University of Groningen

Assessment of professional behaviour in undergraduate medical education

Schonrock-Adema, Johanna; Heijne - Penninga, Marjolein; van Duijn, Marijtje A J; Geertsma, Jelle; Cohen - Schotanus, Janke

Published in:
Medical Education

DOI:
[10.1111/j.1365-2923.2007.02817.x](https://doi.org/10.1111/j.1365-2923.2007.02817.x)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2007

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Schonrock-Adema, J., Heijne - Penninga, M., van Duijn, M. A. J., Geertsma, J., & Cohen - Schotanus, J. (2007). Assessment of professional behaviour in undergraduate medical education: Peer assessment enhances performance. *Medical Education*, 41(9), 836-842. <https://doi.org/10.1111/j.1365-2923.2007.02817.x>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Assessment of professional behaviour in undergraduate medical education: peer assessment enhances performance

JOHANNA SCHÖNROCK-ADEMA,¹ MARJOLEIN HEIJNE-PENNINGA,¹ MARIJTJE A J VAN DUIN,² JELLE GEERTSMA¹ & JANKE COHEN-SCHOTANUS¹

OBJECTIVES To examine whether peer assessment can enhance scores on professional behaviour, with the expectation that students who assess peers score more highly on professional behaviour than students who do not assess peers.

METHODS Undergraduate medical students in their first and second trimesters were randomly assigned to conditions with or without peer assessment. Of the total group of 336 students, 278 students participated in the first trimester, distributed over 31 tutorial groups, 17 of which assessed peers. The second trimester involved 272 students distributed over 32 groups, 15 of which assessed peers. Professional behaviour was rated by tutors on 3 dimensions: Task Performance; Aspects of Communication, and Personal Performance. The rating scale ranged from 1 (poor) to 10 (excellent). Data were analysed using multivariate repeated measures multilevel analysis.

RESULTS Assessment scores were found to have generally increased in the second trimester, especially the personal performance scores of students who assessed peers. In addition, female students were found to have significantly higher scores than male students.

CONCLUSIONS In undergraduate medical education, peer assessment has a positive influence on

professional behaviour. However, the results imply that peer assessment is only effective after students have become adjusted to the complex learning environment.

KEYWORDS humans; male; female; randomised controlled trial [publication type]; *education, medical, undergraduate; professional practice/*standards; peer group; clinical competence/standards.

Medical Education 2007; **41**: 836–842
doi:10.1111/j.1365-2923.2007.02817.x

INTRODUCTION

Adequate professional behaviour has come to be recognised as an essential part of a medical doctor's competence.^{1–3} The acknowledgement of the importance of professional behaviour has fuelled the need to incorporate this subject into undergraduate medical education.⁴ As the literature shows that nothing regulates the learning process as much as the examination programme does,^{5–7} it is recommended that professional behaviour be assessed.

Behaviour assessment has an educational goal: to focus students' attention on their own behaviour and to stimulate them to optimise their professional behaviour. Some studies outside the context of medical education show that students show more improvement when they are asked to assess themselves or their peers.^{8–10} Possibly, the increased awareness of criteria that results from the process of assessing peers deepens elaboration,¹¹ which, in turn, improves learning outcomes.^{12,13}

The current study adds to previous research because of its longitudinal experimental design. In addition,

¹Center for Research and Innovation of Medical Education, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands

²Department of Sociology/Statistics and Measurement Theory, Heijmans Institute/ICS, University of Groningen, Groningen, The Netherlands

Correspondence: Johanna Schönrock-Adema, A Deusinglaan 1, 9713 AV Groningen, The Netherlands. Tel: 00 31 50 363 83 78; Fax: 00 31 50 363 73 90; E-mail: j.schonrock-adema@med.umcg.nl

Overview

What is already known on this subject

Adequate professional behaviour is an essential part of a medical doctor's competence and should therefore be assessed.

Outside the context of medical education, the process of assessing peers has been shown to enhance learning.

What this study adds

In this study, performed in undergraduate medical education, peer assessment turned out to be effective only after a period of adjustment to the complex learning environment. Furthermore, females systematically scored higher than males.

Suggestions for further research

Future research should focus on the effects of peer assessment in later phases of medical education and on the replicability of the current results in situations with summative assessments.

the present study was conducted within medical education, focusing on the effect of peer assessment on professional behaviour scores awarded by tutors. It was expected that students in a condition with peer assessment would obtain higher scores on professional behaviour than students in a condition without peer assessment.

METHODS

Respondents and procedure

The undergraduate medical curriculum at the University of Groningen covers 6 years, each of which is divided into 3 trimesters. In the first 10 trimesters of the study programme, tutorial groups of approximately 10 students meet twice a week to discuss medical problems. This study focused on professional behaviour in the first 2 trimesters of the programme, as rated by the group tutors. Participation in the study was voluntary, but encouraged: it was explained that the study would be important to the medical

school because professional behaviour and assessment thereof would be part of a new curriculum under development. In total, 336 students participated in the first 2 trimesters, approximately two thirds of whom were female. To eliminate any effects of previous learning or experiences, only novice medical students were included in the study. Of the remaining 315 students, data for 278 students in the first trimester (88%) and 272 students in the second trimester (86%) were available.

For each trimester, students were assigned randomly to groups and groups randomly assigned to a condition. It was logistically impossible to balance the groups exactly according to condition. The students in the first trimester were distributed over 31 tutorial groups, with 14 groups assigned to the condition without peer assessment (condition 0) and 17 to the condition with peer assessment (condition 1). The students in the second trimester were distributed over 32 tutorial groups, with 17 groups assigned to condition 0 and 15 to condition 1. A pool of 48 tutors led the 63 tutorial groups, with a maximum of 2 groups per tutor. In order to enhance the reliability of the assessment, tutors received both written and verbal instructions. In addition, the instructors encouraged the tutors to consult them in case of difficulties. The tutors were given supporting forms on which to record their observations during the trimester.

The peer assessment condition implied that each student assessed 3 peers in his or her tutorial group, namely, the 3 students who came next in the randomly established group order. Students were encouraged to observe their peers' behaviours from the beginning of the trimester in order to facilitate their final decisions on professional behaviour scores. They also received supporting forms on which to record their observations. In both conditions, students were given information about the key concepts and criteria used in the 3 areas of assessment before the trimester started, so that the pure influence of peer assessment could be investigated (in condition 1). Although the tutors who rated the students on professional behaviour were not blinded to the treatment condition, they were not aware of the hypotheses of the study. The data were processed confidentially and students received feedback on their professional behaviour afterwards.

Instrument

The instrument was developed according to the recommendations of the Project Team Consilium Abeundi, which has formulated guidelines on the

assessment of professional behaviour for medical schools across the Netherlands.⁴ Professional behaviour was assessed on 3 dimensions based on theoretical notions regarding professional behaviour: Task Performance; Aspects of Communication, and Personal Performance.^{3,4} Illustrative characteristics of the dimensions were listed on the assessment form and more detailed descriptions of the characteristics were provided on the reverse side of the form. Some examples of characteristics pertaining to the dimensions are: 'preparation' and 'taking responsibility' for Task Performance; 'active listening' and 'co-operation' for Aspects of Communication, and 'critical reflection on personal performance' and 'personal growth' for Personal Performance. Professional behaviour was rated on a scale of 1 (poor) to 10 (excellent). This type of marking is commonly used at all levels of education in the Netherlands. A score < 5.5 is considered unsatisfactory. In addition to the marks, the tutors were invited to provide positive feedback and constructive criticism to help students improve their behaviour in the future.

Data analysis

Full data would have comprised 3 assessments per student per tutorial group per trimester. Observations were missing for 80 students who participated in only 1 tutorial group in both trimesters. Missing observations amounted to 5% and approximately 13.5% of observations in trimesters 1 and 2, respectively. These data called for a multivariate multilevel analysis.

Multivariate analyses were performed in order to control for dependencies between individual measures.¹⁴ The dependent variable comprised the 3 scores on Task Performance, Aspects of Communication and Personal Performance.

A multilevel dependence structure was formed by both the tutorial groups and the 2 trimester measurements. Because of the way the curriculum was organised, with a change in tutorial groups each trimester, it was impossible to distinguish changing between-group variances from within-subject (or between-trimester) variance. After establishing that between-group variances were low in comparison with within-subject variance, a longitudinal model was used, ignoring the group structure. The data were analysed with the multilevel computer programme MLwiN (Version 2.02).

In a first step, the variance structure was established in a simple model that allowed for different trimester

means but did not use additional covariate information. This revealed considerable correlation between the observations at and between each trimester, with somewhat decreasing variance over time. Next, we investigated which factors significantly influenced the scores on professional behaviour in a stepwise backward analysis. The largest model contained the variables 'student gender' and 'condition' and the interaction between gender and condition, plus all interactions with trimester, as well as an indicator variable of previous experience with peer assessment in the second trimester. Whether a new model implied a deterioration from the previous larger model was determined on the basis of the difference in deviance. The deviance represents a measure of discrepancy between model and data. The significance of an increase in deviance is tested using a chi-square test, with the number of degrees of freedom equal to the number of left-out model parameters.

The final model was obtained after removing insignificant interaction effects and main effects not included in a significant interaction effect. The significance of the contribution of each separate independent variable was determined by dividing its coefficient by its standard error, resulting in a *z*-value, given the sufficiently large student sample size.¹⁴

RESULTS

An overview of the marks granted by tutors is presented in Table 1. Task Performance yielded the highest scores and Personal Performance the lowest. The scores in the second trimester were higher than those in trimester 1. In the first trimester the scores were approximately the same for both conditions. In the second trimester, the scores in the condition with peer assessment were slightly higher than those in the condition without peer assessment.

The results of the multivariate multilevel analyses with the tutor scores on Task Performance, Aspects of Communication and Personal Performance as the dependent variables showed an effect of time (Table 2). Scores on Task Performance, in particular, and Personal Performance, to some extent, exhibited a learning effect ($z = 3.34$, $P < 0.001$; $z = 1.69$, $P < 0.05$, respectively), although no significant effect emerged for Aspects of Communication. Scores on Task Performance, Aspects of Communication and Personal Performance were different for males and females ($z = 3.34$, $P < 0.001$; $z = 2.83$, $P < 0.01$, and $z = 3.18$, $P < 0.001$, respectively), with female students systematically scoring higher than male

Table 1 Overview of tutor marks: means, standard deviations, minimum and maximum scores by condition* and for the total sample

	Trimester 1					Trimester 2				
	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>n</i>	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Task Performance										
Condition 0	124	7.46	0.87	5.00	9.50	149	7.73	0.70	5.00	9.00
Condition 1	140	7.43	0.68	4.50	9.00	90	7.93	0.73	5.50	9.00
Total	264	7.44	0.78	4.50	9.50	239	7.81	0.71	5.00	9.00
Aspects of Communication										
Condition 0	124	7.42	0.89	5.00	9.50	148	7.52	0.73	5.00	9.00
Condition 1	140	7.35	0.71	5.50	9.00	91	7.71	0.70	6.00	9.00
Total	264	7.38	0.80	5.00	9.50	239	7.59	0.72	5.00	9.00
Personal Performance										
Condition 0	124	7.32	0.76	5.00	9.00	149	7.44	0.61	5.00	9.00
Condition 1	140	7.23	0.54	6.00	8.50	91	7.73	0.56	6.50	9.00
Total	264	7.27	0.65	5.00	9.00	240	7.55	0.61	5.00	9.00

* Condition 0 = tutor assessment; condition 1 = peer and tutor assessment

SD = standard deviation

Table 2 Final model of multivariate multilevel analyses with scores on Task Performance, Aspects of Communication and Personal Performance as dependent variables

Fixed effects	Task Performance		Aspects of Communication		Personal Performance	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
Mean trimester 1	7.28*	0.087	7.27*	0.088	7.18*	0.071
Improvement trimester 2	0.29*	0.086	0.10	0.090	0.13‡	0.074
Gender	0.26*	0.079	0.22†	0.078	0.20*	0.062
Condition	-0.04	0.093	-0.08	0.097	-0.09	0.079
Condition*trimester 2	0.18	0.138	0.25‡	0.140	0.35*	0.113
Random effects	(Co)Variance	SE	(Co)Variance	SE	(Co)Variance	SE
Between-respondents covariance matrix						
Task Performance	0.175	0.038				
Aspects of Communication	0.142	0.033	0.115	0.040		
Personal Performance	0.078	0.026	0.043	0.027	0.034	0.027
Within-respondents covariance matrix						
Trimester 1/2						
Task Performance	0.43/0.30	0.052/0.045				
Aspects of Communication	0.30/0.18	0.046/0.040	0.53/0.39	0.059/0.052		
Personal Performance	0.19/0.17	0.037/0.034	0.26/0.22	0.041/0.037	0.39/0.31	0.043/0.038
Deviance	2582.1					

* $P < 0.001$; † $P < 0.01$; ‡ $P < 0.05$ (1-sided)

SE = standard error

students. The effect of gender and condition and their interaction over time is illustrated in Figs 1–3 for the 3 aspects of assessment. In addition, a multivariate interaction effect of time and condition was found: in the second trimester, students in condition 1 (with peer assessment) scored more highly than students in condition 0. This effect was strongest for Personal Performance ($z = 3.12$, $P < 0.001$), and of medium strength for Aspects of Communication, but not significant for Task Performance. The analyses did not show an effect of experience caused by previous assignment to condition. The variances and covariances shown in the

Figure 1 Overview of scores on Task Performance for males and females for both trimesters and conditions

Figure 2 Overview of scores on Aspects of Communication for males and females for both trimesters and conditions

Figure 3 Overview of scores on Personal Performance for males and females for both trimesters and conditions

lower half of Table 2 illustrate the individual variability (somewhat declining over time) and the strong relationship between scores on the 3 dependent variables over trimesters (within-individual) and within each trimester.

DISCUSSION

The hypothesis that students in a condition with peer assessment score more highly on professional behaviour than students in a condition without peer assessment was only partially supported. Only in the second trimester did students assessing peers score more highly. A possible explanation for the lack of confirmation of the hypothesis in the first trimester bears on the fact that students were at the very start of their medical training at the time and had to acquire several skills related to problem-based learning (PBL). To make learning a constructive, self-directed, collaborative and contextual process,¹⁵ students need to actively plan, monitor and evaluate their learning process and collaborate with peers, among other things. Possibly, the transition to such a complex learning environment may have caused a

cognitive overload, leaving students little mental space to examine their own professional behaviour (beyond focusing on the content of learning). That the scores of students who assessed peers in the first trimester were lower rather than higher (although the difference was far from significant) might be interpreted as support for this explanation. Obviously, advantages such as increased awareness of criteria leading to enhanced learning outcomes^{11–13} do not apply in such cases. Confirmation of the hypothesis in the second trimester could be interpreted as support for the idea that students need time to adjust to the complex learning environment. If this train of thought is accurate, then future research should demonstrate a positive influence of peer assessment on professional behaviour in later phases of the curriculum.

Closer inspection of the results shows a learning effect over time for Task Performance, a condition effect for Aspects of Communication in the second trimester and a time effect as well as a condition effect in the second trimester for Personal Performance. Scores for Task Performance seem to have improved regardless of condition. Possibly the assessment criteria for this dimension are so well defined and clear that peer assessment does not further enhance awareness and therefore does not bring about additional improvement. The effect of condition in the second trimester for Aspects of Communication supports the idea that active use of the criteria is essential for improving the behaviour aspects pertaining to this dimension. Enhancement for Task Performance or for Aspects of Communication also implies higher scores for Personal Performance, for this dimension also includes personal growth. Therefore, the effects found for Personal Performance follow logically from the effects for the other dimensions. The preceding discussion suggests that peer assessment has greater effect in situations where the assessment criteria are less clear.

Although on average the differences were only small, peer assessment seems particularly important for students scoring close to the cut-off point. Given that a score <5.5 is considered unsatisfactory, the results show that none of the students in condition 1 performed unsatisfactorily in the second trimester, which contrasts with findings for condition 0. This is particularly important given the seriousness of reported unprofessional behaviour in medical education.⁴ The minimum scores of students in condition 1 were 0.5–1.5 points higher than those of students in condition 0. Although a difference of 0.5 points on a 10-point scale may not seem very large,

differences of this size are already relevant. As the investments required to reach this effect are only small (just asking students to assess some of their peers), it is recommended that peer assessment be introduced.

A first possible limitation of this study is a confounding of tutor rating stringency with treatment. However, as the majority of the tutors had only 1 group, the small between-group variances also indicate that tutors did not vary much in rating stringency. Secondly, a confounding effect is not plausible because, if such an effect had been present, the scores of tutors in 1 of the conditions, for example in the condition with peer assessment, would have been higher in both trimesters and not only in the second.

A second limitation of this study is that the assessment was formative. This may have led students, and possibly tutors as well, to take the assessment less seriously and thus make less effort to ensure accuracy than they would have if the assessment had been summative. However, as far as the criterion scores (tutor scores) are concerned, this problem yielded both conditions. As far as student dedication is concerned, this limitation may have made the effects of peer assessment weaker than they would have been in cases of summative assessment. Future research might examine whether the present results are replicable in situations with summative assessments. Research has repeatedly shown that judgements with far-reaching consequences for the assessee are milder than judgements that are not intended for summative decisions.^{16,17} In addition, it should be noted that the generalisability of the results of this study to other educational situations is limited as the tutorial setting is structured with clear goals and tasks.

No cumulative effects of condition were found in the current study, implying that assessing peers in the first trimester did not enhance performance in the second. However, it might be that repeatedly assessing peers results in better professional behaviour in later phases of the study, when students have already become accustomed to the learning environment. This effect might be increased by another positive effect of peer assessment that is well known from multi-source rating: deploying several raters also provides ratees with feedback from different perspectives, which can provide unique and meaningful information about performance and may therefore help to increase the acceptance and usefulness of the feedback.^{18,19} Therefore, it would be interesting to

investigate whether an accumulative effect of peer assessment on professional behaviour exists in later phases of medical education.

Main effects of student gender were found, with females scoring systematically higher than males. Females have repeatedly been shown to gain higher marks than males.²⁰ This may be because females tend to work in a more disciplined and conscientious way than males.^{21,22} Discipline and conscientiousness have consistently been demonstrated to be related to achievement and performance.^{23,24} The fact that these qualities are reflected in the assessment criteria for assessing professional behaviour may have magnified the differences in scores between males and females: dealing with tasks and people in a disciplined and conscientious manner may dovetail very well with the female approach.

Two interesting questions arose as a result of the gender effects found in this study. Does gender composition within groups influence the results of individual students? Does tutor gender interact with student gender and group composition? These questions may be relevant as group work is an important ingredient in problem-based learning and deserves further research.

In conclusion, peer assessment enhanced scores for professional behaviour in the second trimester. Although on average the differences were only small, peer assessment seems particularly important for students who score close to the cut-off point and is therefore a promising method of preventing unprofessional behaviour. In addition, the costs involved in peer assessment are relatively low. In view of the discussion on this finding and because of the additional positive effects of peer assessment mentioned earlier in this paper, it is recommended that peer assessment be introduced. However, it seems wise to restrict implementation to later phases of the curriculum in order to give students some time to become accustomed to the learning environment. Future research should shed light on any differences between formative and summative assessments with respect to the influence of peer assessment and on the replicability of the findings in other educational settings.

Contributors: JS-A was principally responsible for conceptualising and conducting the study, and writing the manuscript. MH-P contributed to the study design, acquisition of data and data entry. MAJvD conducted the

multilevel analysis. JG contributed to the data analysis. JC-S supervised the study. All authors contributed to the interpretation of results, commented on successive drafts and approved the final manuscript.

Acknowledgements: the authors thank the students who participated in the study, and J Bouwkamp-Timmer, for her constructive comments on the manuscript.

Funding: none.

Conflicts of interest: none.

Ethical approval: not required.

REFERENCES

- 1 CanMEDS (Canadian Medical Education Directions for Specialists). CanMEDS 2000: extract from the CanMEDS 2000 Project Societal Needs Working Group Report. *Med Teach* 2002;**22**:549–54.
- 2 Crossley J, Humphris G, Jolly B. Assessing health professionals. *Med Educ* 2002;**36**:800–4.
- 3 Van Luijk SJ, Smeets JGE, Smits J, Wolfhagen I, Perquin MLF. Assessing professional behaviour and the role of academic advice at the Maastricht Medical School. *Med Teach* 2000;**22**:168–72.
- 4 Van Luijk SJ, ed. *Professional Behaviour: Teaching, Assessing and Coaching Students*. Final Report of the Project Team Consilium Abeundi, appointed by the Netherlands Council of Medical Deans of the Dutch Federation of University Medical Centres. Maastricht: University Press Maastricht 2005.
- 5 Cohen-Schotanus J. Student assessment and examination rules. *Med Teach* 1999;**21**:318–21.
- 6 Eraut M. A wider perspective on assessment. *Med Educ* 2004;**34**:803–4.
- 7 Frederiksen N. The real test bias: influences of testing on teaching and learning. *Am Psychol* 1984;**39**:193–202.
- 8 Dominick P, Reilly R, McGourty J. The effects of peer feedback on team member behaviour. *Group Organization Manage* 1997;**22**:508–20.
- 9 Sluijsmans DMA, Brand-Gruwel S, Van Merriënboer JJG. Peer assessment training in teacher education: effects on performance and perceptions. *Assessment Eval Higher Educ* 2002;**27**:443–54.
- 10 Topping K. Peer assessment between students in colleges and universities. *Rev Educational Res* 1998;**68**:249–76.
- 11 Somervell H. Issues in assessment, enterprise and higher education: the case for self, peer and collaborative assessment. *Assessment Eval Higher Educ* 1993;**18**:221.
- 12 Bruch M, Pearl L, Giordano S. Differences in the cognitive processes of academically successful and unsuccessful test-anxious students. *J Counsel Psychol* 1986;**33**:217–9.
- 13 Svensson L. On qualitative differences in learning. III. Study skill and learning. *Br J Educ Psychol* 1977;**47**:233–43.
- 14 Snijders TAB, Bosker RJ. *Multilevel Analysis: an Introduction to Basic and Advanced Multilevel Modelling*. London: Sage 1999.
- 15 Dolmans DHJM, De Grave W, Wolfhagen IHAP, van der Vleuten CPM. Problem-based learning: future challenges for educational practice and research. *Med Educ* 2005;**39**:732–41.
- 16 Decker P, Cornelius E. The effect on leniency of justifying performance ratings to a supervisor. *J Psychol* 1981;**108**:211–8.
- 17 Harris M, Smith D, Champagne D. A field study of performance appraisal purpose: research- versus administrative-based ratings. *Personnel Psychol* 1995;**48**:151–60.
- 18 Harris MM, Schaubroeck J. A meta-analysis of self-supervisor, self-peer and peer-supervisor ratings. *Personnel Psychol* 1988;**41**:43–62.
- 19 London M, Smither JW. Can multi-source feedback change perceptions of goal accomplishment, self-evaluations and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychol* 1995;**48**:803–39.
- 20 Duckworth AL, Seligman MEP. Self-discipline gives girls the edge: gender in self-discipline, grades and achievement test scores. *J Educational Psychol* 2006;**98**:198–208.
- 21 Grabill KM, Lasane TP, Povitsky WT *et al*. Gender and study behaviour: how social perception, social norm adherence, and structured academic behaviour are predicted by gender. *North Am J Psychol* 2005;**7**:7–24.
- 22 Rubinstein G. The big five among male and female students of different faculties. *Pers Individ Diff* 2005;**38**:1495–503.
- 23 Busato VV, Prins FJ, Elshout JJ, Hamaker C. Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Pers Individ Diff* 2000;**29**:1057–68.
- 24 Furnham A, Moutafi J, Chamorro-Premuzic T. Personality and intelligence: gender, the big five, self-estimated and psychometric intelligence. *Int J Select Assessment* 2005;**13**:11–24.

Received 24 November 2006; editorial comments to authors 12 February 2007; accepted for publication 3 April 2007