates in specialist training are affected by three themes (resident, work, and training) which interact in four different ways. In an optimal situation the three themes would be fully integrated, i.e. overlap completely at the centre of the figure. The areas away from the middle depict situations where two themes interact

Residents' needs and behaviours constituted the third theme, where the lack of confidence was an important issue, as were strategies to cope with that deficiency (for instance asking colleagues or supervisors for support). Another trainee-related theme was awareness and protection of personal boundaries.

Residents explained how these three themes interact in facilitating or hindering an optimal learning climate. This is depicted in figure 1. We will discuss the interactions and illustrate them by quotations from interviews and focus groups.

Work-Training-Resident

Residents considered features promoting a good match between work, training, and residents' personal needs to be conducive to a positive learning climate. A first often described feature is the discussion around a specific patient: consultants can stimulate residents by spontaneous explanations or by discussing patients' management with each other and expecting the resident to contribute to the discussion as well.
Another feature was a 'safe environment', which was perceived to promote patient safety, because it made residents less hesitant to admit mistakes, fostered effective training – where no questions were derided - and enhanced residents' personal development. The same was said about attending physicians' openness and approachability or lack thereof, which might negatively affect the quality of patient care. According to residents departments existed where after duty hours attending physicians were difficult to get hold of or to persuade to come to the hospital when trainees needed senior help.

"Well, I think that especially at night it happens that you call a gynaecologist and say well 'I have a problem and you need to come now' (..) that the gynaecologist started by saying 'I don't think I need to come at all' and then I think well, if I need someone now it does not matter what I need them for. I just need help right now, no matter for what, that he just comes and does not start to argue with you in the middle of the night." (FG3-Trainee Obs/gyn).

A third feature is getting to know each other on a more personal level. Knowing an attending physician, and being known, as a 'person' positively influenced work, training and residents' personal development.

Work-Resident

Residents concurred that, given their differing backgrounds, work and training should be tailored to their personal needs. In practice, however, residents often felt they had little say in

the planning of their rotations. Many felt that service imperatives prevailed to the detriment of their learning needs.

"For I often think that it is all about production and not about learning and, er, I notice (…) that it is more important that I do an outpatient clinic because it, er, reduces our waiting list and that it matters less that I should do the outpatient clinic with a certain attending who has a lot to teach me." (FG1, paediatrics trainee).

At the same time, residents said they learned most from seeing patients.

"The whole training program revolves around learning by doing and you learn from the patients you see. In the end I think that that is the biggest learning event." (I12, paediatrics trainee).

Residents valued seeing multiple patients with similar problems, because this fostered a more thorough understanding of a certain disease but they also appreciated task diversity and complexity to help them obtain enough experience and gain confidence in different areas of their specialty.
Another important work-related factor was co-operation with the team. Working well with fellow residents (sharing feelings of insecurity or helping out with night shifts) and nurses (appreciating their expertise) fostered a good learning climate. Residents occasionally perceived some imbalance in co-operating with nurses: they felt they had to make an effort to gain nurses' assistance and input, for instance by doing chores or being overly polite.

"(..) when you arrive as a resident (..) you are always the one who has to introduce him/herself and that every time it is you who has to take the initiative and yes, that you are expected to behave yourself and be polite but that that doesn't apply so much the other way around, (…) and then I think, well with all due respect I am willing to do my part but it has to be a two-way street and then things may get a little more awkward and then you notice it immediately in day-to-day work." (FG1, surgical trainee)

Residents, the more senior ones in particular, said that having some influence over their work rotations led to an enhanced learning climate. However, they differed in the opportunities they saw for themselves to adapt their work and interactions to their personal preferences.

Training-resident

This theme comprised the same ideas about tailor-made programs as did the work-resident theme. It was suggested that, in order to promote personalised training, departments should

organise an introductory meeting where resident and trainer explicate their expectancies and a personal training program can be worked out accordingly.

"(..) that's what she [the trainer] said, 'What do you want? What are your goals at the end of this year? Where do you want to be? What are your objectives? Which way, er, do you want to go?'" (I11, surgical trainee)

This was in sharp contrast to the state of affairs in departments where trainees were made to feel they were only an extra pair of hands, and where virtually no explicit attention was paid to their 'learner status'.
Doctors in training said they could undertake various activities to further their personal development. These activities included seeking feedback, both implicitly (by monitoring) and explicitly (by asking for feedback). Second, they tried looking things up and studying, which they felt they should do more often. Interesting patients, enough time, and anticipated assessment triggered active studying. Another activity was pro-active behaviour.

"(..) you can look everything up, you can choose a topic for a patient presentation, you can ask for more explanations. These are ways to structure your own training program." (I8, internal medicine trainee)

Residents also described contributions to a good learning climate from supervisory strategies, such as formal teaching by supervisors and feedback on trainees' activities. They received this too infrequently, residents said. Many said they just assumed they were doing all right when no feedback was given. Being assessed by supervisors was another suggestion: trainees thought that supervisors should put them to the test more often.

"One supervisor wanted you to recount what you were doing [during surgery] and I think that is the way it should be done, like 'What exactly are you doing, (…) what are you looking for, why are you looking for that?' (…) And, yes that would make one study a lot harder." (110 surgical trainee).

Work-training

Situations were mentioned where work and training coincided but where no account was taken of individual trainees' needs, such as elaborate (planned) discussions of patients seen in the outpatient clinic or on rounds. Although these meetings were not tailored to individual trainees, the trainees said these meetings fulfilled a clear dual goal of supporting patient care as well as their training.

CONCLUSIONS

Our results suggest that the learning climate in postgraduate medical education can be optimised when work and training are appropriately integrated and adapted to the differing learning needs of individual doctors in training. These findings resonate with earlier research of factors facilitating residents' learning.[8;17-19] The novelty of this study lies in the diverse backgrounds of the participants; this resulted in the description of an optimal learning climate valid for every resident. Moreover, we employed interviews and focus groups, which are suitable to get a profound impression of an ill-defined concept such as the postgraduate learning climate.[20] Furthermore, our 'two-step' approach (consisting of interviews resulting in a preliminary model, and focus groups checking and adapting this into the final model) adds to the robustness of our findings. A possible weakness is that one researcher conducted most of the interviews. This may have affected the course of the interviews, although comparison with the interview conducted by the other researcher yielded no major discrepancies. Another weakness is we interviewed trainees only, where supervisors, or other co-workers, could have added valuable input as well.

This study started with the assumption that (supporting) participation is pivotal to learning. Residents confirmed this assumption but recognised value in 'formal' learning as well. This resonates with Sfard's description of two major outlooks on learning: the acquisition metaphor and the participation metaphor.[21] The acquisition metaphor represents a 'traditional' concept of learning as acquiring knowledge which "..makes us think about the human mind as a container to be filled with certain materials and about the learner as becoming owner of these materials" (p.5). The participation metaphor represents a relatively recent take on learning, where "..learning is viewed as a process of becoming part of a greater whole" (p.6), with a strong emphasis on the constant flux of doing -as opposed to the permanence of having in the acquisition metaphor-, and on the importance of context. In evaluating learning climates medical educationists have tended to focus on the 'acquisition forms of learning'.[22] Lave and Wenger, in contrast, more or less reject formal learning, and focus on learning by participation only[11]; a controversial assumption that is debated by other authors.[23] The participants in this study stressed that both outlooks have worth and various other researchers share this view.[24-28]

Residents are not clean slates

Our study shows that consideration of residents' personal attributes can improve their learning climate. New trainees arrive with their previous experiences and should not be treated as if their mind were a tabula rasa.[22] This view has two implications for departments. Firstly, they should see residents not as passive recipients of the training delivered to them but as

active learners who are well able to decide which learning opportunities are useful to them.[29] In a good learning climate residents are empowered to make these choices. This will benefit not only residents but the department as well: fresh input from a newcomer can inspire indispensable innovations.[30] Secondly, it seems unwise for departments to aim for uniformity of training programs, because, as Ellström states: "the learning processes and outcomes of different people placed in the same job with the same learning potential will be expected to differ based on their personal readiness."[31] In an optimal learning climate evaluation of residents' personal readiness determines which types of work and training are most appropriate for them.

Reciprocity

Learning climates were affected by interactions between various parties. In many interactions reciprocity (the equality of perceived investments and outcomes relative to the person's internal standards[32]) appeared to play a role. Residents mentioned being at a disadvantage in interactions with nurses and attending physicians, for example when attending physicians were unavailable for consultation after working hours. This is in line with earlier publications on reciprocity in specialist training[33] and suggests that the learning climate for residents stands to gain from increased reciprocity of relationships within ward teams. More reciprocity might even diminish the number of residents with burn-out, which has been shown to be related to an imbalance in give-and-take.[33]

A small step?

Residents gave ample illustrations of situations where improvement would be easy to achieve. A few extra minutes for questions and background information could greatly improve the learning value of discussions about patients and one conversation between resident and trainer could improve learning when it resulted in an appropriate personal learning program. The wish may be the father of the thought in this case because many studies have shown how difficult it is to change.[34;35]

Conclusions and recommendations

This study offers suggestions for what departments can do to optimise their learning climate. An important message is that learning can consist of participating in communities of practice but also, more traditionally, of studying and tuition (on or off the job). Additionally, residents should be allowed to (partly) determine their own training programs and be offered (somewhat) flexible programs. Fostering more reciprocity in relationships with the ward team

would also be beneficial. A final interesting insight is that small changes may be effective in bringing about substantial improvement.

Future studies might use the framework presented in this study as an empirical basis for developing (multidimensional) questionnaires to examine the learning climate in postgraduate medical education. Moreover, a recent study by Rodriguez-Paz et al. stresses that, as long as a safe culture is fostered, much can be learned from mistakes and patient safety can improve.[36] It would, thus, be worthwhile to research whether the learning climate in specialist training affects residents' functioning and whether this, in turn, affects the quality of patient care.

Acknowledgments

The authors thank all the residents who participated in this study. We are also grateful to Mereke Gorsira for her editorial assistance.

REFERENCES

1. Roff S, McAleer S. What is educational climate? Med Teach 2001; 23(4):333-334.
2. Roff S. Education environment: a bibliography. Med Teach 2005; 27(4):353-357.
3. Seabrook MA. Clinical students' initial reports of the educational climate in a single medical school. Med Educ 2004; 38(6):659-669.
4. Boor K, Scheele F, van der Vleuten CPM, Teunissen PW, den Breejen EME, Scherpbier AJJA. How undergraduate clinical learning climates differ: a multi-method case study. Med Educ 2008; 42(10):1029-1036.
5. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. Med Teach 2005; 27(4):326-331.
6. Lyon P. A model of teaching and learning in the operating theatre. *Med.Educ* 2004;**38**:1278-87.
7. Pope C, Smith A, Goodwin D, Mort M. Passing on tacit knowledge in anaesthesia: a qualitative study. *Med Educ* 2003;**37**:650-5.
8. Stok-Koch L, Bolhuis S, Koopmans R. Identifying factors that influence workplace learning in postgraduate medical education. *Educ.Health (Abingdon.)* 2007;**20**:8.
9. Musselman LJ, MacRae HM, Reznick RK, Lingard LA. 'You learn better under the gun': intimidation and harassment in surgical education. *Med Educ.* 2005;**39**:926-34.
10. Hoff TJ, Pohl H, Bartfield J. Creating a learning environment to produce competent residents: the roles of culture and context. Acad Med 2004; 79(6):532-539.
11. Snadden D. Clinical education: context is everything. Med Educ 2006; 40(2):97-98.
12. Lave J, Wenger E. Situated Learning: Legitimate peripheral participation. Cambridge: Cambridge University Press, 1991.
13. Robson C. Real world research. A resource for social scientists and practitioner-researchers. 2nd ed. Blackwell Publishing; 2002.
14. Smith, MK. 'Communities of practice', *the encyclopedia of informal education*. www.infed.org/biblio/communities_of_practice.htm. Last updated: 28-12-2007.
15. Barbour RS. Making sense of focus groups. Med Educ 2005; 39(7):742-750.
16. Miles MB, Huberman AM. Qualitative data analysis: an expanded sourcebook. 2nd edition ed. Thousand Oaks, California: Sage; 1994.
17. Elliott RL, Juthani NV, Rubin EH, Greenfeld D, Skelton WD, Yudkowsky R. Quality in residency training: toward a broader, multidimensional definition. *Acad.Med.* 1996;**71**:243-7.
18. Schultz KW, Kirby J, Delva D, Godwin M, Verma S, Birtwhistle R *et al*. Medical Students' and Residents' preferred site characteristics and preceptor behaviours for learning in the ambulatory setting: a cross-sectional survey. *BMC.Med.Educ.* 2004;**4**:12.
19. Thrush CR, Hicks EK, Tariq SG, Johnson AM, Clardy JA, O'Sullivan PS *et al*. Optimal learning environments from the perspective of resident physicians and associations with accreditation length. *Acad.Med* 2007;**82**:S121-S125.
20. Britten N, Jones R, Murphy E, Stacy R. Qualitative research methods in general practice and primary care. Fam Pract 1995; 12(1):104-114.
21. Sfard A. On two metaphors for learning and the dangers of choosing just one. Educational Res 1998; 27(2):4-13.
22. Swanwick T. Informal learning in postgraduate medical education: from cognitivism to 'culturism'. Med Educ 2005; 39(8):859-865.

23. Fuller A, Hodkinson H, Hodkinson P, Unwin L. Learning as peripheral participation in communities of practice: a reassessment of key concepts in workplace learning. *Br Educ Res J* 2005;**31**:49-68.
24. Fuller A, Unwin L. Learning as apprentices in the contemporary UK workplace: creating and managing expansive and restrictive participation. J Educ Work 2003; 16(4):407-426.
25. Billett S. Situated learning: bridging sociocultural and cognitive theorising. Learning and Instruction 1996; 6(3):263-280.
26. Teunissen PW, Scheele F, Scherpbier AJJA, van der Vleuten CPM, Boor K, Van Luijk SJ et al. How residents learn: qualitative evidence for the pivotal role of clinical activities. Med Educ 2007; 41(8):763-770.
27. Teunissen PW, Boor K, Scherpbier AJ, van der Vleuten CP, Diemen-Steenvoorde JA, Van Luijk SJ et al. Attending doctors' perspectives on how residents learn. Med Educ 2007; 41(11):1050-1058.
28. Bleakley A. Broadening conceptions of learning in medical education: the message from teamworking. Med Educ 2006; 40(2):150-157.
29. Billett S. Workplace participatory practices. Conceptualising workplaces as learning environments. J Workpl Learn 2004; 16(6):312-324.
30. Bleakley A. Pre-registration house officers and ward-based learning: a 'new apprenticeship' model. Med Educ 2002; 36(1):9-15.
31. Ellström P. Integrating learning and work: problems and prospects. Human Resource Development Quarterly 2001; 12(4):421-435.
32. Schaufeli WB, van Dierendonck D, van Gorp K. Burnout and reciprocity, towards a dual-level social exchange model. Work Stress 1996; 10(3):225-237.
33. Prins JT, Gazendam-Donofrio SM, Dillingh GS, van de Wiel HB, van der Heijden FM, Hoekstra-Weebers JE. The relationship between reciprocity and burnout in Dutch medical residents. Med Educ 2008; 42(7):721-728.
34. Bland CJ, Starnaman S, Wersal L, Moorehead-Rosenberg L, Zonia S, Henry R. Curricular change in medical schools: how to succeed. *Acad.Med.* 2000;**75**:575-94.
35. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. *Med J Aust.* 2004;**180**:S57-S60.
36. Rodriguez-Paz JM, Kennedy M, Salas E, Wu AW, Sexton JB, Hunt EA *et al*. Beyond "see one, do one, teach one": toward a different training paradigm. *Qual.Saf Health Care* 2009;**18**:63-8.

6

The development and psychometric quality of D-RECT, a multi-factorial instrument to measure postgraduate residents' learning climate

Klarke Boor
Cees van der Vleuten
Pim Teunissen
Albert Scherpbier
Fedde Scheele

Submitted

ABSTRACT

Context

Many instruments have been designed to measure learning climates but most of them lack an explicit theoretical foundation and their psychometrical properties have scarcely been tested.

Objective

This paper describes the development of an instrument to measure the learning climate in postgraduate specialist training: the Dutch Residency Educational Climate Test (D-RECT). D-RECT's psychometric qualities are also reported.

Design

In April-May 2008 a preliminary questionnaire (based on a theoretical framework) was evaluated in a modified Delphi procedure. In May 2008 all residents in the Netherlands were invited to fill out the preliminary questionnaire. Exploratory factor analysis was performed to analyze the outcomes and construct the definitive D-RECT. Confirmatory factor analysis tested the questionnaire's goodness of fit. Generalisability studies tested the number of residents needed for a reliable outcome.

Setting: Postgraduate specialist training in the Netherlands.

Participants: 38 (of 40) experts participated in the Delphi procedure. 1278 (of 4835) residents representing 26 specialties completed the questionnaire.

Results

In two rounds, the Delphi panel reached consensus concerning removal or retention of questionnaire items. Exploratory factor analysis of the completed surveys led to the definitive D-RECT, consisting of 50 items in 11 subscales. Cronbach's α varied from 0.68 to 0.82 for the different subscales. Confirmatory factor analyses confirmed the construct (CFI=0.90, RMSEA=0.04, CMIN/DF=2.9). The results showed that eight residents can reliably assess most subscales for a single department. Four residents per department can assess groups of departments reliably.

Conclusions

The D-RECT appears to be a valid, reliable and feasible tool to measure the quality of the postgraduate learning climate.

INTRODUCTION

Governments as well as medical training boards expect, and progressively *in*spect, a high quality of residency training.[1-3] One way to evaluate the quality of training programs is to evaluate the learning climate. As early as the 1960s and 1970s, many instruments to measure learning climates were developed in the field of organizational psychology.[4] In the same period, measuring instruments for undergraduate learning environments were also described within the domain of medical education.[5-7] More recently, Roff and McAleer put the educational climate on the medical education map by asserting its importance[8;9] and stimulating the development of measurement instruments, such as for the undergraduate setting (DREEM)[10], the Operating Room (STEEM and OREEM)[11;12], and the postgraduate setting (PHEEM)[13]. Despite widespread usage of these instruments, it has proven difficult to replicate the original subscales within PHEEM[14] and DREEM (Isba, personal communication) empirically and there is debate on the best fitting subscale structure.[15;16] An absent or controversial subscale structure hinders the possibilities for specific feedback to departments. Other limitations of these instruments are that most studies fail to report the number of participants needed for a reliable outcome and the interpretation of the findings is hampered by a lack of an explicit theoretical foundation.
Some other studies have described questionnaires that were developed to measure the *overall* quality of training programs but these questionnaires are tailored to specific specialties[17;18] or a specific medical school.[19]

Early in 2008, we performed a qualitative study, reported elsewhere[20], among residents at different levels of training in various specialties to explore how an optimal learning climate can be created. The results showed that an optimal learning climate is characterized by integration of work and training and tailored to individual residents' needs. We used these three interacting domains (work, training, and residents' needs) as the theoretical basis for a learning climate questionnaire, the Dutch Residency Educational Climate Test (D-RECT) that is the subject of this study. We expected that a sound theoretical basis would help us avoid some of the shortcomings of earlier questionnaires. We narrowed down the outcomes of the qualitative study[20] in a modified Delphi procedure and by psychometric analysis to test the validity and reliability of the questionnaire.

METHODS

Setting

In the Netherlands, medical students obtain a basic medical degree after six years of undergraduate medical training. This entitles them to apply for a place in a training program in one of 27 specialties. Depending on the specialty, training lasts from four to six years. In this paper we use the term resident to refer to a junior doctor who is undertaking postgraduate specialist training. Specialist training programs consist of rotations in a university hospital and an affiliated general hospital. In every hospital department where training is offered, one specialist is the 'specialty tutor' and formally responsible for residents' education and training as well as (bi-)annual assessments

Ethical considerations

This study was exempt from Institutional Board Review under Dutch law. However, we made sure that no possible harm could come to our participants. In the invitation to the Delphi panelists and in the letter inviting the residents to take part in the questionnaire study, we explicitly stated that participation was voluntary and full anonymity was guaranteed.

Creating the preliminary D-RECT

In an earlier study we had found that an optimal learning climate was characterized by the integration of work and training and tailored to residents' needs.[20] We used input from this qualitative study in designing the first version of the learning climate questionnaire, which consisted of 83 items. These 83 items were discussed in an expert group of two medical education experts (CvdV and AS), two medical doctors (PT and KB), and one specialty tutor (FS). Additionally, four residents, two specialty tutors, and three medical educationalists (all of whom were not on the research team) checked the items for face validity and made suggestions for removal or rewording. This resulted in a 75-item *preliminary* D-RECT. Every item invited agreement on a five-point Likert scale (1= totally disagree and 5= totally agree); we also included a not applicable option.

Creating the *definitive* D-RECT

We submitted the 75 items of the preliminary D-RECT to a Delphi panel. In addition we performed several psychometric analyses on a large pool of completed questionnaires.

Delphi procedure

A Delphi procedure is aimed at achieving consensus among experts in a systematic manner.[21] In multiple consultation rounds, experts indicate their (dis)agreement with statements or

concepts. After the first round the experts can change their own rating in light of the summarized (anonymous) ratings of the other panel members.[22] In a modified Delphi procedure, the statements or items are not generated by the expert group but -as in this study - carefully selected based on earlier research.[23]

In April and May 2008 we invited ten medical educationalists, ten policymakers, ten residents, and ten specialists (all specialty tutors) for our Delphi panel; the latter two groups represented different specialties as well as university and non-university hospitals. The 40 panelists were chosen for their involvement in postgraduate specialist training. They were offered a monetary incentive for participating in the whole Delphi procedure. Experts received the preliminary D-RECT and rated every item's relevance in relation to postgraduate learning climate on a 5-point scale (1= not relevant and 5= highly relevant). For the analysis the ratings were dichotomized (1, 2, and 3 were interpreted as not relevant; 4 and 5 as relevant). In the absence of undisputed guidelines to decide when consensus is reached,[24] we decided that agreement among 80% of participants would lead to inclusion or exclusion of an item and that consensus was considered to have been reached when 90% of the items that were included or excluded in one round remained unchanged in the subsequent round.

Questionnaire mailing

In May 2008 a letter was sent to all residents in the Netherlands, asking them to complete the web-based preliminary D-RECT and answer some demographic questions. Psychiatric residents were not included for logistical reasons. Some of the respondents could win an incentive provided they completed the questionnaire. No reminders were sent.

Exploratory Factor Analysis

Factor analysis is used to identify clusters of related variables.[25] We first performed exploratory factor analysis using both Varimax rotation (which presupposes no correlations between the factors) and Oblimin rotation (which presupposes some correlations between the factors).[25] Because the 'Component Correlation Matrix' indicated correlations between the factors, from then on we only used Oblimin rotations.[25] Items with weak factor loadings were eliminated and internal consistency of the factors was determined by calculating Cronbach's α. Our theoretical framework[20] and the outcomes of the Delphi procedure guided our decisions on the inclusion of ambiguous items. This resulted in a multi-factorial model: the *definitive* D-RECT.

Testing the *definitive* D-RECT

Confirmatory Factor Analysis

Using structural equation modeling we tested the goodness-of-fit of the multi-factorial model. The Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RM-

SEA) and chi-square (CMIN) were used as indices of the goodness-of-fit.[26] Because of the large sample size (which can influence the CMIN result) we also report the Relative Chi-Square (CMIN/DF).

Reliability analysis

In order to determine generalisability, variance component analysis was performed to measure the contributions of all relevant components (in this case: residents, departments, items and their interactions) to the variance in an outcome measure.[27] We performed generalisability analysis for the mean total score and each separate subscale, to estimate the number of residents needed to obtain reliable test scores.

We treated the total number of items as fixed, because in this analysis we did not want to assess how modifying the instrument's length would influence score integrity. The number of residents within a single department and the number of departments were allowed to vary. Following variance component estimation, we estimated the Standard Error of Measurement (SEM) for a single department. The SEM can be interpreted on the original scoring scale (in this case 1 to 5) and we decided to accept a maximum 'noise level' of 1.0 on the scale. We therefore used a SEM < 0.26 (1.96 x 0.26 x 2 ≈ 1.0) as the smallest admissible value for a 95% confidence interval interpretation.

To use D-RECT across a group of departments, we estimated the Root Mean Square Error (RMSE), which can be interpreted in the same way as the SEM but now at group level.

We used Amos structural equation modeling software for the confirmatory factor analysis, URGENOVA software to analyze generalisability, and SPSS for all the other analyses.

RESULTS

Creating the *definitive* D-RECT

Delphi procedure

Two respondents did not complete the Delphi procedure: one specialist was unable to take part because of health problems and one resident failed to respond in time; the other 38 respondents completed the full Delphi procedure. In the 1st round, one of the researchers (KB) sent the preliminary D-RECT to all participants asking them to rate every item's relevance for the postgraduate learning climate. If desired, they could add qualitative remarks to explain their ratings. The completed questionnaires were returned to KB, who summarized the ratings and remarks. In the 2nd round, the preliminary D-RECT was sent to all participants, with the item mean scores and standard deviations of all items in the 1st round as well as (anonymous) summaries of the qualitative remarks. The participants were also asked to (dis)

agree to some slight adjustments to seven items, which were simplified to improve clarity. The panel again completed the questionnaire and returned it to KB, who analyzed the new ratings and remarks. All participants approved the seven adjusted items. Moreover, consensus was reached in this round: 91% of the items included and excluded in the 1st and 2nd rounds were identical. The Delphi procedure led to elimination of 30 items, leaving a 45-item questionnaire.

Questionnaire mailing

We invited 4835 residents to complete the preliminary D-RECT; 1276 residents in 26 specialties returned the questionnaire, 53.1% were female. Twenty-five questionnaires that had more than 25 unanswered items were excluded from the analysis. 591 respondents checked 'not applicable' on some items but the response rate per item was never lower than 94%. The values of these items were replaced for the psychometric analysis using two-way imputation, a method that corrects both for person effects and item effects.[28]

Exploratory Factor Analysis

We first analyzed the data using an exploratory factor analysis. We eliminated items with weak factor loadings and followed the advice of the Delphi panel for 70% of the items. This led to the definitive 50-item D-RECT with 11 subscales (table 1). Cronbach's α varied between 0.68 and 0.82 for the different factors.

Testing the *definitive* D-RECT

Confirmatory Factor Analysis

The definitive 50-item D-RECT was tested using confirmatory factor analysis. The goodness-of-fit indices were: CFI 0.90 (≥0.90 indicates adequate fit); RMSEA 0.04 (<0.05 indicates adequate fit) and CMIN which was significant (significance implies poor fit) but this was probably due to the large sample size, because the CMIN/DF (which corrects for large sample sizes) was 2.9 (<5 indicates adequate fit). Overall, the indices showed a good fit (table 2).

Reliability analysis

We examined the generalisability of the 50 items of the definitive D-RECT. In order to obtain reliable outcomes based on the overall score for one department, four residents had to fill out D-RECT. Fifteen residents were needed for reliable outcomes for every subscale in one department (see table 3). For reliable outcomes for groups of departments, two residents from three departments were sufficient to obtain a reproducible total score. For a reliable outcome for every subscale six residents from ten departments were required (table 3).

Table 1: D-RECT questionnaire

	Items	N	Mean	SD	α	Delphi
	Subscale: Supervision				0.74	
1	When I need a consultant, I can always contact one.	1251	4.51	0.70		A
2	When I need to consult a consultant, they are readily available.	1249	4.50	0.72		A
3	The guidelines clearly outline when to request input from a supervisor.	1250	3.77	1.04		A
4	The amount of supervision I receive is appropriate for my level of experience.	1250	4.09	0.92		A
5	It is clear which attending supervises me.	1251	4.22	0.94		A
	Subscale: Coaching and assessment				0.80	
6	I am asked on a regular basis to provide a rationale for my management decisions and actions.	1249	3.86	0.88		A
7	My consultants coach me on how to communicate with difficult patients.	1197	3.40	1.01		NA
8	My consultants take the initiative to explain their actions.	1250	3.42	0.97		NA
9	My consultants take the initiative to evaluate my performance.	1251	3.06	1.05		NA
10	My consultants take the initiative to evaluate difficult situations I have been involved in.	1237	2.97	1.05		A
11	My consultants evaluate whether my performance in patient care is commensurate with my level of training.	1219	3.35	1.07		A
12	My consultants occasionally observe me taking a history.	1154	2.80	1.26		A
13	My consultants assess not only my medical expertise but also other skills such as teamwork, organization or professional behavior.	1245	3.64	1.12		A
14	My consultants give regular feedback on my strengths and weaknesses.	1245	3.33	1.12		A
	Subscale: Observation forms				0.82	
15	Observation forms (i.e. Mini-CEX) are used to structure feedback.	1240	3.79	1.18		NA
16	Observation forms (i.e. Mini-CEX) are used periodically to monitor my progress.	1218	3.43	1.23		NA
	Subscale: Teamwork				0.69	
17	Consultants, nursing staff, other allied health professionals, and residents work together as a team.	1248	3.80	1.03		NA
18	Nursing staff and other allied health professionals make a positive contribution to my training.	1229	3.67	1.07		NA
19	Nursing staff and other allied health professionals are willing to reflect with me on the delivery of patient care.	1225	3.71	1.04		NA
20	Teamwork is an integral part of my training.	1242	2.98	1.04		A
	Subscale: Peer collaboration				0.78	
21	Residents work well together.	1238	4.49	0.76		A
22	Residents, as a group, make sure the day's work gets done.	1225	4.15	1.01		NA
23	Within our group of residents it is easy to find someone to cover or exchange a call.	1226	4.26	0.91		NA
	Subscale: Professional relations between consultants				0.75	
24	Continuity of care is not affected by differences of opinion between consultants.	1236	3.72	1.12		A
25	Differences of opinion between consultants about patient management are discussed in such a manner that is instructive to others present.	1244	3.49	1.10		A
26	Differences of opinion are not such that they have a negative impact on the work climate.	1244	3.62	1.17		A

Table 1: D-RECT questionnaire (continued)

	Items	N	Mean	SD	α	Delphi
	Subscale: Work is adapted to residents' competence				0.68	
27	The work I am doing is commensurate with my level of experience.	1250	4.10	0.83		A
28	The work I am doing suits my learning objectives at this stage of my training.	1249	3.95	0.94		A
29	It is possible to do follow up with patients.	1210	3.96	1.06		A
30	There is enough time in the schedule for me to learn new skills.	1235	3.19	1.08		A
	Subscale: Consultants' role				0.80	
31	My consultants take time to explain things when asked for advice.	1251	4.10	0.77		A
32	My consultants are happy to discuss patient care.	1249	4.12	0.82		A
33	There is (are) NO attending physician(s) who have a negative impact on the educational climate.	1250	3.30	1.27		NA
34	My consultants treat me as an individual.	1250	3.46	1.16		NA
35	My consultants treat me with respect.	1250	4.44	0.86		A
36	My consultants are all in their own way positive role models.	1251	3.44	1.06		NA
	Subscale: Formal education				0.74	
37	Residents are generally able to attend scheduled educational activities.	1248	3.56	1.07		A
38	Educational activities take place as scheduled.	1247	3.88	0.98		A
39	Consultants contribute actively to the delivery of high-quality formal education.	1247	3.57	1.13		A
40	Formal education and training activities are appropriate to my needs.	1247	3.55	1.08		NA
	Subscale: Role of the specialty tutor				0.76	
41	The specialty tutor monitors the progress of my training.	1249	4.46	0.85		A
42	The specialty tutor provides guidance to other consultants when needed.	1203	3.46	1.23		A
43	The specialty tutor is actively involved in improving the quality of education and training.	1247	4.11	1.07		A
44	In this rotation evaluations are useful discussions about my performance.	1191	3.58	1.13		A
45	My plans for the future are part of the discussion.	1179	3.58	1.14		NA
46	During evaluations, input from several consultants is considered.	1188	3.78	1.14		A
	Subscale: Patient sign out				0.74	
47	When there is criticism of a management plan I have developed in consultation with my attending physician, I know the attending physician will back me up.	1196	3.65	1.14		A
48	Sign out takes place in a safe climate.	1190	4.00	1.01		A
49	Sign out is used as a teaching opportunity.	1186	3.90	1.01		A
50	Consultants encourage residents to join in the discussion during sign out.	1184	3.67	1.40		A

A professional translator rendered the original Dutch questionnaire into English. This version was checked by a British medical specialist for clarity. A native speaker translated the questionnaire back into a Dutch version: this version was comparable to the original version.

N= number of respondents, M= mean, SD= standard deviation, α= Cronbach's alpha, Delphi= item accepted (A) or not accepted (NA) by Delphi panel

Table 2: Goodness-of-fit indices

	CFI	RMSEA	CMIN	CMIN/DF
D-RECT	0.90	0.04	Significant	2.9
Conventional cut-off for adequate fit	□0.90	>0.05	Not significant	<5

Table 3: Generalisability analysis of D-RECT total scores and subscales

	SEM	RMSE
	n (residents)	n (residents)/n (departments)
Total Score	4	2/3
Supervision	5	3/3
Coaching and assessment	6	3/3
Observation forms	15	3/10
Team work	7	4/4
Peer collaboration	7	3/4
Professional relations between consultants	10	6/5
Work is adapted to residents' competence	7	3/4
Consultants' attitudes	6	4/4
Formal education	8	4/5
Role of the specialty tutor	7	3/4
Patient handover	9	3/5

SEM= Standard Error of Measurement; number of residents needed to get a reliable result for one department,
RMSE= Root Mean Square Error; number of residents needed to get a reliable outcome for a group of departments (for instance, 3 residents from 3 departments are needed to get a reliable outcome for the factor 'Coaching and assessment).

DISCUSSION

Principal findings

Three different approaches guided the development and validation of D-RECT, a questionnaire to evaluate the postgraduate learning climate. The theoretical foundation of the questionnaire was a model of the clinical learning climate based on earlier empirical findings.[20] A modified Delphi procedure among experts determined the final inclusion and exclusion of items and extensive psychometric analyses revealed a multi-factorial questionnaire with a feasible number of residents required for reliable outcomes.

Strengths and weaknesses

It is a strength of this study that different strategies were used in developing the questionnaire. This lends stronger support to the questionnaire's validity and reliability. Other studies reporting on the development and/or validation of learning climate questionnaires used an undefined 'literature review', parts of previously described questionnaires, or focus groups with ill-described analytical methods as the input for the development of their instrument, and Cronbach's α as the sole indicator of questionnaire stability.[10;13;29] Some studies included an exploratory factor analysis,[17;30] but, to our knowledge, none of the earlier studies combined theoretical input, a Delphi procedure, exploratory and confirmatory analyses, and generalisability studies in developing and validating their instrument. Another strength of this study is that the data for the psychometric analyses were obtained from residents in 26 different specialties, at different levels of training, and from 76 different hospitals. This strengthens the comprehensive applicability of D-RECT. Furthermore, the number of residents needed for a reliable outcome for one department lies between four and fifteen, but most subscales can be judged reliably by as few as eight residents. This supports the feasibility of the instrument. As for groups of departments, even fewer residents (four from four different departments) would yield a reliable impression of most subscales. Other studies have shown similar generalisability outcomes or required (much) larger numbers of participants for a reliable outcome.[14;19;31]

Despite some unique advantages of this study, there are some caveats. First, the 26% response rate was rather low. Nevertheless, the number of respondents is sufficiently high for the psychometric analyses. Second, we conducted our study within the context of postgraduate specialist training in the Netherlands. Whether D-RECT can also be used in international settings remains to be investigated. Another weakness is that we followed the 'advice' of the Delphi panel only partially. Although the majority of the items is in agreement with their judgment, in 30% of the items residents' opinions outweighed the Delphi panel's input. Thus the residents' input may be overrepresented.

Implications and future research

Our study resulted in a multi-factorial questionnaire that can yield reliable results that can be used to give specific feedback to departments on their local learning climate. Based on the questionnaire outcome, a (group of) department(s) can be informed which elements (i.e. subscales) scored highly and should be continued and which factors need to be improved in order to offer a better learning climate for residents. In addition, D-RECT offers benchmarking opportunities. Departments who devote time and effort to creating an optimal learning climate can substantiate their qualities, while departments in need of help can be made aware of their performance in relation to that of other departments. This is also relevant in light of

governmental and public calls for transparency and monitoring of the quality of postgraduate medical education.[2;3]

From a research point of view, it would be interesting to examine D-RECT's consequential validity to strengthen the instrument's evidence base.[32] D-RECT is intended to improve the educational climate. Does it succeed? Are there unintended positive or adverse effects? Another line of investigation would be to validate D-RECT for use in international settings. Finally, D-RECT may enable us to study the relationship between a good learning climate and patient and resident outcomes. Are high D-RECT scores related to the delivery of better patient care? Do they result in better specialists? Research could also examine whether some subscales in particular are associated with the delivery of better patient care or better doctors.

Acknowledgments

The authors thank the participants of the Delphi-panel and all residents who completed D-RECT. We also thank Ron Hoogenboom, research assistant at the University of Maastricht, and Henk van Berkel, PhD and assistant professor at the University of Maastricht, for their assistance in analyzing the data. In addition we thank Mereke Gorsira for her valuable support in editing this paper's English.

The CBOG (an institute involved in residents' training) dispatched the questionnaire and offered administrative support in analyzing the filled out surveys; they had no influence on the content or analysis of the data.

REFERENCES

1. Iverson DJ. Meritocracy in graduate medical education? Some suggestions for creating a report card. Acad.Med 1998;**73**:1223-5.
2. Afrin LB, Arana GW, Medio FJ, Ybarra AF, Clarke HS, Jr. Improving oversight of the graduate medical education enterprise: one institution's strategies and tools. Acad.Med 2006;**81**:419-25.
3. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. Med.Teach. 2001;**23**:445-54.
4. Ashkanasy NM, Wilderom CPM, Peterson MF. Introduction. In Ashkanasy NM, Wilderom CPM, Peterson MF, eds. Handbook of organizational culture & climate, pp 1-18. Thousand Oaks CA: Sage Publications, 2000.
5. Hutchins EB. The 1960 medical school graduate: his perception of his faculty, peers, and environment. J Med Educ 1961;**36**:322-9.
6. Marshall RE. Measuring the medical school learning environment. J Med Educ 1978;**53**:98-104.
7. Rothman AI, Ayoade F. The development of a learning environment: a questionnaire for use in curriculum evaluation. J.Med.Educ. 1970;**45**:754-9.
8. Roff S, McAleer S. What is educational climate? Med Teach 2001;**23**:333-4.
9. Roff S. New resources for measuring educational environment. Med.Teach. 2005;**27**:291-3.
10. Roff S. The Dundee Ready Educational Environment Measure (DREEM)-a generic instrument for measuring students' perceptions of undergraduate health professions curricula. Med.Teach. 2005;**27**:322-5.
11. Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. Med Teach. 2004;**26**:260-4.
12. Kanashiro J, McAleer S, Roff S. Assessing the educational environment in the operating room-a measure of resident perception at one Canadian institution. Surgery 2006;**139**:150-8.
13. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. Med.Teach. 2005;**27**:326-31.
14. Boor K, Scheele F, van der Vleuten CP, Scherpbier AJ, Teunissen PW, Sijtsma K. Psychometric properties of an instrument to measure the clinical learning environment. Med Educ 2007;**41**:92-9.
15. Schonrock-Adema J, Heijne-Penninga M, Hell EA, Cohen-Schotanus J. Necessary steps in factor analysis: Enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. Med Teach. 2008;1-7.
16. Nagraj S, Wall D, Jones E. The development and validation of the mini-surgical theatre educational environment measure. Med Teach. 2007;1-6.
17. Bligh J, Slade P. A questionnaire examining learning in general practice. Med Educ 1996;**30**:65-70.
18. Seelig CB, DuPre CT, Adelman HM. Development and validation of a scaled questionnaire for evaluation of residency programs. South.Med.J 1995;**88**:745-50.
19. Bierer SB, Fishleder AJ, Dannefer E, Farrow N, Hull AL. Psychometric properties of an instrument designed to measure the educational quality of graduate training programs. Eval.Health Prof. 2004;**27**:410-24.
20. Boor, K., Scherpbier, A. J. J. A., Teunissen, P. W., van der Vleuten, C. P. M., and Scheele, F. The postgraduate learning climate: integration of work, training, and residents' needs. 2008 (*under review*)
21. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. Am.J Public Health 1984;**74**:979-83.

22. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995;**311**:376-80.
23. Rowe G, Wright G, Bolger F. Delphi. A reevaluation of research and theory. Technol Forecast Soc Change 1991;**39**:235-51.
24. Holey EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. BMC.Med Res Methodol. 2007;**7**:52.
25. Field A. Exploratory Factor Analysis. Discovering statistics using SPSS, pp 619-80. London: Sage, 2005.
26. McDonald RP, Ho MH. Principles and practice in reporting structural equation analyses. Psychol. Methods 2002;**7**:64-82.
27. Crossley J, Davies H, Humphris G, Jolly B. Generalisability: a key to unlock professional assessment. Med Educ 2002;**36**:972-8.
28. Sijtsma K, van der Ark LA. Investigation and treatment of missing item scores in test and questionnaire data. Multivariate Behav Res 2003;**38**:505-28.
29. Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. Med Teach. 2004;**26**:260-4.
30. Pololi L, Price J. Validation and use of an instrument to measure the learning environment as perceived by medical students. Teach.Learn Med 2000;**12**:201-7.
31. van der Hem-Stokroos HH, van der Vleuten CP, Daelmans HE, Haarman HJ, Scherpbier AJ. Reliability of the clinical teaching effectiveness instrument. Med Educ 2005;**39**:904-10.
32. Beckman TJ, Cook DA, Mandrekar JN. What is the validity evidence for assessments of clinical teaching? J Gen.Intern.Med 2005;**20**:1159-64.

7

General Discussion

INTRODUCTION

Medical students and residents receive their clinical training in preparation for entry into the medical profession within the clinical context of hospital departments. This context, which constitutes the clinical learning climate, has a decisive impact on the learning of students and residents. It is at the centre of the research in this thesis. In order to gain a better understanding of the clinical learning climate we formulated the following central research questions:

Which elements play a part in the clinical learning climate? How can they be measured? What is the clinical learning climate? How can it be defined?

The introductory chapter of this thesis presented a review of studies of organisational climates and cultures and of medical educational climate research. Based on this review we formulated three prerequisites for research into the clinical learning climate. Firstly, it was deemed important to use a combination of qualitative and quantitative methodologies. Secondly, it was important to connect comparative (climate type) research with historically and socially based (culture type) research. And the final prerequisite was to examine and develop instruments for evaluating clinical learning climates. These are the guiding principles of the studies in this thesis.

MAIN FINDINGS AND CONCLUSIONS

Main findings

Following the Introduction, chapter 2 presented a study examining the psychometric properties of an existing questionnaire to measure the clinical learning climate, the Postgraduate Hospital Educational Environment Measure (PHEEM)[1]. On the positive side, we found that PHEEM was reliable with a feasible number of respondents, but on the negative side we found only one sub-scale instead of the three sub-scales hypothesised by the authors of PHEEM, which forced us to conclude that empirical replication of the internal structure had failed. This diminished the strength of PHEEM as an instrument to produce targeted feedback and weakened its construct validity.
We then initiated in-depth explorations of the concept of the clinical learning climate, which are reported in chapters 3-5.
Chapter 3 focused on a person highly influential of the learning climate: the clinical teacher. Residents expressed different and sometimes even contradictory wishes in relation to the 'ideal' clinical teacher but basically they preferred clinical teachers who are responsive to the needs of individual residents. This study also showed that almost half of the characteristics

of the ideal teacher are related to the role of the teacher as a 'person'. Apparently, residents value personal contact with their clinical teachers above the other roles, 'physician', 'supervisor', and 'teacher'.

Combining outcomes of PHEEM and of semi-structured interviews with clinical medical students, the study described in chapter 4 showed that (a) departments and (b) medical students *together* determine the learning climate. Departments with an 'expansive' approach to student participation in department routines, where students are offered opportunities for legitimate participation in multiple activities, evoked positive responses about the department's learning climate. By contrast, the 'restrictive' approach to participation, characterised by limited attention to the learner status of medical students and few opportunities for students to participate in activities, yielded less positive evaluations. Medical students could affect the learning climate by taking (no) initiatives or by (not) expressing their learning needs. These processes were also subject to impact from student motivation, partly governed by 'clerkship fatigue'.

Chapter 5 examined residents' views of the ideal learning climate. The main outcome was that the ideal learning climate integrates work, training, and residents' individual needs. This emphasises the significance of daily work: seeing patients and learning professional skills are pivotal to the learning climate. Another finding was that training and work should be tailored to the needs of individual residents: both should be relevant and at a level that is commensurate with individual residents level of experience and training.

Chapter 6 synthesised the findings from chapters 2 and 5. The PHEEM study, reported in chapter 2, revealed the problem of a one-dimensional measure which lacked a sound theoretical foundation. The study among residents reported in chapter 5 offered a theoretical framework that can potentially remedy this problem. Based on the insights gained from these studies we developed a new instrument, the Dutch Residency Educational Climate Test (D-RECT). After being subjected to testing with a wide array of analytical methods, D-RECT turned out to be a reliable instrument with a sound internal structure.

Conclusions

The studies presented in this thesis examined the clinical learning climate from different angles. In answer to the research question "What elements play a part in the clinical learning climate?" we identified several important elements.

A significant aspect is the '*approach to participation*' that is prevalent in a hospital department. Learning takes place through participation. Emphasis on participation as a core condition for workplace learning is not new.[2-5] Lave & Wenger's work on 'situated learning' - which looks upon all learning as contextual and embedded in a social and physical environment - has been influential in this respect. Research specifically aimed at understanding workplace learning in medical settings underlines the central position of learning by doing.[6-8] Our stud-

ies showed how departments could shape their approaches to trainee participation to create an effective and supportive learning climate. Detailed descriptions of these approaches are given in this chapter under the heading 'The implications for educational practice.'
Another important finding is that the clinical learning climate depends on contributions from both learners and departments; *reciprocal investment* is required. Learners' responsibilities for creating an optimal learning climate include, for instance, pro-active behaviour and an active attitude in seeking attention for their learning needs. It is the task of the departments, for instance, to offer residents and students possibilities for participating in meaningful activities and to promote 'tailor made' approaches to individual residents. Other studies have also stressed how important it is for individual teachers identifying ways to determine the best way to help students and residents.[6;9]

A third key component of the clinical learning climate is the role of *clinical teachers*. It is their role to serve as role models, pass on knowledge and skills, supervise residents' delivery of patient care, and build and maintain a personal bond with each resident. In this way they influence (in)formal residency training, patient care, and residents' personal development.

Finally, the qualitative studies among medical students and residents highlighted *similarities and differences* in what these groups value in the clinical learning climate. Both medical students and residents indicate that work and training should be tailored to their needs, that they themselves should take initiatives to improve the learning climate, and that their 'learner status' should be acknowledged. Differences can be found in roles medical students and residents fulfil: medical students are newcomers, who in the brief time span of a clerkship rotation are rarely able to become full participants in a department, whereas residents can participate fully in many ways, although some would value access to a more differentiated range of activities. Moreover, as 'old-timers' residents have a much bigger impact on teamwork and service delivery. Also, diminished motivation due to a high apprenticeship turnover ('clerkship fatigue') plays a less dominant role in residency, since Dutch residents spend a considerable amount of time in one department.

Besides identifying elements that are important to the clinical learning climate, this thesis aimed to answer the question "How can the learning climate be measured?" Despite its effectiveness and feasibility as an instrument for measuring differences between departments, PHEEM was unfortunately not suitable because it lacked a sufficiently sound internal structure, as appeared from the description of the development process by the original authors[1] as well as our analyses. We developed and tested a new questionnaire, D-RECT, to overcome these deficiencies. The content and internal structure of D-RECT were grounded in the theoretical model described in chapter 5 and fine tuned by input from a group of experts. Factor analysis validated the hypothesised internal structure and generalisability studies revealed

the number of respondents needed for a reliable outcome. This resulted in a questionnaire with 11 sub-scales that could provide reliable outcomes with a feasible number of residents. D-RECT thus appears to be a better alternative to PHEEM for measuring the clinical learning climate for residents.

The final central research questions of this thesis were: "What is the clinical learning climate? And how can it be defined?" Our studies revealed common characteristics and added to the understanding and definition of the clinical learning climate. Integrating definitions of organisational climate[10] and results from this thesis resulted in the following description. The clinical learning climate:

- reflects (restrictive or expansive) approaches to (legitimate) participation,
- is produced by reciprocal interactions between learners and healthcare workers who constitute a clinical department;
- incorporates perceptions shared by these actors of themes like the atmosphere, supervision, the status of learners, and the relations between team members, such as clinical teachers and nurses;
- is shaped by organisational arrangements and artefacts that facilitate or impede learners' participation or more formal forms of learning (such as planned (off the job) teaching);
- influences learners' behaviour and affect;
- becomes manifest in daily practice and activities in which learners participate;
- can differentiate one department from another.

In formulating a theory-based definition of the clinical learning climate, Sfard's distinction between two metaphors of learning may be helpful. According to Sfard, today's discourse on learning is caught between two equally valid metaphors: the *acquisition* metaphor, where learning is viewed as the acquisition and possession of knowledge, and the *participation* metaphor, representing the focus on learning by doing and becoming a part of a greater whole.[11] The acquisition metaphor pictures the traditional outlook on learning, which was dominant in studies on the learning climate in the past. The use of phrases like "attainment of clinical proficiency"[12], "identify lack of knowledge"[13] or "acquisition of factual knowledge"[14] characterises this (implicit) discourse on learning. In this thesis, the clinical learning climate has been examined explicitly from a 'participational' viewpoint, such as by considering how people act in accordance with the norms of a community (intersubjectivity) and the collaboration of team members. Nevertheless, the 'acquisitional' viewpoint has been studied as well (although less extensively), because only one viewpoint would yield an incomplete picture.[11] As an example, table 1 shows how the D-RECT sub-scales relate to either the participation or the acquisition metaphors, or to both.

Apparently, the components of the clinical learning climate as exemplified by the D-RECT sub-scales are predominantly informed and coloured by the participation metaphor. It should

Table 1: Relation of D-RECT subscales to learning metaphors

D-RECT subscales	Participation metaphor	Acquisition metaphor
Supervision	√	√
Coaching and assessment	√	√
Observation forms	√	√
Teamwork	√	
Peer collaboration	√	
Professional relation between consultants	√	
Work adapted to residents' competence	√	√
Consultants' attitudes	√	
Formal education	√	√
Role of specialty tutor	√	√
Patient handover	√	√

be borne in mind, however, that this interpretation of the clinical learning climate should be regarded with some reservations as such descriptions are bound to sway as (educational) insights evolve over time. Nevertheless, this description illuminates which aspects of the learning climate are valued by today's medical students and residents.

STRENGTHS AND LIMITATIONS

The strengths of this thesis are the richness of the data, the methodological thoroughness, and its relevance.

The richness of the data was the result of the use of multiple qualitative approaches. Individual interviews with medical students and residents illuminated the clinical learning climate in ways not published before. Additionally, the input from residents was verified in three focus groups, which yielded further insights. The open-ended questionnaire asking respondents to describe characteristics of the ideal clinical teacher and the Delphi procedure soliciting experts' opinions on D-RECT gave stakeholders yet another opportunity to express their views. The main strength of qualitative research, also in this thesis, is that it enables respondents to answer freely and intuitively, producing a wealth of informative data.

The research methodologies also added to the strength of this thesis. Psychometrically, our analytical methods are far more thorough than those employed in most climate survey studies. As shown in the introductory chapter, many studies use only Cronbach's α to test the reliability of a questionnaire.[15-22] And although validation studies occasionally include

exploratory factor analysis, the outcomes are often used opportunistically (ranging from ignoring the outcomes[23] to following the results without testing them against a theoretical framework[18-20]). In this thesis, however, we performed multiple analyses in addition to those mentioned above (generalisability theory studies, Mokken scale analysis, and confirmatory factor analysis) and our analyses were underpinned by a sound theoretical base.
As for our qualitative analyses, we have consistently used the method described by Miles and Huberman[24] and as a result progressed in expertise in performing and analysing qualitative data. The use of several methods to prevent the drawing of false conclusions, such as member checking or repetitive discussion of findings with an expert group, also contributed to the trustworthiness of the findings.

Finally, we think this thesis presents results that can be of relevance to educational practice. The qualitative studies show ways in which an optimal learning climate can be created. They offer valuable lessons to learners and departments about their contributions to the learning climate.
Our studies of PHEEM and D-RECT yielded information on the questionnaires' pitfalls and possibilities. D-RECT is a new instrument that is not only suitable for comparing different departments but also for giving specific feedback which departments can use to improve their learning climate.

The limitations of this thesis are related to generalisability and bias.

The fact that all the studies were conducted in the Netherlands potentially affects the generalisability of the findings and necessitates further testing of the validity and utility of D-RECT in international settings. Furthermore, the evidence base of the studies was limited because the participants were mostly medical students and residents, although the final study also sought the opinions of other stakeholders (staff members, policy makers). Because D-RECT was validated for residency settings only, its applicability in undergraduate education remains to be investigated.

One should always be on the alert for potential sources of bias in qualitative research, such as false interpretations and disproportionate influence of one researcher on the outcomes. In the studies in this thesis, the main researcher did indeed play a central role in all analyses so bias cannot be ruled out, even though multiple strategies were used, as described above, in order to prevent this.

IMPLICATIONS FOR EDUCATIONAL PRACTICE

Educational practice is changing in undergraduate and postgraduate settings. In the Netherlands, for instance, all 27 Dutch medical specialty boards have developed or are developing new competency-based postgraduate curricula.[25,26] These curricula include 'revolutionary' components, such as mandatory portfolios and clinical assessment of habitual performance, requiring major reforms of residency training programmes.[25,26] Similar approaches have been advocated by accreditation councils in the United States[27] and Canada[28]. The credibility of these approaches is diminished, however, by the sparseness of published studies on workplace learning in clinical settings, notable exceptions being the work by Dornan[29] on medical student learning and Teunissen[30] on resident learning.
Another major change in educational practice is related to residents workload due to the implementation of working time directives which limit residents' working hours.[31,32] In Europe and the United States the number of hours residents are allowed to work has been reduced significantly in the last decade.[33,34] Although this has resulted in decreased exposure to clinical practice during training, it is hoped that this reduction will lead to safer and better patient care as a result of diminished resident fatigue.[33-36]

So curricula and working conditions in clinical medical education are changing and much remains to be investigated about (effective and safe) workplace learning. This thesis offers several insights that can be helpful in understanding and preferably improving clinical educational practice. In the following paragraphs, we present recommendations concerning the use of D-RECT, the value of participation, and the value of being responsive to learners' needs.

Using D-RECT for comparison and improvement

As argued in chapter 1, curriculum changes are costly, both literally and figuratively, as changing daily routines is an arduous task.[37-39] It is therefore essential that such costly changes be evaluated. Moreover, in order for programme changes to be successful, it is essential that they are assessed, any shortcomings are remedied and required adaptations implemented.[40-42] One way of assessing educational programmes is to obtain data on the perceived educational climate.[43] D-RECT has been shown to be a good instrument for this purpose. Not only can it be used to make comparisons between departments with a feasible number of respondents, the sub-scales can alert departments to specific areas where they should direct their improvement efforts. In summary, D-RECT is a feasible instrument to be used by departments and external institutions to compare between departments and to guide improvements.

Value of participation

The findings of this thesis support participation of students and residents in daily activities in the workplace to promote an optimal learning climate.

To shape *medical students'* learning climate departments can consciously decide which approach to participation they prefer. Do they wish to invest in an expansive or a restrictive approach to participation[44]? Programme directors or specialty tutors can foster an expansive approach to clerkships:

- by allocating responsibility for guiding medical students to a specific person or persons;
- by reifying clerkships (e.g. by comprehensive introductory documents, symbols or tools like special 'clerk coffee mugs');
- by offering a schedule with a) possibilities to participate in multiple communities of practice and b) enough off-the-job time for scheduled teaching and reflection.
- by stimulating recognition of the learner status of medical students by all team members (including nursing staff and residents). This implies involving students in daily activities with the express purpose that they should learn from it, as opposed to involvement for convenience's sake.

As for *residency training,* the central role of learning by doing was confirmed once again in this thesis. There are several ways in which departments can optimise the learning climate for residents. First of all, they should be aware of the value of participation in daily activities. Residents place a high value on being given tasks in patient care. This implies that time off-the-job for well-intended formal educational activities does not necessarily enhance their appreciation of the learning climate. What does affect their appreciation, however, is the way in which participation in patient care is implemented. Preferably residents work should be in alignment with their personal learning objectives at that specific stage of their training. Supervisors and specialty tutors should take care to gather information (directly or indirectly) about the training requirements of individual residents in order to provide learning experiences, cases, and feedback tailored to individual residents' stage of training. Supervisors' awareness of individual residents competencies and learning needs is not only beneficial to residents, it will also enhance patient care and patient safety. Secondly, departments should pay attention to team collaboration. (Lack of) cooperation with supervisors, nurses, and peers can (impede or) facilitate full participation.

The importance of responding to learners' needs

The study that focused on the characteristics of the ideal clinical teacher revealed that there is not one uniform ideal teacher for all students. Residents differ in their preferences for dif-

ferent teacher characteristics depending on their personal needs. This is in line with results from the qualitative studies among residents and students showing that in an 'ideal' learning climate learners' activities are fitted to their personal needs. So, it seems wise for educational practitioners to show an interest in the differing needs of individual learners and after establishing these adjust their educational strategies accordingly.

RELEVANCE FOR FUTURE RESEARCH

The results presented in this thesis call for further research in two areas: research related to D-RECT and research to examine the learning climate from different points of view.

The introductory chapter sketched five forms of validity evidence: (1) content, (2) response process, (3) internal structure, (4) relations to other variables, and (5) consequences.[45] In terms of psychometrical research further studies might address D-RECT's *relation to other variables,* such as the association between a good learning climate and the quality of patient care or residents' performance. It should be kept in mind, however, that these associations can be difficult to demonstrate because of the complex and multi-factorial constructs involved.[46] Other potential subjects of further study are *consequential validity* and validation of D-RECT as a measure of the undergraduate learning climate, because so far it has only been validated for residency settings. Moreover, D-RECT will not have eternal life: concepts of the clinical learning climate will probably change over time. In order to anticipate this, we recommend regular 'updating' of D-RECT in response to changing situations to ensure continued content validity. Finally, although further research is required to improve D-RECT, in its current form the instrument can be used as an outcome measure in (intervention) studies aimed to improve the clinical learning climate.

The studies described in this thesis predominantly used residents' and medical students' contributions to make sense of the clinical learning climate. Input from other groups that play a role in the clinical workplace would also be relevant. In our studies clinical nurses affected the learning climate of both residents and medical students considerably. Studies to examine their perceptions of the learning climate and their power to influence it, could increase our understanding of this concept. Another interesting group to study are clinical teachers: how do they picture the ideal clinical learning climate? Different methods, such as observational research, could also afford surprising insights into the learning climate. In this thesis, we have focused on perceptions and descriptions of the respondents' reality, but ethnographic methodologies are likely to expand our understanding of the clinical learning climate in meaningful ways.

FINAL REMARKS

When we undertook this thesis we were convinced that conceptualising the clinical learning climate would be a task that was as intriguing as it was demanding. We have been proven right on both counts. All the hard work has resulted not only in clarifying insights into the nature of the clinical learning climate and factors contributing to it, but it has also produced an instrument to measure the learning climate. These results can be helpful in assessing and understanding the functioning of (changed) curricula. Additionally, there are several practical implications for educational practitioners that can be implemented today. Still, further study is necessary. The clinical learning climate will remain fertile ground for medical education researchers.

REFERENCES

1. Roff S, McAleer S, Skinner A. Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Med.Teach.* 2005;**27**:326-31.
2. Billett S. Situated learning: bridging sociocultural and cognitive theorising. *Learning and Instruction* 1996;**6**:263-80.
3. Billett S. Workplace participatory practices. Conceptualising workplaces as learning environments. *J Workpl Learn* 2004;**16**:312-24.
4. Lave J, Wenger E. Situated Learning: Legitimate peripheral participation. Cambridge: Cambridge University Press, 1991.
5. Wenger E. Communities of Practice. Learning, Meaning, and Identy. Cambridge: Cambridge University Press, 1998.
6. Dornan T, Boshuizen H, King N, Scherpbier A. Experience-based learning: a model linking the processes and outcomes of medical students' workplace learning. *Med Educ* 2007;**41**:84-91.
7. Teunissen PW, Scheele F, Scherpbier AJJA, van der Vleuten CPM, Boor K, Van Luijk SJ *et al*. How residents learn: qualitative evidence for the pivotal role of clinical activities. *Med Educ* 2007;**41**:763-70.
8. Teunissen PW, Boor K, Scherpbier AJ, van der Vleuten CP, Diemen-Steenvoorde JA, Van Luijk SJ *et al*. Attending doctors' perspectives on how residents learn. *Med Educ* 2007;**41**:1050-8.
9. Elliott RL, Juthani NV, Rubin EH, Greenfeld D, Skelton WD, Yudkowsky R. Quality in residency training: toward a broader, multidimensional definition. *Acad.Med.* 1996;**71**:243-7.
10. Moran ET, Volkwein JF. The cultural approach to the formation of organizational climate. *Hum Relat* 1992;**45**:19-47.
11. Sfard A. On two metaphors for learning and the dangers of choosing just one. *Educational Res* 1998;**27**:4-13.
12. Hutchins EB. The 1960 medical school graduate: his perception of his faculty, peers, and environment. *J Med Educ* 1961;**36**:322-9.
13. Hutchinson L. Educational environment. *BMJ* 2003;**326**:810-2.
14. Patel VL, Dauphinee WD. The clinical learning environments in medicine, paediatrics and surgery clerkships. *Med.Educ.* 1985;**19**:54-60.
15. Cassar K. Development of an instrument to measure the surgical operating theatre learning environment as perceived by basic surgical trainees. *Med Teach.* 2004;**26**:260-4.
16. Kanashiro J, McAleer S, Roff S. Assessing the educational environment in the operating room-a measure of resident perception at one Canadian institution. *Surgery* 2006;**139**:150-8.
17. Marshall RE. Measuring the medical school learning environment. *J Med Educ* 1978;**53**:98-104.
18. Nagraj S, Wall D, Jones E. The development and validation of the mini-surgical theatre educational environment measure. *Med Teach.* 2007;1-6.
19. Oliveira Filho GR, Vieira JE, Schonhorst L. Psychometric properties of the Dundee Ready Educational Environment Measure (DREEM) applied to medical residents. *Med.Teach.* 2005;**27**:343-7.
20. Pololi L, Price J. Validation and use of an instrument to measure the learning environment as perceived by medical students. *Teach.Learn Med* 2000;**12**:201-7.
21. Roff S, McAleer S, Harden RM, Al-Qahtani M, Ahmed AU, Deza H *et al*. Development and validation of the Dundee Ready Education Environment Measure (DREEM). *Med Teach* 1997;**19**:295-9.
22. Rotem A, Godwin P, Du J. Learning in hospital settings. *Teach Learn Med* 1995;**7**:211-7.
23. Holt MC, Roff S. Development and validation of the Anaesthetic Theatre Educational Environment Measure (ATEEM). *Med Teach.* 2004;**26**:553-8.

24. Miles MB, Huberman AM. Qualitative data analysis: an expanded sourcebook. Thousand Oaks, California: Sage, 1994.
25. Scheele F, Teunissen P, Van Luijk S, Heineman E, Fluit L, Mulder H *et al*. Introducing competency-based postgraduate medical education in the Netherlands. *Med.Teach.* 2008;**30**:248-53.
26. Ten Cate O. Medical education in The Netherlands. *Med.Teach.* 2007;**29**:752-7.
27. Swing SR. The ACGME outcome project: retrospective and prospective. *Med.Teach.* 2007;**29**:648-54.
28. Societal Needs Working Group. Skills for the new millennium: CanMEDS 2000 project. *Ann R Coll Phys Surg Can* 1996;**29**:206-16.
29. Dornan, T. Experience based learning. Learning clinical medicine in workplaces. 2006. University of Maastricht. (thesis)
30. Teunissen, PW. Unravelling learning by doing. A study of workplace learning in postgraduate medical education. 2009. Free University, Amsterdam. (thesis)
31. Dola C, Nelson L, Lauterbach J, Degefu S, Pridjian G. Eighty hour work reform: faculty and resident perceptions. *Am.J Obstet.Gynecol.* 2006;**195**:1450-6.
32. Yedidia MJ, Lipkin M, Jr., Schwartz MD, Hirschkorn C. Doctors as workers: work-hour regulations and interns' perceptions of responsibility, quality of care, and training. *J.Gen.Intern.Med.* 1993;**8**:429-35.
33. Fletcher KE, Underwood W, III, Davis SQ, Mangrulkar RS, McMahon LF, Jr., Saint S. Effects of work hour reduction on residents' lives: a systematic review. *JAMA* 2005;**294**:1088-100.
34. Waurick R, Weber T, Broking K, Van Aken H. The European Working Time Directive: effect on education and clinical care. *Curr.Opin.Anaesthesiol.* 2007;**20**:576-9.
35. Fletcher KE, Davis SQ, Underwood W, Mangrulkar RS, McMahon LF, Jr., Saint S. Systematic review: effects of resident work hours on patient safety. *Ann Intern.Med.* 2004;**141**:851-7.
36. Lockley SW, Landrigan CP, Barger LK, Czeisler CA. When policy meets physiology: the challenge of reducing resident work hours. *Clin.Orthop.Relat Res.* 2006;**449**:116-27.
37. Bland CJ, Starnaman S, Wersal L, Moorehead-Rosenberg L, Zonia S, Henry R. Curricular change in medical schools: how to succeed. *Acad.Med.* 2000;**75**:575-94.
38. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. *Med J Aust.* 2004;**180**:S57-S60.
39. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;**362**:1225-30.
40. Wilkes M, Bligh J. Evaluating educational interventions. *BMJ* 1999;**318**:1269-72.
41. Hayden J, McKinlay D. Monitoring standards of training. *Med.Educ.* 2001;**35**:68-72.
42. Davis DJ, Ringsted C. Accreditation of undergraduate and graduate medical education: how do the standards contribute to quality? *Adv.Health Sci.Educ.Theory.Pract.* 2006;**11**:305-13.
43. Genn JM. AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education - a unifying perspective. *Med.Teach.* 2001;**23**:445-54.
44. Fuller A, Unwin L. Learning as apprentices in the contemporary UK workplace: creating and managing expansive and restrictive participation. *J Educ Work* 2003;**16**:407-26.
45. American Education Research Association and American Psychological Association. Standards for Educational and Psychological Testing. Washington DC: American Education Research Association, 1999.
46. Scott T, Mannion R, Marshall M, Davies H. Does organisational culture influence health care performance? A review of the evidence. *J.Health Serv.Res.Policy* 2003;**8**:105-17.

Summary

SUMMARY

Medical students and residents learn in specific contexts. These contexts constitute their clinical learning climates, which are generally regarded as an important indicator of the educational quality of a department that provides clinical training. The clinical learning climate is the subject of this thesis.
The setting of the studies in this thesis is the clinical phase of undergraduate medical training, also referred to as clerkships, and postgraduate specialist training in the Netherlands. In this thesis postgraduate trainees at all levels of training are referred to as residents.

To provide some insight into what is already known about learning climates, **Chapter 1** summarises the organisational psychology literature and the medical education literature on (learning) climates.
Within the *organisational psychology literature* the study of organisational climates started more than half a century ago. The most frequently used methodology is questionnaires for quantitative comparisons between organisations. Less attention is given to defining the organisational climate. Later, organisational culture researchers focus on the study of individual organisations as opposed to the comparative studies favoured by researchers of the organisational climate. Culture researchers predominantly employ observational and anthropological methodologies. The last decade these two research groups have tended to integrate and incorporate ideas and research methodologies from each other.
The *medical education literature* comprises multiple studies of the educational climate in medical schools. Most studies describe instruments to evaluate the learning climate or use these instruments to compare between different groups. The psychometric properties of instruments are only rarely described and this impedes evaluation of the instruments' psychometric quality. A clear definition of the learning climate is lacking. In addition, several authors have published qualitative studies on this subject.
Despite the abundance of studies on the concept of (clinical) learning climates, it remains to be explored what this climate is and which factors play a part. This thesis aims to answer the following research questions: What elements play a part in the clinical learning climate? How can it be measured? What is the clinical learning climate? And how can it be defined?

Chapter 2 describes the psychometric properties of an existing instrument: the Postgraduate Hospital Educational Environment Measure (PHEEM). The instrument is relevant for this thesis because it is specifically designed to measure the clinical learning climate. The 40-item questionnaire consists of items such as feedback, atmosphere and facilities. The original designers of the questionnaire identified three sub-scales: perception of autonomy, perception of teaching, and perception of social support, but failed to test PHEEM's *construct validity* (the agreement between the hypothesised sub-scales and the measuring device) or *reliability*

(number of respondents needed for a reproducible evaluation of departments or groups of departments). This chapter describes these two psychometric properties of PHEEM as determined in studies among medical students and residents.
PHEEM questionnaires completed by 256 medical students and 339 residents are analysed. Factor analysis and a Mokken-scale analysis fail to replicate the hypothesised three subscales. The analyses show only one dimension. Generalisability analysis shows that a feasible sample size suffices for a reliable outcome.
To sum up, this study indicates that PHEEM can demonstrate differences between departments with a feasible number of respondents. The (factor and Mokken-scale) analyses, however, do not confirm the construct hypothesised by the original designers: PHEEM measures only one construct. What this construct exactly consists of is a topic for further research.

Chapter 3 reports the results of a study of the characteristics of the ideal clinical teacher. Clinical teachers are important in the education of medical students and residents and are a major determinant of the clinical learning climate. In 1994 and 2003 residents in obstetrics-gynaecology were asked to state the characteristics of the ideal clinical teacher. Their answers are analysed using two strategies.
The first analysis is based on the 'traditional' medical education literature. Within this field many studies describe the desired characteristics of an ideal (clinical) teacher. Residents' answers fit within an earlier described framework of roles: the role of Physician, Supervisor, Teacher, and Person. Interestingly, almost half of all the responses correspond to the Person role. The second most frequently mentioned role in 1994 is the Physician role, whilst in 2003 it is the Supervisor role.
Next, a sociocultural perspective is used to analyse the data. This outlook emphasises the importance of social contexts and relationships within which learning takes place: a resident decides whether a teacher is ideal while interacting with this teacher. The characteristics of the ideal clinical teacher are related to direct interaction in 70% of the cases; 30% of the responses concern indirect interactions. An interesting finding is that residents value different, even opposite, characteristics of their ideal clinical teacher, depending on their personal needs.
This study emphasises the importance of personal interaction between residents and clinical teachers. The first analysis shows that residents value the Person role of their teachers. Between 1994 and 2003 the second most valued teacher characteristic changes, possibly due to the implementation of a more learner-centred curriculum. The second analysis shows that residents appreciate different characteristics. The ideal clinical teacher uses interaction to determine the specific needs of individual residents.

Chapter 4 studies medical students' perceptions of the clinical learning climate. Medical students from twelve different obstetrics-gynaecology departments complete the PHEEM.

Students from the highest and lowest scoring department are invited to participate in semi-structured individual interviews. An analysis method described by Miles and Huberman guides the interpretation of the data.
The results show that the department with the highest PHEEM score differs significantly from the department with the lowest PHEEM score. From these two departments fourteen medical students participate in interviews lasting 45-90 minutes. Analysis of these interviews shows that departments and medical students together determine the learning climate. Departments leave their mark on the learning climate by their approach towards student participation. This approach is manifested in the legitimacy of student participation in the department, clerkship organisation, and the involvement of teachers in medical students' personal development. Medical students can influence the learning climate in the initial phase of a clerkship by developing initiatives and feeling out the local dos and don'ts, and, later in the clerkship, by increasingly showing initiative or, by contrast, ceasing to do so. Some medical students talk about 'clerkship fatigue' caused by too many previous clerkship experiences and severely reducing their motivation to once again try and adapt to new local customs.
In summary, this study shows that departments and medical students *together* determine the clinical learning climate. How medical students participate is central to their perceptions of clinical learning climates. Departments can approach student participation 'expansively' or 'restrictively'. Another finding is that medical students seek to understand the explicit and implicit expectations and customs that characterise a department. If they can comply with the local conventions of a department, their learning climate is better. Departments can make it easier for medical students to understand departmental customs by being open about them.

In addition to *medical students'* perceptions, described in chapter 4, **chapter 5** depicts *residents'* perceptions of the learning climate. This two-step study aims to answer how optimal learning climates of residents are shaped. The first step consists of semi-structured interviews with fourteen residents from different specialties. In the second step of this study the interview findings are tested and further refined in three focus group discussions. The analysis of the data is again guided by the method described by Miles and Huberman.
Three interrelated themes are important in the eyes of the residents. An optimal learning climate combines (1) work and (2) training, tailored to (3) residents' specific attributes and needs. These three themes converge when a patient is discussed properly, when there is a safe atmosphere, and when there is a positive relationship with the supervisor (as is stressed in chapter 3 as well). Within the work theme residents value a well-balanced distribution of service and education, good collaboration with the team, and opportunities to influence daily work and adjust it to their personal developmental needs. Within their training residents value active learners (who seek feedback and are pro-active) and active supervisors (who

take initiative to give feedback, teach, and assess) as well as a mutual introduction. Meetings where work related issues are discussed in an informative way improve the learning climate. To sum up, this study illustrates that an optimal learning climate facilitates both 'traditional' learning (such as self study) and learning through participating in daily practice. In addition departments should appreciate residents' previous experiences and bring work and training in alignment with these experiences. Reciprocity with, for instance, nurses is an important theme as well: an imbalance can have a negative impact on the learning climate. Finally, residents state that small changes in daily practice can improve the learning climate.

Chapter 6 gives an account of the development and psychometric properties of a newly developed questionnaire to test the educational climate of residents, the Dutch Residency Educational Climate Test (D-RECT). This study aims to overcome the shortcomings described in chapters 1 and 2 of many questionnaires to measure the educational climate, such as absence of theoretical underpinnings and incompletely described psychometric properties. The definition of the learning climate described in chapter 5 serves as the theoretical basis for the development of the *preliminary* 75-item D-RECT. Thirty-eight experts judge this instrument in a Delphi procedure: in multiple rounds they indicate whether an item is to be included in the definitive D-RECT. Concurrently, 1278 residents in 26 specialties complete the preliminary D-RECT. The results of exploratory factor analysis of these questionnaires combined with the results of the Delphi panel lead to the *definitive* 50-item D-RECT. Further analysis of the instrument involves confirmatory factor analysis and a generalisability study. This results in eleven sub-categories representing 50 items and the finding that a reliable outcome can be achieved with a feasible number of residents.
In conclusion, chapter 6 describes the development of a valid and reliable instrument to measure the learning climate of residents. D-RECT differs from other instruments, like the PHEEM, in having a solid theoretical basis, experts' input, and well-described and promising psychometric properties.

Chapter 7 summarises the previous chapters and draws several conclusions; describes the strengths and weaknesses of the studies, and presents educational implications and suggestions for further research.
As for conclusions, it is stressed that it is impossible to underestimate the importance of the value of learning through participation in daily activities and through becoming part of a team. In addition this thesis shows that the learning climate demands a reciprocal investment from learner and department. Also the pivotal role of clinical teachers is illuminated. Studies of instruments to measure the learning climate show D-RECT to be superior to the PHEEM, because of D-RECT's theoretical basis and psychometric properties and its possibilities to offer focused feedback to departments. Finally, this chapter presents a definition of the clinical learning climate based on the combined findings of the studies in this thesis.

These conclusions show that 'learning' is seen as more than the acquisition of knowledge and skills (so called 'acquisition learning'); 'learning' is also a continuous process in which a person becomes a part of a greater whole (so called 'participation learning').

The analysis of the strengths and weaknesses shows that the strengths of this thesis are the information-rich data, the various methodologies used to analyse the qualitative data and evaluate the measuring instruments, and the relevance of the findings to the daily practice of medical education. Weaknesses are the limited generalisability of the findings and the possibility of bias in the data analysis.

There are several implications for educational practice. In a time of drastic changes in post-graduate training, D-RECT can offer insights into the consequences of these changes and provide recommendations for which changes to retain or adjust so as to optimise the learning climate. Additionally, this thesis shows how departments can take a constructive approach to participation by learners. Finally, multiple studies stress that it is of major importance for supervisors to take in interest in the needs of learners and respond to them appropriately.

As for research, the psychometric quality of the D-RECT requires further study in other settings and D-RECT can serve as an outcome measure for intervention studies aimed at improving the educational climate. Studies of the perceptions of other parties involved in clinical training, such as nurses, may offer valuable insights, while observational research is likely to illuminate other interesting aspects of the clinical learning climate.

Samenvatting

SAMENVATTING

Co-assistenten en aios leren in een bepaalde context het vak van basisarts of specialist. Deze context is het klinisch opleidingsklimaat. Dit klimaat wordt algemeen erkend als een belangrijke indicator van de onderwijskwaliteit van een bepaalde afdeling en is het onderwerp van dit proefschrift.
De onderzoeken beschreven in dit proefschrift vonden allen plaats in het klinische deel van de opleiding tot basisarts of specialist in Nederland. Medisch studenten die stages lopen op afdelingen in ziekenhuizen worden 'co-assistenten' genoemd; basisartsen die in opleiding zijn tot specialist heten 'arts in opleiding tot specialist' kortweg aios.

Om inzicht te krijgen in wat er reeds bekend is omtrent het (opleidings)klimaat wordt in **hoofdstuk 1** zowel de organisatie psychologische als de medisch onderwijskundige literatuur geraadpleegd.
De *organisatie psychologische literatuur* laat zien dat er al meer dan een halve eeuw onderzoek wordt gedaan naar organisatie klimaten. De meest gebruikte onderzoeksmethode is het gebruik van vragenlijsten die het mogelijk maken organisaties kwantitatief met elkaar te vergelijken. Wat het organisatie klimaat precies is, krijgt minder aandacht. Later komt een tegenbeweging op van organisatie cultuur onderzoekers, die zich niet richten op het meten en vergelijken van verschillende organisaties, maar meer verdieping zoeken in het begrijpen van individuele organisaties. Hier past men met name observationele, antropologische onderzoeksmethoden toe. De laatste jaren is integratie zichtbaar tussen deze twee voorheen gescheiden onderzoekscholen.
De *medisch onderwijskundige literatuur* toont een veelheid aan klimaatstudies die zich voornamelijk richten op het opleidingsklimaat aan medische faculteiten (in de niet klinische fase). De meeste onderzoeken beschrijven meetinstrumenten of rapporteren uitkomsten van vergelijkingen tussen afdelingen met behulp van dergelijke meetinstrumenten. De psychometrische kenmerken van de vragenlijsten zijn vaak nauwelijks beschreven wat het moeilijk maakt de psychometrische kwaliteit te beoordelen. Een heldere omschrijving van het opleidingsklimaat ontbreekt. Hiernaast zijn de laatste jaren zijn ook enkele kwalitatieve onderzoeken gepubliceerd.
Er is dus reeds veel over dit concept geschreven, maar wat het opleidingsklimaat precies is en welke factoren een rol spelen bij de totstandkoming ervan blijft ongewis. Dit proefschrift probeert de volgende onderzoeksvragen te beantwoorden: Welke factoren spelen een rol in het klinisch opleidingsklimaat? Hoe kan het gemeten worden? Wat is het klinisch opleidingsklimaat? En hoe kan het worden gedefinieerd?

Hoofdstuk 2 beschrijft een studie naar de psychometrische kenmerken van een bestaand instrument: de 'Postgraduate Hospital Educational Environment Measure' (de PHEEM). Deze

lijst is relevant voor dit proefschrift omdat het zich richt op het meetbaar maken van klinische opleidingsklimaten. De vragenlijst bestaat uit 40 stellingen en beslaat onderwerpen als feedback, atmosfeer en faciliteiten. De ontwikkelaars van de lijst herkenden drie subschalen: perceptie van autonomie, perceptie van onderwijs en perceptie van sociale steun. Zij hebben echter noch de *construct validiteit* (de overeenkomst tussen vooraf bedachte subschalen en het daadwerkelijke instrument) noch de *betrouwbaarheid* (de hoeveelheid respondenten nodig voor een reproduceerbaar oordeel over een (groep) afdeling(en)) onderzocht. Dit hoofdstuk beschrijft deze twee psychometrische kenmerken voor zowel co-assistenten als aios.

In totaal hebben 256 co-assistenten en 339 aios de PHEEM ingevuld. Deze vragenlijsten zijn nader geanalyseerd met verschillende statistische methodieken. De drie subschalen bleken niet met een factoranalyse of Mokkenschaal analyse te repliceren. De analyses toonden aan dat de PHEEM een uni-factoriële (en dus niet multi-factoriële) vragenlijst is. De generaliseerbaarheidsanalyses lieten zien dat met een redelijk aantal co-assistenten of aios (groepen) afdelingen betrouwbaar beoordeeld konden worden.

Concluderend toont deze studie dat de PHEEM in staat is verschillen tussen afdelingen aan te tonen met een redelijk aantal respondenten, maar dat het onderliggende construct anders is dan werd verondersteld door de ontwikkelaars van de lijst: psychometrisch bleek de PHEEM één samenhangend construct te meten. Wat dit onderliggende construct precies is, is een onderwerp wat verder onderzocht moet worden.

Hoofdstuk 3 schetst een onderzoek naar 'de ideale opleider'. Opleiders spelen een belangrijke rol in de opleiding van co-assistenten en aios en beïnvloeden het opleidingsklimaat. In 1994 en 2003 is aios obstetrie en gynaecologie gevraagd wat kenmerken zijn van ideale opleiders. Hun antwoorden zijn geanalyseerd op twee verschillende manieren.

Allereerst is geput uit de 'traditionele' medisch onderwijskundige literatuur waar reeds veel onderzoek gepubliceerd is over de karakteristieken die ideale opleiders zouden moeten bezitten. De antwoorden van de aios bleken te passen binnen de volgende, in de literatuur beschreven, rollen: de Dokter, de Supervisor, de Leraar en de Mens. Opvallend was dat bijna 50% van de antwoorden, in beide onderzoeksjaren, paste binnen de rol Mens. Als tweede meest genoemde rol kwam in 1994 de rol van Dokter naar voren; in 2003 was dit de rol van Supervisor.

In de tweede plaats werd vanuit een socio-culturele invalshoek geanalyseerd. Dit perspectief geeft meer aandacht aan hetgeen er tussen aios en opleiders gebeurt: in interactie wordt de ideale opleider vastgesteld. De karakteristieken van de ideale opleider bleken in 70% van de antwoorden gerelateerd aan directe interactie, de resterende 30% ging over indirecte interactie. Een interessante bevinding was dat aios heel verschillende karakteristieken als ideaal bestempelden, afhankelijk van hun persoonlijke behoefte.

Samenvattend benadrukt deze studie het belang van persoonlijk contact tussen aios en opleiders. De eerste analysemethode toonde dat aios veel belang hechten aan de rol van Mens in hun opleider. Ook toonde deze analyse dat de verwachtingen door de tijd heen veranderden: mogelijk ten gevolge van de meer leerling gecentreerde opleiding vond een verschuiving plaats van waardering voor Leraar karakteristieken naar Supervisor karakteristieken. Uit de tweede analyse bleek dat aios niet allemaal hetzelfde zoeken in hun opleider. In interactie kan een ideale opleider vaststellen wat een specifieke aios nodig heeft.

Hoofdstuk 4 onderzoekt welke percepties co-assistenten van twee zeer verschillende afdelingen hebben van hun opleidingsklimaat. Om twee contrasterende afdelingen te kunnen identificeren werden op 12 gynaecologische afdelingen PHEEM's verspreid die per afdeling door 14 co-assistenten werden ingevuld. Co-assistenten van de hoogst en de laagst scorende afdelingen werden benaderd voor deelname aan individuele, semigestructureerde interviews. Een analyse methode beschreven door Miles en Huberman leidde de interpretatie van de data.
De resultaten toonden dat er een significant verschil bestond tussen de afdeling met de hoogste PHEEM score en de afdeling met de laagste PHEEM score. Van deze twee afdelingen werden 14 co-assistenten geïnterviewd gedurende 45 tot 90 minuten. De analyse van deze interviews toonde dat zowel afdelingen als co-assistenten zelf het uiteindelijk ervaren opleidingsklimaat bepalen. Afdelingen beïnvloedden het opleidingsklimaat met hun benadering van participatie: dit uitte zich in de mate van legitimiteit die co-assistenten ervoeren, de organisatie van het co-schap en de betrokkenheid bij de persoonlijke ontwikkeling van de co-assistent. Co-assistenten konden het opleidingsklimaat beïnvloeden zowel in de initiële fase- door initiatieven te ontplooien en af te tasten wat er van je verwacht werd- als tijdens het verdere co-schap- door steeds meer initiatief te nemen of juist steeds meer ge*de*motiveerd te raken. Enkele co-assistenten vertelden over zogenaamde 'co-schap moeheid', wat betekende dat ze bij het zoveelste co-schap weinig gemotiveerd bleken zich opnieuw te verdiepen in de lokale mores.
Uit dit onderzoek blijkt dus dat afdelingen en co-assistenten *samen* het uiteindelijke opleidingsklimaat bepalen. De manier waarop co-assistenten participeren speelt een centrale rol in hun perceptie van dit opleidingsklimaat. Afdelingen kunnen deze participatie 'expansief' dan wel 'restrictief' benaderen. Een andere bevinding is dat co-assistenten ernaar streven te begrijpen wat een afdeling van hen verwacht. Het levert een positiever opleidingsklimaat op als co-assistenten de ongeschreven regels leren herkennen. Afdelingen kunnen op hun beurt open zijn over de afdelingsgewoontes waarmee ze het makkelijker maken voor co-assistenten om zich aan te passen.

Naast de in het vorige hoofdstuk beschreven percepties van *co-assistenten* van hun opleidingsklimaat, is ook studie naar *aios*' percepties gedaan; **hoofdstuk 5** zet deze percepties

uiteen. In twee stappen werd antwoord gezocht op de vraag hoe optimale opleidingsklimaten tot stand komen. De 1e stap bestond uit semigestructureerde interviews met 14 aios van diverse disciplines. Vervolgens werden de bevindingen uit deze interviews getest in de 2e stap van het onderzoek: in 3 focusgroepen (geleide groepsdiscussies) werd er met aios gesproken over hun visie op het opleidingsklimaat. De bevindingen uit de interviews konden naar aanleiding van deze focusgroepen verder worden aangescherpt. De analyse vond opnieuw plaats met behulp van de analyse strategie beschreven door Miles en Huberman.
Uit de analyse bleek dat aios drie thema's van belang vonden binnen het opleidingsklimaat. Idealiter komt (1) het werk overeen met (2) de opleiding die (3) een specifieke aios op dat moment nodig heeft. Een goede en relevante patiëntbespreking, een veilig klimaat en een prettig persoonlijk contact met de supervisor (waarvan het belang ook al in hoofdstuk 3 bleek) waren uitingen van de samenkomst van deze drie thema's. Binnen het werk hechtten aios belang aan een goede balans tussen productie en opleiding, een goede samenwerking met het gehele team en mogelijkheden om je dagelijkse werkzaamheden te beïnvloeden en aan te passen een je persoonlijke ontwikkeling behoeften. Binnen hun opleiding vonden aios een actieve aios-rol (zelf feedback zoeken, studeren, pro-actief zijn) en een actieve supervisor-rol (zelf initiatief nemen om onderwijs, feedback en beoordelingen te geven) belangrijk, evenals een wederzijdse introductie. Daarnaast waren bijeenkomsten waarin werkgerelateerde onderwerpen illustratief werden besproken van belang voor het opleidingsklimaat.
Concluderend blijkt uit deze studie dat een optimaal opleidingsklimaat zowel het 'traditionele' leren (zoals zelfstudie) faciliteert, als leren door participatie in de dagelijkse praktijk. Daarnaast is het van belang te beseffen dat aios ervaringen meebrengen naar de werkvloer waarbij hun opleiding en werkzaamheden moeten aansluiten. Op de werkvloer bleek wederkerigheid in relatie tot bijvoorbeeld verpleegkundigen een belangrijk thema; aios ervoeren soms een disbalans wat het opleidingsklimaat negatief beïnvloedde. Aios verwoordden verder dat met kleine veranderingen in de dagelijkse praktijk het opleidingsklimaat zou kunnen verbeteren.

Hoofdstuk 6 beschrijft de ontwikkeling en psychometrische kenmerken van de Dutch Residency Educational Climate Test (D-RECT), een nieuw instrument om het opleidingsklimaat van aios te meten. Dit onderzoek tracht de in hoofdstuk 1 en 2 beschreven tekortkomingen van veel klimaat meetinstrumenten (het ontbreken van een theoretisch construct, gebrekkige psychometrische kenmerken) te ondervangen. Bij de ontwikkeling is gebruik gemaakt van de omschrijving van het opleidingsklimaat beschreven in hoofdstuk 5: dit vormde de theoretische basis voor de *voorlopige* D-RECT, bestaande uit 75 items. Deze lijst is beoordeeld door 38 experts in een Delphi-procedure. Per item werd bepaald of dit item behouden moest blijven voor de definitieve lijst of niet. De lijst werd tegelijkertijd ingevuld door 1278 aios van 26 verschillende specialismen. Deze ingevulde lijsten werden met behulp van een exploratieve factoranalyse geanalyseerd. De uitkomsten van deze analyse en van het Delphi panel

leidde tot de *definitieve* D-RECT, bestaande uit 50 items. Deze lijst werd nader onderzocht met een confirmatorische factor analyse en een generaliseerbaarheid analyse. Hieruit bleek dat de 50 items 11 subcategorieën representeerden. Ook bleek dat met een beperkt aantal aios een betrouwbaar oordeel over een afdeling kan worden verkregen.
Concluderend beschrijft hoofdstuk 6 de ontwikkeling van een valide en betrouwbaar instrument om het opleidingsklimaat van aios te meten. De D-RECT onderscheidt zich van andere instrumenten- zoals de PHEEM- door een theoretische fundering, input van experts en uitgebreid statistisch onderzoek naar de psychometrische kwaliteit met goede uitkomsten.

Hoofdstuk 7 vat de bevindingen van de vorige hoofdstukken samen en trekt op basis daarvan enkele conclusies, geeft een sterkte-zwakteanalyse van het verrichte onderzoek en geeft adviezen op onderwijskundig en onderzoek gebied.
Dit hoofdstuk concludeert dat de rol van het leren door te participeren in dagelijkse activiteiten en onderdeel te worden van een team niet onderschat moet worden. Daarnaast komt naar voren dat het opleidingsklimaat een investering vraagt van de zowel de 'leerling' als de afdeling en blijkt de belangrijke rol van de supervisor als opleider. Het onderzoek van meetinstrumenten van het opleidingsklimaat toont dat de D-RECT, door de theoretische en psychometrische onderbouwing en de mogelijkheden om specifiek feedback aan afdelingen te geven, een beter instrument lijkt te zijn om het opleidingsklimaat te meten in vergelijking met de PHEEM. Hiernaast geeft hoofdstuk 7 een definitie van het opleidingsklimaat gebaseerd op de bevindingen uit alle onderzoeken.
Uit bovenstaande conclusies blijkt dat 'leren' binnen dit proefschrift niet alleen gezien wordt als het je eigen maken van kennis of vaardigheden (het zogenaamde acquisitie leren), maar 'leren' wordt ook gezien als een continue ontwikkeling door deel uit te (gaan) maken van een groter geheel (het zogenaamde participatie leren).
Uit de sterkte-zwakteanalyse blijkt dat de kracht van de onderzoeken ligt in de informatierijke data die is verzameld, in de methodologie die is gebruikt zowel om de kwalitatieve data te analyseren als om de meetinstrumenten te onderzoeken en in de relevantie van de bevindingen voor de dagelijkse praktijk. Zwakke punten zijn de generaliseerbaarheid van de bevindingen en mogelijke bias die is opgetreden bij de analyse van de data.
De onderzoeken van dit proefschrift hebben onderwijskundige implicaties. In deze tijd met veel veranderingen in de klinische fase van de medische opleiding, kunnen metingen met behulp van de D-RECT inzicht geven in hoe die veranderingen uitpakken en welke zaken behouden of verbeterd zouden kunnen worden om het opleidingsklimaat te verbeteren. Daarnaast geven de studies inzicht hoe afdelingen op een constructieve manier om kunnen gaan met participatie. Tenslotte tonen meerdere onderzoeken het belang voor begeleiders om zich te verdiepen in hetgeen de leerling nodig heeft en daarop te reageren.
Dit proefschrift geeft ook aanleiding tot verder onderzoek. Met betrekking tot de D-RECT kan de psychometrische kwaliteit nog verder worden onderzocht (onder meer in andere

settings), en kan de D-RECT gebruikt worden om het effect te meten van interventies gericht op klimaat verbetering. Verder zou onderzoek naar percepties van andere betrokkenen, zoals bijvoorbeeld verpleegkundigen, informatie kunnen toevoegen evenals observationeel onderzoek.

Dankwoord

DANKWOORD

De eerste kennismaking met de medische onderwijswereld staat me nog helder voor de geest. Ik mocht een poster presenteren op het AMEE congres in Zwitserland. In twee auto's togen we naar Bern, waar we ons kamp opsloegen op de plaatselijke camping. Godzijdank regende het vier dagen vrijwel niet en ik koester mooie herinneringen aan wandelingen langs de rivier, gesprekken op terrassen, omkleedpartijen in het toiletgebouw en- o, ja- de geweldige inhoud van het congres. Het spreekt vanzelf dat na deze glamoureuze introductie mijn liefde voor dit vakgebied en de mensen die er actief in zijn werd geboren.

Beste Fedde, jij ben een van de redenen geweest waarom ik dit onderzoek heb gekozen, het leek me verfrissend met zo'n onconventionele promotor het promotieavontuur aan te gaan. Je hebt die verwachting meer dan waar gemaakt. Ik kreeg vanaf het begin je vertrouwen en je liet me mijn eigen pad volgen. Je nam me serieus en vroeg me soms zelfs om advies. Je vertelde open over je eigen twijfels en beslommeringen en at ijsjes met mij en Pim op de trappen van het SLAZ. Ik heb veel opgestoken van onze gesprekken en wil je danken voor je (schijnbaar) onuitputtelijke energie en input. Ik hoop nog vele avonturen met je te beleven.

Ook Cees en Albert voldeden niet aan mijn standaard indruk van promotoren. Ik weet nog goed dat ik na het rondmailen van mijn eerste artikel binnen een dag feedback kreeg, inclusief enkele complimenten!? Dat was ik in de medisch setting toch minder gewend. Cees, ik heb veel bewondering voor je wetenschappelijke expertise, je oprechte interesse in anderen en je laagdrempeligheid, een zeldzame combinatie. Albert, ik waardeer je betrokkenheid (telefoon na vrijwel elke bijeenkomst, "was het naar je zin?"), neus voor interessante onderzoeksvragen en je pragmatische aanpak. Ik kijk er naar uit ook als clinicus samen met jullie te blijven nadenken over écht goed medisch onderwijs en misschien nog een keertje een weekend te gaan zeilen (en dat ik deze keer dan niet (letterlijk) in het water val..).

Lieve Pim, van onderzoeksmaatje ben je verworden tot een goede vriend. Uiteraard ben je mijn paranimf. Wat ben ik vaak bij je binnengevallen om leuke en minder leuke zaken met je te delen. Het was een voorrecht met je samen te werken want je bent niet alleen slim maar ook erg geestig. Onze eindeloze gesprekken hebben me onder meer wetenschappelijke inzichten, geweldige presentaties, de slappe lach en essentiële klusweetjes opgeleverd. Van vrijwel elk congres wisten we, soms samen met Jessica en Niels, een echt feestje te maken. Ik denk met lichte weemoed terug aan onze tijd op de 7e verdieping in het SLAZ, maar ik kijk er nu al naar uit straks als aios met je samen te werken.

Tevens wil ik de leden van mijn leescommissie, prof. dr. H.A.M. Brölmann, prof. dr. D.D.M. Braat, prof dr. R.J. Stolker, prof. dr. T.L. Dornan en dr. M.J.M.H. Lombarts hartelijk danken voor de zorgvuldige beoordeling van mijn manuscript.

Tim, I would like to thank you particularly for your never-ending enthusiasm and support and for being a member of my reading committee. I've come to know you and Ceri as two very friendly and special people with a great sense of humour and hospitality. It was inspiring to discuss work related issues, but other issues as well and I hope to stay in touch.

Beste Hans van der Schoot, samen met Fedde heb je me overtuigd van de kansen die in dit onderzoek verborgen lagen. Je heb je me altijd gesteund bij het uitvoeren van mijn onderzoek en toonde interesse hoe het met me ging. Ik vind het geweldig dat het SLAZ het experiment met mij als onderwijscoördinator/onderzoeker is aangegaan en ik heb het gevoel dat jij daar een belangrijke rol in hebt gespeeld. Dank hiervoor.

Mijn directe collegae van het opleidingscentrum en met name Susanne, Anke Nel en Joyce wil ik danken voor de gezelligheid, wederzijdse betrokkenheid en de goede inhoudelijke discussies. Ik heb veel van jullie geleerd.

Michiel, jij hebt als 'de nieuwe Klarke' mijn werk overgenomen en bent daadkrachtig aan de slag gegaan met zowel je onderzoek als onderwijs gerelateerde projecten. We hebben kort samengewerkt maar je enthousiasme en steun als ik (hoogzwanger) enigszins puffend de gang op slofte staan me nog helder voor de geest. Dank je wel.

De 'meisjes van de bieb' (Marjan, Chantal, Marianne): dank voor de vele artikelen en boeken die jullie voor me hebben opgevraagd en de kopjes koffie.

Daarnaast wil ik de Journal Club bedanken en dan in het bijzonder Debbie: we hebben elkaar aan het begin van ons onderzoek ontmoet bij de cursus van Albert en Arno en contact gehouden. Het is altijd erg gezellig je te spreken en ik vind je, als medemoeder met ambitie, inspirerend. Dank dat je mijn paranimf wilt zijn.

Rachel, thank you for being such a wonderful research colleague. It was great to work together with you on 'our' chapter on the learning environment (and check out the coffee bars in Amsterdam and shopping venues in Manchester).

Thomer, dank voor je hulp bij het Engels van mijn 2 eerste hoofdstukken. Mereke, jij hebt de rest van het proefschrift 'ge-edit': dank voor je opbouwende commentaar.

Clasien van der Houwen, aan het begin van mijn onderzoekstijd hebben we ons samen verdiept in het doen van onderzoek met behulp van focusgroepen. De rijke opbrengst van kwalitatief onderzoek heeft me geïnspireerd deze insteek ook later in mijn onderzoek te gebruiken. Dank voor de kennismaking met deze methodiek en met jou als persoon.

Klaas Sijtsma, jij hebt gepoogd me het een en ander uit te leggen over de wondere wereld der statistiek. Dank voor je geduld.

Jonas van de Lande, jij verzamelde deels de data onder aios gynaecologie over hun beeld van de ideale opleider en dacht mee over het uiteindelijke artikel, dank voor je inbreng.

Elvira den Breejen, je hebt verschillende interviews met andere co-assistenten gehouden en geholpen met de analyse van de (enorme berg) verzamelde data, dank daarvoor.

Scheltus van Luijk, dank voor je zorgvuldige inbreng bij het tot stand komen van de D-RECT, je aanwezigheid bij de overleggen op het Onderwijsinstituut en natuurlijk je uitstekende tips op het gebied van theatersport voorstellingen die ik moest gaan zien.

Daarnaast wil ik Ed van der Veen en Jian-Lie Ku van het CBOG en Marco Varkevisser van het ECRi danken voor de logistieke ondersteuning van het D-RECT project.

Alle co-assistenten, a(n)ios, opleiders, onderwijskundigen en anderen die hebben meegewerkt aan mijn onderzoeken wil ik hartelijk danken voor hun tijd en energie bij het invullen van vragenlijsten of deelnemen aan interviews.

Lieve Sanne, Mark, Sebastiaan, Marjolein, Wendy, Xander, Jansje en Gert Jan: dank voor jullie luisterend oor en immer opbeurende woorden. De fijne kopjes koffie, etentjes, mailtjes en uitjes/vakanties mogen wat mij betreft tot in de eeuwigheid doorgaan.

Lieve Ellen, Leen en Yarah: wat bof ik toch met zo'n geweldige schoonfamilie! Dank voor jullie steun en interesse en het ook in praktische zin bijspringen (oppassen..) zodat ik nu dit dankwoord kan schrijven.

Lieve Abel & Eelco en Margriet & Marco, het was geweldig dat jullie er altijd waren om een geaccepteerd artikel te vieren of me even op te peppen als ik me onzeker of gedemotiveerd voelde. Ik heb het gevoel dat ik altijd bij jullie terecht kan en dat is een groot goed.

Lieve papa en mama. Jullie zijn de beste ouders die ik me maar kan wensen: ik dank jullie voor het optimisme, daadkrachtig optreden ('en NU aan het werk') en het vertrouwen in mijn kunnen. Ik groei er nog steeds door. Ik ben heel blij met jullie!

Lieve Japie, een deel van dit proefschrift schreef ik toen je vrolijk tegen mijn ribben trapte en later toen je kraaiend in je bedje lag. Je maakt mijn leven elke dag een stukje mooier.

Liefste Niels. Je bent het. *In your arms I feel, sunshine.*

Curriculum Vitae

CURRICULUM VITAE

Klarke Boor was born on the 1st of January 1977 in Kirinda, Rwanda. When she was three years old she moved back with her parents and two sisters to the Netherlands and grew up in Haarlem. After graduating high school (Stedelijk Gymnasium, Haarlem) in 1996, she commenced her Medicine study at the VU University in Amsterdam.

During her study she trained younger peers in courses on communication skills and medical ethics and worked as a volunteer at the Kids Help Phone. Her clerkships sparked her interest in the field of Obstetrics and Gynaecology. She graduated cum laude as a medical doctor in 2003 and worked as an obstetrics-gynaecology resident in the St. Lucas Andreas Hospital in Amsterdam (prof. dr. F. Scheele). In November 2004 she started her research project on the clinical learning climate described in this thesis. Klarke combined this with the coordination of all clerkships in the St. Lucas Andreas Hospital and the development of a monitoring system assessing clerkship quality.
Since April 2009 she works as a specialist trainee in obstetrics and gynaecology at the VU University Medical Center in Amsterdam (prof. dr. H.A.M. Brölmann). Occasionally, she gives 'teach the teacher' courses to consultants and residents.

In her spare time she loves to play impov theatre, walk in the mountains or read novels. She lives in Haarlem with Niels and is the mother of a baby boy, Jaap.