On the role of biomedical knowledge in the acquisition of clinical knowledge

Stefan K Schauber,[1,2] Martin Hecht,[3] Zineb M Nouns[2] & Susanne Dettmer[1]

CONTEXT Basic science teaching in undergraduate medical education faces several challenges. One prominent discussion is focused on the relevance of biomedical knowledge to the development and integration of clinical knowledge. Although the value of basic science knowledge is generally emphasised, theoretical positions on the relative role of this knowledge and the optimal approach to its instruction differ. The present paper addresses whether and to what extent biomedical knowledge is related to the development of clinical knowledge.

METHODS We analysed repeated-measures data for performances on basic science and clinical knowledge assessments. A sample of 598 medical students on a traditional curriculum participated in the study. The entire study covered a developmental phase of 2 years of medical education. Structural equation modelling was used to analyse the temporal relationship between biomedical knowledge and the acquisition of clinical knowledge.

RESULTS At the point at which formal basic science education ends and clinical training begins, students show the highest levels of biomedical knowledge. The present data suggest a decline in basic science knowledge that is complemented by a growth in clinical knowledge. Statistical comparison of several structural equation models revealed that the model to best explain the data specified unidirectional relationships between earlier states of biomedical knowledge and subsequent changes in clinical knowledge. However, the parameter estimates indicate that this association is negative.

DISCUSSION Our analysis suggests a negative relationship between earlier levels of basic science knowledge and subsequent gains in clinical knowledge. We discuss the limitations of the present study, such as the educational context in which it was conducted and its non-experimental nature. Although the present results do not necessarily contradict the relevance of basic sciences, we speculate on mechanisms that might be related to our findings. We conclude that our results hint at possibly critical issues in basic science education that have been rarely addressed thus far.

Medical Education 2013: *47*: 1223–1235
doi: 10.1111/medu.12229

Discuss ideas arising from the article at www.mededuc.com 'discuss'

[1]Institute of Medical Sociology, Charité – Universitätsmedizin Berlin, Berlin, Germany
[2]Dieter Scheffner Center for Medical Teaching and Educational Research, Charité – Universitätsmedizin Berlin, Berlin, Germany
[3]Institute for Educational Quality Improvement, Humboldt–Universität zu Berlin, Berlin, Germany

Correspondence: Stefan K Schauber, Institute of Medical Sociology, Charité Universitätsmedizin Berlin, Luisenstrasse 57, Berlin 10117, Germany. Tel: 00 49 30 450 576 037;
E-mail: stefan.schauber@charite.de

INTRODUCTION

Abraham Flexner's report on medical education in the USA and Canada,[1] published more than 100 years ago, had an indisputable influence on the way medical curricula were conceived. Flexner described how the biomedical sciences – or what he referred to as 'the laboratory branches' of medical education[1] – were separated from the practical, clinical aspects. He emphasised the fundamental and integral role of the biomedical sciences and argued that the division was made 'for purposes of convenience […] as the work is carried on mainly in laboratories or mainly in the hospital'.[1] In other words, according to Flexner, '…the distinction is only superficial'.[1] Flexner argued that basic sciences and clinical practice are tightly connected, that teaching must be conducted with the critical method of science in mind, and that 'undergraduate instruction [should] be throughout explicitly conscious of its professional end and aim'.[1] He saw medical education as being in 'close association with contiguous, contributory, or overlapping sciences'[1] and as providing an interdisciplinary learning environment in which 'at any moment a lucky stroke may transfer a problem from pathology to chemistry or biology'.[1]

Indeed, medical education has undergone pronounced changes since the early 20th century. One major development in recent decades has concerned the integration of content from the neighbouring fields of the behavioural or social sciences into the medical curriculum.[2,3] Overall, the medical curriculum has widened considerably since Flexner's report, resulting in a truly multidisciplinary perspective on the knowledge, skills and competencies to be acquired by future doctors.[4]

Needless to say, these developments have not been smooth. The reform of medical education toward a competence-based approach has been accompanied by a shift away from the basic sciences.[5–8] In Scotland, for example, the number of teaching hours for anatomy has decreased by 50% in recent decades.[9] Other authors have reported similar findings in other regions.[6,10] Indeed, it seems that the relevance of detailed basic science knowledge has increasingly been called into question in recent decades.[6,7,10,11] This debate on the relevance of biomedical knowledge follows several related discussions on, for example, the deficits of medical school graduates' basic science knowledge (for a review, see Bergman *et al.*[5]), the role of dissection in the undergraduate curriculum[12,13] and how basic science knowledge can be transferred to clinical practice.[14,15]

At first, this challenge to the role of basic sciences in undergraduate medical training seems to contrast with the general belief that it is important for doctors to have elaborate biomedical knowledge.[7,16] However, research has found that expert doctors rarely use biomedical knowledge in daily diagnosis. Rather, non-analytic reasoning or pattern recognition is used to process the vast majority of routine cases.[17,18] It seems that these non-analytic strategies may be the result of direct patient contact and clinical experience, rather than the outcome of factual knowledge of, for example, anatomy, biochemistry and physiology.[17] Against this background, why do doctors need basic science knowledge at all?

Relating expert performance to biomedical knowledge

Theories on the role of clinical reasoning in medical expertise address the relationships among biomedical knowledge, clinical knowledge and clinical reasoning and usually highlight the importance of basic sciences in the medical curriculum. Two of the most influential approaches are discussed in the following paragraphs.

According to Kaufman *et al.*[7] and Patel *et al.*,[19] clinical knowledge and biomedical concepts are stored separately. In general, experts do not need to access biomedical knowledge; rather, they draw on their specific knowledge of a particular clinical presentation. However, biomedical knowledge becomes essential when pattern recognition no longer applies, when cases are particularly difficult or when unexpected scenarios occur. Patel and colleagues[19] assume that biomedical knowledge is characterised by more abstract causal relations. By contrast, clinical knowledge is assumed to be more 'concrete' or 'observational'[19] and to describe 'how a set of symptoms is consistent with a diagnosis'.[20] Therefore, the approach described by Patel *et al.*[19] and Kaufman *et al.*[1,7] is often referred to as the 'two worlds hypothesis'.

By contrast, the model proposed by Boshuizen, Schmidt, Norman and Rikers[21–24] suggests that the development of medical expertise follows a meaningful sequence of stages and that every stage involves structural changes in the knowledge base. In the first step, students build an elaborate base of biomedical knowledge. This biomedical knowledge is then increasingly subsumed under clinical

knowledge and hence encapsulated. An extensively connected network of biomedical and clinical knowledge provides the basis for the development of illness scripts. These scripts include knowledge of enabling conditions, malfunctions and signs and symptoms,[24] which allow for the rapid processing of clinical cases, similarly to pattern recognition.[23] According to Schmidt and Rikers,[24] biomedical and clinical knowledge are integrated into a single common knowledge base. Biomedical knowledge is assumed to be encapsulated under clinical knowledge and is thus thought to be implicitly activated in the handling of routine cases, by contrast with the 'two worlds' view.[19]

The theories of both Patel *et al.*[19] and Schmidt *et al.*[23] leave no doubt as to the relevance of basic science content in medical education and have stimulated a long tradition of research on the relationships among clinical reasoning, clinical knowledge and biomedical knowledge.[21,25–32] Critically, both conceptions have far-reaching implications for how undergraduate medical education should be structured. Kaufman *et al.* emphasise the 'special status'[7] of biomedical knowledge in medical education, in line with the traditional curriculum according to Flexner.[1] By contrast, the works of Boshuizen and Schmidt[21] imply the need for an integrated curriculum in which problem-based learning (PBL) is the central instructional approach.[22]

Although highly relevant to this discussion, studies applying theories of expertise to the daily practice of undergraduate medical education were surprisingly rare in the early years of research on the role of biomedical knowledge. Of course, various studies have related specific instructional approaches to cognitive benefits[33] or to differences in problem-solving styles.[34] However, Woods and colleagues were the first to systematically relate the unique value of biomedical knowledge to learning and instruction.[35–38] In a series of papers,[35–38] these authors conceived of basic science knowledge as a mnemonic device,[16] a tool enabling students to better learn and integrate clinical knowledge, or to make sense of signs and symptoms. Various studies found that the causal explanation of a disease in the form of its underlying pathological mechanisms represents an efficient tool that facilitates both memorising[36,39] and correct diagnosis.[35] Woods *et al.* concluded that biomedical knowledge plays an indirect role in the development of a novice to an expert: '[c]ausal understanding leads to more coherent understanding of clinical conditions, which in turn leads to expert-like behaviour.'[38] This conception of biomedical knowledge as a tool that creates coherence in the representation of clinical content[36] is directly related to the process of encapsulation proposed by Schmidt and Rikers.[24]

What this study adds

As noted by Woods,[16] most relevant research to date has been carried out in highly controlled laboratory settings. In addition, previous studies have generally drawn on rather small samples and have not allowed changes in student performances to be tracked across intervals longer than a few days or weeks. We are, of course, aware of the large body of research on the development of medical knowledge on a broader timescale, especially in the contexts of curricular comparisons[40–43] or the effectiveness of PBL.[44] In this tradition, however, biomedical and clinical knowledge have generally been treated as independent domains. Given the implications of the work of Schmidt and Rikers,[24] among others, which suggest that biomedical knowledge and clinical knowledge are closely related, this approach is questionable.

The purpose of the present study was therefore to model the temporal interaction between biomedical knowledge and changes in the development of clinical knowledge. To this end, we applied structural equation modelling (SEM) techniques to longitudinal progress test data. In accordance with the conception of basic science knowledge as a mnemonic device, we hypothesised that higher levels of biomedical knowledge at the beginning of the clinical phase are associated with higher gains in clinical knowledge in the subsequent academic year.

METHODS

Participants

Data for a total of 598 undergraduate medical students participating in an educational research study at Charité Medical University Berlin were included in the analysis. In order to cover a developmental phase of 2 years of medical education, we collected data in an accelerated longitudinal[45] or sequential cohort[46] design in three consecutive waves in, respectively, October 2010, April 2011 and October 2011. This approach combines longitudinal and cross-sectional data in order to cover a broad span of development sufficiently. Although participants provided longitudinal information on their development, their data do not necessarily span the 2 years

of interest. The sampling structure is presented in Table 1. Before test administration, students completed a short self-report questionnaire in which they were asked to consent to the inclusion of their data in the analyses. The procedure required the students to generate a unique personal ID and to transcribe a multi-digit dataset number to the questionnaire. Thus, participants' data could no longer be traced to individual people. This procedure was chosen after consulting the local data protection authorities. Students were given the option of participating in a lottery to win audio players, digital cameras and book tokens.

Educational context

All participants were enrolled in a traditional curriculum with a typical structure of 2 years of basic science instruction followed by 3 years of clinical content. At the end of the second academic year, students are required to pass an extensive national licensing examination assessing the entire basic science content covered in the first 2 years. From the third year onwards, there is no further systematic teaching of basic sciences. Progress tests are mandatory for traditional curriculum students from the beginning of the third academic year (semester 5). Sufficient data were available for the third and fourth academic years (i.e. semesters 5–8). We therefore restricted our analyses to this phase of medical education.

Table 1 Sample structure showing cohorts, measurement occasions and number of participants per semester of study

	Semester			
Cohort	**5**	**6**	**7**	**8**
1	C_1T_2 n = *90*	C_1T_3 n = *88*		
2	C_2T_1 n = *150*	C_2T_2 n = *149*	C_2T_3 n = *126*	
3		C_3T_1 n = *96*	C_3T_2 n = *115*	C_3T_3 n = *100*
4			C_4T_1 n = *96*	C_4T_2 n = *118*

C_1–C_4 = cohorts 1–4; T_1–T_3 = measurement occasions 1–3; n = *total number of students per cohort at that measurement occasion*

Missing data

A total of 926 students sat the Berlin Progress Test in Medicine (PTM) during the period of interest. Of these, 772 (83%) participated in the study. However, data for 177 students could not be included in the analysis mainly because of transcription errors arising as a result of the data protection procedure described above. These difficulties led to partially incomplete data and we therefore employed full information maximum likelihood[47,48] estimation in our analysis.

Measures

The Berlin PTM was used to assess the development of undergraduates' clinical and basic science knowledge over time. Progress testing monitors the acquisition of medical knowledge by testing students repeatedly on the content they are expected to have mastered by the end of their education. Test items are sampled from an item pool covering the entire undergraduate medical curriculum across medical disciplines and functional systems. In the PTM, 200 items per test are drawn randomly from a database containing more than 5000 items. Items that have been administered in a test are excluded from the sampling procedure for 2 years.

All students take the same test, regardless of their level of training. Participants are able to choose a 'don't know' option to 'dismiss' content they cannot readily answer. Guessing is penalised and the total test score is obtained by negative marking of incorrect answers. As the PTM is implemented as a formative progress test, no pass or fail decisions are made on the basis of the results. However, it is compulsory for students to take the test twice per year. (For further details, see Nouns and Georg[49]).

All items are administered in single best-answer multiple-choice question format and, when possible, consist of clinical vignettes (see Appendix S1, online). Independent expert committees have judged each item with regard to the appropriateness of its content and level of formality. The number of distractors ranges from three to a maximum of eight. Item and distractor analysis is conducted after the administration of each progress test. Items with extreme difficulty values or negative discrimination and those flagged as problematic by students are submitted to post review and recoded or excluded from the analysis. The basic science items cover anatomy (including gross anatomy, histology and embryology), biochemistry and chemistry, physiology and physics. The mean ± standard

deviation (SD) Cronbach's alpha for the basic science subset (n = 25 items) across each combination of test occasion and cohort was 0.76 ± 0.05. The clinical items cover 12 clinical disciplines, including: internal medicine; surgery; paediatrics; gynaecology; psychiatry; ophthalmology; dermatology; emergency medicine; orthopaedics; family medicine; urology, and neurology. The mean ± SD Cronbach's alpha for the clinical subset (n = 120 items) across each combination of test occasion and cohort was 0.89 ± 0.03. Items that did not exclusively assess clinical or basic science content (e.g. behavioural sciences, biometrics, epidemiology) were excluded from the analysis (n = 55 items).

Both subsets were coded according to the percentage of correct answers and scaled to the T-metric (mean ± SD: 50 ± 10) at the first point of measurement (October 2010). Variables at consecutive points of measurement were rescaled with reference to the first measurement occasion. This rescaling does not alter psychometric properties or statistical comparisons, but puts the variables on a comparable metric and enhances presentation.[50]

Statistical procedure

We used a relatively new statistical procedure, a bivariate dual change score model (DCSM),[50,51] in order to analyse the developmental relationship between students' biomedical knowledge and clinical knowledge. The bivariate DCSM itself relates to the technique of cross-lagged correlation[52] and to a special application of confirmatory factor analysis to longitudinal data (latent growth curve modelling [LGM][53]). We give details on the methodological procedure in Appendix S1 and focus here on the conceptual features of this statistical technique. We will first detail how the development of a single variable is characterised in a DCSM. Thereafter, we will focus on the bivariate case, in which two changing variables are analysed.

In a DCSM, the change of *one* variable is primarily characterised by the (latent) differences between two consecutive measurement occasions ('time lags'). Those differences (e.g. gains or losses) are decomposed into two components: firstly, an additive (i.e. 'stable' or 'time-independent') component, which is similar to the conventional slope in a regression analysis, and secondly, a multiplicative or auto-proportional component accounts for temporal dependencies of sequential occasions. For example, it might be assumed that students have a certain body of medical knowledge that grows in a steady fashion. In this scenario, students show some stable accumulation of medical knowledge, which could be expressed by the finding that 'every semester, students are able to answer a further 20 items correctly' (i.e. the constant, additive component). Furthermore, students might not be able to fully retrieve (or simply partially 'forget') what they have learned in earlier terms. This could be designated as a different component which accounts for an amount of medical knowledge that was learned at an earlier time-point and is only partially retrievable at a later moment and would hence form the multiplicative or proportional component; this could be expressed in the finding that 'students typically forget half of the content they have learned within a semester'. Of course, there is no claim that the parameters estimated in a DCSM are 'true' representations of the concrete fractions that such components might represent. The estimated components do not unambiguously permit conclusions of the type '80% of the learned knowledge is forgotten and the most is learned completely new'. Importantly, the decomposition into two change components permits a statistically more flexible formalisation of developmental patterns and the analysis of correlates of the several change components.

For our analysis, however, the most worthwhile feature is the inclusion of longitudinal dependencies or 'couplings' between *two* changing variables (i.e. performances on biomedical and clinical knowledge tests). Hence, we used a bivariate extension of dual change score models. Here, gains and losses in performance are modelled for each variable, as we have described. Moreover, two coupling components are added to the model to account for the interrelatedness of the two change processes. Consequently, the performance on the biomedical items on one measurement occasion is supposed to be related to changes (e.g. gains or losses) in performance on the subset of clinical items within the following semester, and vice versa.

In the following text, we use more technical abbreviations that are common in the literature on DCSMs. The additive, proportional and coupling components are labelled with the Greek letters α, β and γ, respectively. The full model is given as a path diagram in Fig. 1.

As our research question is directly related to the coupling components between performances in basic science (BSc) and clinical knowledge (CK), we explicate several models that statistically formalise different related developmental patterns. The logic of our analysis is based on the comparison of the

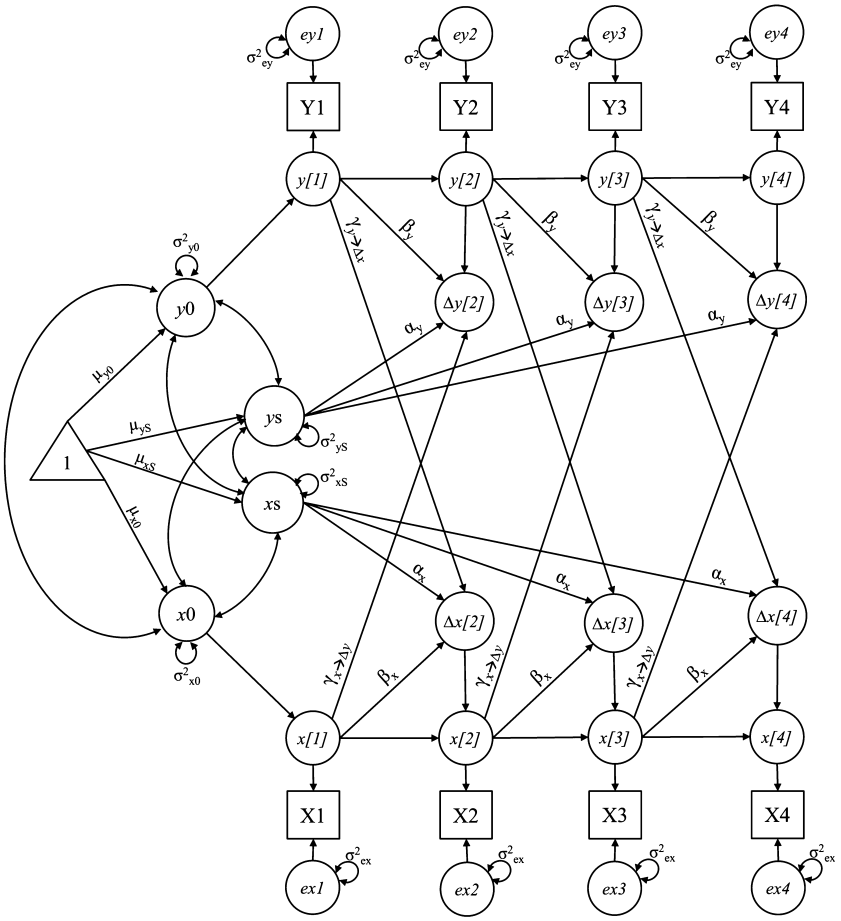

Figure 1 Graphical representation of the bivariate dual change score model for two variables X and Y (50, 51); Triangles indicate constants (i.e., means and intercepts). Observed (manifest) variables are represented by squares. All unlabelled paths–except for the covariances–are fixed to 1. Latent variables are indicated by circles. Regression weights are represented by one-headed arrows; variance and covariances, by two-headed arrows.

empirical fit of different plausible models, as is common in SEM approaches. We estimated five models and compared their relative fits to the data.

Three models were specified to explore bidirectional relations between biomedical and clinical knowledge. Firstly, model 1-a estimates two coupling parameters between BSc and CK that are allowed to differ in their magnitude ($\gamma_{BSc\rightarrow\Delta CK}$ = ? and $\gamma_{CK\rightarrow\Delta BSc}$ = ?). [Question marks refer to the parameters that are freely estimated in a specific model. The γ-value designates the coupling of one variable on the later change in the other variable. Hence, $\gamma_{BSc\rightarrow\Delta CK}$ reads as: 'couplings of earlier state of BSc on subsequent change in CK'.]

A restriction of this model is the equal coupling model 1-b. Here again, reciprocal dependencies are assumed, but they are expected *not* to differ in their magnitude ($\gamma_{BSc\rightarrow\Delta CK} = \gamma_{CK\rightarrow\Delta BSc}$ = ?). Hence, there is no 'leading' effect of one variable on change in the other variable. The most restrictive model 1-c assumes that there are no coupling effects (i.e. the two variables develop independently of each other). Hence, both coupling parameters are fixed to zero ($\gamma_{BSc\rightarrow\Delta CK} = \gamma_{CK\rightarrow\Delta BSc} = 0$).

By contrast with these bidirectional models, unidirectional models specify the hypothesis that states in one variable precede changes in the other. Model 1-d specifies unidirectional couplings from basic science knowledge to changes in clinical knowledge ($\gamma_{BSc\rightarrow\Delta CK}$ = ?; $\gamma_{CK\rightarrow\Delta BSc} = 0$). Conversely, model 1-e specifies unidirectional couplings from clinical knowledge to changes in biomedical knowledge ($\gamma_{BSc\rightarrow\Delta CK} = 0$; $\gamma_{CK\rightarrow\Delta BSc}$ = ?).

Models 1-a to 1-e were analysed using Mplus 6.1.[54] Model fit to the data was determined by the root mean square error of approximation (RMSEA), the Tucker–Lewis index (TLI) and the comparative fit index (CFI). The RMSEA compares the model-

implied covariances with those observed in the sample. Larger RMSEA values indicate a larger discrepancy between the specified model and the data. Values of $\leq$ 0.06 are considered good; values of < 0.10 may still be considered acceptable.[55,56] Both the TLI and the CFI test the specified model against a 'null model' or 'independence model', which basically assumes the observed variables to be uncorrelated. In general, values close to 0.95 or higher are considered acceptable.[55,56] In-depth discussions of fit indices are available in, for instance, Kline[55] and Hu and Bentler.[56] Models were compared directly based on the statistical significance of differences in chi-squared values ($\Delta\chi^2$) and the corresponding change in degrees of freedom (Δd.f.).

RESULTS

Descriptives

Descriptive statistics for the observed variables across the covered educational phase are given in Table 2. In general, the data show a decline in the amount of retrieved basic science knowledge and a steady growth in clinical knowledge.

Table 2 Descriptive statistics for total scores for clinical knowledge (CK) and basic science (BSc) knowledge

Domain	Semester	Score (T-*scaled*), *mean* ± *SD*	Correct, %, mean ± SD
CK	5	50.00 ± 10.00	20.02 ± 8.33
	6	68.70 ± 12.62	27.51 ± 10.52
	7	91.34 ± 14.42	36.58 ± 12.01
	8	104.90 ± 15.78	42.01 ± 13.15
BSc	5	50.00 ± 10.00	65.30 ± 15.52
	6	39.25 ± 12.00	51.26 ± 18.63
	7	37.45 ± 11.83	48.90 ± 18.36
	8	34.51 ± 12.17	45.07 ± 18.89

T-scaled scores are scaled to a T-metric with a mean = 50 and SD = 10 at the first measurement occasion (semester 5)
SD = standard deviation

Dual change score model

Measures of model fit are given in Table 3. Tests of chi-squared differences among the models (Table 3, $\Delta\chi^2$) indicated that the full coupling model and the unidirectional model with couplings from BSc to ΔCK fitted the data equally well. All other models showed a statistically significant loss in fit relative to the full coupling model. We therefore consider the full coupling (1-a) and the unidirectional BSc→ΔCK (1-d) models to provide the best explanation for our data.

Parameter estimates for the full coupling and the unidirectional BSc→ΔCK model are given in Table 4. The main difference is that the coupling parameter from CK on ΔBSc is fixed to zero in the unidirectional BSc→ΔCK model (1-d), but estimated in the full coupling model. Although this parameter is statistically significant in the full coupling model,

Table 3 Comparisons of goodness of fit for models 1-a to 1-e

Model		Goodness of fit					
		χ^2	d.f.	RMSEA (90% CI)	CFI	TLI	$\Delta\chi^2$ (Δd.f.)
1-a	Full coupling	48.60	19	0.051 (0.034–0.069)	0.97	0.96	–
1-b	Equal coupling	66.98	20	0.063 (0.046–0.080)	0.95	0.94	18.38 (1)*
1-c	No coupling	67.13	21	0.061 (0.045–0.077)	0.95	0.94	18.53 (2)*
1-d	Unidirectional BSc→ΔCK	52.02	20	0.052 (0.035–0.069)	0.97	0.96	3.42 (1) (p = 0.064)
1-e	Unidirectional CK→ΔBSc	67.00	20	0.063 (0.047–0.080)	0.95	0.94	18.40 (1)*

* p < *0.001 indicates significant loss in fit relative to the full coupling model*
RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval; CFI = comparative fit index; TLI = Tucker–Lewis index; BSc = basic science; CK = clinical knowledge

Table 4 Parameter estimates for the full coupling and unidirectional BSc→ΔCK models

Parameter	Model	
	1-a full coupling (SE)	1-d unidirectional BSc→ΔCK (SE)
Regression coefficients		
Auto-proportion BSc (β_{BSc})	− 0.97 (0.11)‡	− 0.76 (0.06)‡
Auto-proportion CK (β_{CK})	− 0.41 (0.07)‡	− 0.35 (0.07)‡
Coupling BSc→ΔCK ($\gamma_{BSc\rightarrow\Delta CK}$)	− 1.07 (0.23)‡	− 0.87 (0.23)‡
Coupling CK→ΔBSc ($\gamma_{CK\rightarrow\Delta BSc}$)	− 0.08 (0.04)*	0 (–/–)
Means		
BSc intercept ($\mu_{BSc(0)}$)	49.55 (0.63)‡	49.41 (0.61)‡
BSc slope ($\mu_{BSc(S)}$)	41.35 (6.99)‡	27.30 (2.40)‡
CK intercept ($\mu_{CK(0)}$)	48.85 (0.61)‡	48.71 (0.61)‡
CK slope ($\mu_{CK(S)}$)	92.43 (14.38)‡	80.01 (14.38)‡
Variances		
BSc intercept ($\sigma^2_{BSc(0)}$)	52.46 (11.70)‡	45.76 (12.02)‡
BSc slope ($\sigma^2_{BSc(S)}$)	80.72 (22.47)‡	44.66 (8.21)‡
BSc residual ($\sigma^2_{BSc(e)}$)	62.01 (4.14)†	62.92 (4.26)‡
CK intercept ($\sigma^2_{CK(0)}$)	26.59 (10.51)*	26.07 (10.40)*
CK slope ($\sigma^2_{CK(S)}$)	212.2 (73.60)†	153.31 (61.77)*
CK residual ($\sigma^2_{CK(e)}$)	73.48 (5.81)†	73.83 (5.76)‡

*$p < 0.05$; †$p < 0.01$; ‡$p < 0.001$ (two-tailed)
BSc = basic science; CK = clinical knowledge; SE = standard error

it is relatively small ($\gamma_{CK\rightarrow\Delta BSc}$, − 0.08; $p < 0.05$). As indicated by the non-significant loss of model fit, the coupling from CK on ΔBSc ($\gamma_{CK\rightarrow\Delta BSc}$) can be omitted without losing substantial information. It is important to note that fixing a small parameter (− 0.08) causes relatively large shifts in the other parameters estimated. For example, the coupling effect $\gamma_{BSc\rightarrow\Delta CK}$ decreases from − 1.07 in the full coupling model to − 0.87 in the unidirectional BSc→ΔCK model. Indeed, this shift is accompanied by smaller slopes in both BSc and CK. Hence, the different components (additive, auto-proportion, coupling) retain their relative importance and the implied means for the consecutive measurement moments do not differ drastically. (For details on the calculation, please refer to Appendix S1.)

Our results can be described as follows: given the full coupling model, which takes influences of clinical knowledge on the retention of basic science knowledge into account, there is a relatively large negative effect of basic science knowledge on gains in clinical knowledge ($\gamma_{BSc\rightarrow\Delta CK} = -1.07$). This effect is counterbalanced by both an auto-proportional component ($1 + \beta_{CK} = 1 - 0.41 = 0.59$; see Equation 1 in Appendix S1) and constant gains in clinical knowledge ($\mu_{CK(S)} = 92.43$). By contrast, the couplings from CK to ΔBSc are marginal but also negative, indicating that students who perform better on clinical content items tend to have weaker retention of basic science knowledge.

Figure 2 illustrates various combinations of high or low initial achievements in the two domains and relationships to subsequent development. Here, lower initial average achievement in basic science knowledge is associated with higher gains in clinical knowledge (Fig. 2a, c). By contrast, higher initial average achievement in basic science knowledge is

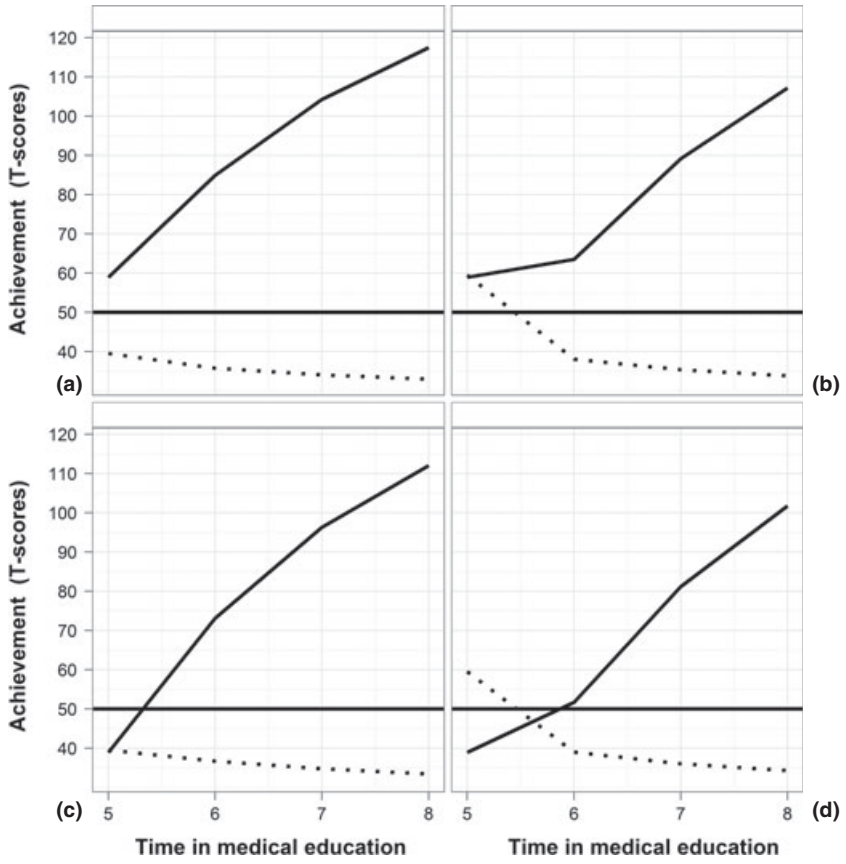

Figure 2 Combinations of high and low achievement in biomedical knowledge (dotted line) and clinical knowledge (unbroken line) in progress tests in 598 undergraduate medical students and relationships with subsequent development. The unbroken horizontal line represents the average T-scaled score of 50 at semester 5. Higher/lower initial achievement is characterised by plus/minus one standard deviation which is scaled to *SD = 10* at semester 5.

associated with lower gains in clinical knowledge (Fig. 2b, d).

DISCUSSION

Using the work of Boshuizen, Schmidt, Norman and Rikers[21–24] and Woods *et al.* [35–38] as our starting platform, we hypothesised that biomedical knowledge functions as a critical variable for the construction of a coherent base of clinical knowledge. However, our findings did not confirm this hypothesis; rather, they contradicted it, indicating that biomedical knowledge seems to be negatively related to the acquisition of clinical knowledge. Given this unexpected effect, our findings warrant careful discussion.

It is possible that individual differences in motivational or cognitive characteristics function as an unobserved 'cause' and thus explain our results. Students may vary in the degree of effort they put into acquiring knowledge in different domains, depending on their personal interests and the professions they aspire to enter. In this case, our results may simply reflect personal preferences. Likewise, cognitive characteristics such as learning strategies or behaviours may be differentially useful at different stages of medical education. As has been noted elsewhere, the amount of knowledge to be learned in pre-clinical courses, in particular, may force medical students to revert to rote learning and memorisation.[57,58] Although the ability to memorise extensive amounts of factual knowledge may be an important factor in high-stakes assessments, such as national licensing examinations, it may be a far less efficient strategy for the acquisition of clinical knowledge, in which cases are processed not in isolation, but in a more holistic manner.[28]

Another interpretation of our results may be that students with less biomedical knowledge have fewer

problems constructing a coherent base of clinical knowledge. Similar processes of interference (i.e. 'performance decrements caused by irrelevant information'[59]) are well-known phenomena in the learning sciences (for a review, see Dempster and Corkill[59]). One explanation may be that the acquired biomedical knowledge induces a proactive interference effect in the form of 'a general disruptive effect of prior learning on the ability to retrieve more recently learned information'.[60] Such an effect suggests a negative influence on the processing of information stored in long-term memory.[61,62] In other words, it is possible that students with more elaborate biomedical knowledge have difficulties encoding and retrieving clinical content.

Another form of problematic interaction may be related to inappropriate transfer. Specifically, students may try to apply their basic science knowledge to create a coherent representation of clinical cases. However, this transfer of knowledge may be ambiguous as the similarities and differences between underlying pathophysiological mechanisms and the signs and symptoms of clinical presentations remain rather inconclusive. Similar findings have been reported in the literature on analogous transfer,[63] in which even the successful identification of correct correspondences does not necessarily lead to appropriate transfer.[64]

Limitations and strengths

Critically, our study was conducted in a specific educational context. All students were following a traditional curriculum based almost entirely on classic course structures such as lectures and seminars, and teaching in the pre-clinical years was conducted exclusively in a massed teaching approach. This approach is likely to have a considerable influence on the patterns of acquisition of medical knowledge.

Furthermore, we used a relatively specific form of knowledge assessment. Although progress test data have been used previously in related research scenarios,[44,65–67] they are summaries of responses to multiple-choice items and thus reflect only a relatively specific form of assessment of medical knowledge.

A common problem in longitudinal research is the analysis of incomplete data. This issue is relevant to the current study. However, the participation rates seem to be acceptable and the absence of data mainly reflected 'technical' issues, as described in the Methods section. We therefore argue that it is plausible to assume that these missing values are 'missing at random'.[68] In this scenario, full information maximum likelihood estimation is known to perform well and to be superior to other techniques for dealing with the consequences of missing data.[69,70]

Finally, as a result of the non-experimental nature of our study, it is important to note that the observed effect may be mediated by other, unobserved factors beyond those discussed above. Consequently, the interrelationship between students' biomedical and clinical knowledge cannot be interpreted as causal, but strictly correlational.

Despite these limitations the analysis presented here may still offer a 'glimpse at the directional dynamics within the considered system of variables'.[71] Our study complements previous laboratory research to give a developmental perspective using an SEM-based approach. To the best of our knowledge, this material represents the first analysis of longitudinal data on the temporal interaction between students' states of biomedical knowledge and subsequent changes in performance on clinical knowledge tests. In addition, we covered a time span of 2 years of medical education.

CONCLUSIONS

Our findings suggest that the worst case scenario in pre-clinical medical education and the teaching of basic sciences may involve not just the perceived inappropriateness – a 'null effect' – of detailed biomedical knowledge, but a negative relationship, at least under the specific conditions presented here. However, our results do not necessarily contradict the conception of the basic sciences as a mnemonic device, as outlined by Woods.[16] By contrast, the negative relationship between the two domains observed in our study underpins the need for research on how to transfer the work of Woods and colleagues[35–38] to the daily practice of pre-clinical medical education. However, our findings hint at possible difficulties that have been largely neglected thus far.

As we mentioned in the Introduction, the theories on clinical reasoning and medical expertise proposed by Boshuizen, Schmidt, Norman and Rikers,[21–24] on the one hand, and Patel *et al.* [19] and Kaufman *et al.*,[7] on the other, imply different reference standards for the structure of undergraduate medical education. Consequently, we would certainly expect differences between integrated and traditional curricula in terms of interrelations between

biomedical and clinical knowledge in the development of medical knowledge. In our study, we analysed data from a traditional curriculum, in which the two domains of knowledge are addressed sequentially and separately. According to Boshuizen and Schmidt, the process of encapsulation can be facilitated by an integrated teaching approach, in which both biomedical knowledge and clinical knowledge are addressed in parallel or simultaneously.[22]

Against this background, we would expect that applying the methodology used in the present analysis to other contexts using data from curricular comparisons[41,44,67] or spanning a broader phase of medical education might yield further insights into the developmental relationship between biomedical and clinical knowledge.

Contributors: All authors contributed substantially to the conception and design of the study. SKS and MH conducted the statistical analysis. All authors contributed to the interpretation of the results. SKS drafted the manuscript. All authors contributed to its critical revision and approved the final manuscript for submission.
Acknowledgements: We thank Jana Struzena and Melanie Hunnenmörder for their important role in test administration, collection and preparation of the analysed data.
Funding: The work of Stefan K. Schauber is funded by the German Federal Ministry of Education and Research (BMBF) within the project "Competence Acquisition and Learning Trajectories in Medical Training" (Grant 01JG1055).
Conflicts of interest: none.
Ethical approval: The processing of personalised data was approved by the local data protection authorities. The study was deemed exempt from full review requirements by the institutional ethical review board of Charité Medical University Berlin.

REFERENCES

1 Flexner A. *Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching*. New York, NY: Carnegie Foundation for the Advancement of Teaching 1910.
2 Kuper A, D'Eon M. Rethinking the basis of medical knowledge. *Med Educ* 2011;**45** (1):36–43.
3 Dornan T, McKendree J, Robbé IJ. Medical education in an age of complexity, uncertainty and reflection. A coda to the Flexner centenary. *Med Educ* 2011;**45** (1):2–6.
4 Frank JR. *The CanMEDS 2005 Physician Competency Framework: Better Standards, Better Physicians, Better Care*. Ottawa, ON: Royal College of Physicians and Surgeons of Canada 2005.
5 Bergman EM, van de Vleuten CPM, Scherpbier AJJA. Why don't they know enough about anatomy? A narrative review. *Med Teach* 2011;**33** (5):403–9.
6 Drake RL, McBride JM, Lachman N, Pawlina W. Medical education in the anatomical sciences: the winds of change continue to blow. *Anat Sci Educ* 2009;**2** (6):253–9.
7 Kaufman DR, Yoskowitz NA, Patel VL. Clinical reasoning and biomedical knowledge: implications for teaching. In: Higgs J, Jones MA, Loftus S, Christensen N, eds. *Clinical Reasoning in the Health Professions*, 3rd edn. Amsterdam; Boston, MA: Butterworth Heinemann/Elsevier 2008;137–49.
8 Raftery AT. Anatomy teaching in the UK. *Surgery* 2007;**25** (1):1–2.
9 Pryde FR, Black SM. Anatomy in Scotland: 20 years of change. *Scott Med J* 2005;**50** (3):96–8.
10 Patel K, Moxham B. Attitudes of professional anatomists to curricular change. *Clin Anat* 2006;**19** (2):132–41.
11 Whitcomb ME. The teaching of basic sciences in medical schools. *Acad Med* 2006;**81** (5):413–4.
12 Winkelmann A. Anatomical dissection as a teaching method in medical school: a review of the evidence. *Med Educ* 2007;**41** (1):15–22.
13 Fahrer M. Art macabre: is anatomy necessary? *ANZ J Surg* 2001;**71** (6):333–4.
14 Norman G. Teaching basic science to optimise transfer. *Med Teach* 2009;**31** (9):807–11.
15 Eva KW, Neville AJ, Norman GR. Exploring the aetiology of content specificity. *Acad Med* 1998;**73** (10):S1–5.
16 Woods NN. Science is fundamental: the role of biomedical knowledge in clinical reasoning. *Med Educ* 2007;**41** (12):1173–7.
17 Norman G. Research in clinical reasoning: past history and current trends. *Med Educ* 2005;**39** (4):418–27.
18 Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: the role of experience. *Med Educ* 2007;**41** (12):1140–5.
19 Patel VL, Evans DA, Groen GJ. Reconciling basic science and clinical reasoning. *Teach Learn Med* 1989;**1** (3):116–21.
20 Patel VL, Groen GJ, Norman GR. Reasoning and instruction in medical curricula. *Cogn Instr* 1993;**10**:335–78.
21 Boshuizen HP, Schmidt HG. On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cogn Sci* 1992;**16** (2):153–84.
22 Boshuizen HPA, Schmidt HG. The development of clinical reasoning expertise. In: Higgs J, Jones MA, Loftus S, Christensen N, eds. *Clinical Reasoning in the Health Professions*, 3rd edn. Amsterdam; Boston, MA: Butterworth Heinemann/Elsevier 2008;113–21.

23 Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise. *Acad Med* 1990;**65** (10):611–21.
24 Schmidt HG, Rikers RMJP. How expertise develops in medicine: knowledge encapsulation and illness script formation. *Med Educ* 2007;**41** (12):1133–9.
25 Rikers RM, Schmidt HG, Boshuizen HP. Knowledge encapsulation and the intermediate effect. *Contemp Educ Psychol* 2000;**25** (2):150–66.
26 Wimmers PF, Schmidt HG, Verkoeijen PPJL, van de Wiel MWJ. Inducing expertise effects in clinical case recall. *Med Educ* 2005;**39** (9):949–57.
27 Schmidt HG, Boshuizen HP. On the origin of intermediate effects in clinical case recall. *Mem Cogn* 1993;**21** (3):338–51.
28 Eva KW, Norman GR, Neville AJ, Wood TJ, Brooks LR. Expert–novice differences in memory: a reformulation. *Teach Learn Med* 2002;**14**:257–63.
29 van de Wiel MWJ, Boshuizen HPA, Schmidt HG. Knowledge restructuring in expertise development: evidence from pathophysiological representations of clinical cases by students and physicians. *Eur J Cogn Psychol* 2000;**12** (3):323–56.
30 Rikers RMJP, Loyens SMM, Schmidt HG. The role of encapsulated knowledge in clinical case representations of medical students and family doctors. *Med Educ* 2004;**38** (10):1035–43.
31 Rikers RMJP, Schmidt HG, Moulaert V. Biomedical knowledge: encapsulated or two worlds apart? *Appl Cogn Psychol* 2005;**19** (2):223–31.
32 Rikers RMJP, Loyens S, te Winkel W, Schmidt HG, Sins PHM. The role of biomedical knowledge in clinical reasoning: a lexical decision study. *Acad Med* 2005;**80** (10):945–9.
33 Hmelo CE. Problem-based learning: effects on the early acquisition of cognitive skill in medicine. *J Learn Sci* 1998;**7** (2):173–208.
34 Patel VL, Groen GJ, Norman GR. Effects of conventional and problem-based medical curricula on problem solving. *Acad Med* 1991;**66** (7):380–9.
35 Woods NN, Brooks LR, Norman GR. The value of basic science in clinical diagnosis: creating coherence among signs and symptoms. *Med Educ* 2005;**39** (1):107–12.
36 Woods NN, Brooks LR, Norman GR. It all makes sense: biomedical knowledge, causal connections and memory in the novice diagnostician. *Adv Health Sci Educ Theory Pract* 2007;**12** (4):405–15.
37 Woods NN, Brooks LR, Norman GR. The role of biomedical knowledge in diagnosis of difficult clinical cases. *Adv Health Sci Educ Theory Pract* 2007;**12** (4):417–26.
38 Woods NN, Howey EHA, Brooks LR, Norman GR. Speed kills? Speed, accuracy, encapsulations and causal understanding. *Med Educ* 2006;**40** (10): 973–9.
39 Woods NN, Neville AJ, Levinson AJ, Howey EH, Oczkowski WJ, Norman GR. The value of basic science in clinical diagnosis. *Acad Med* 2006;**81**: S124–7.
40 Muijtjens AMM, Schuwirth LWT, Cohen-Schotanus J, Thoben AJNM, van der Vleuten CPM. Benchmarking by cross-institutional comparison of student achievement in a progress test. *Med Educ* 2008;**42** (1):82–8.
41 Muijtjens AMM, Schuwirth LWT, Cohen-Schotanus J, van der Vleuten CPM. Differences in knowledge development exposed by multi-curricular progress test data. *Adv Health Sci Educ Theory Pract* 2008;**13** (5):593–605.
42 Schauber S, Nouns ZM. Using the cumulative deviation method for cross-institutional benchmarking in the Berlin progress test. *Med Teach* 2010;**32** (6):471–5.
43 Verhoeven BH, Snellen-Balendong HA, Hay IT *et al*. The versatility of progress testing assessed in an international context: a start for benchmarking global standardisation? *Med Teach* 2005;**27** (6):514–20.
44 Verhoeven BH, Verwijnen GM. An analysis of progress test results of PBL and non-PBL students. *Med Teach* 1998;**20** (4):310–6.
45 Bell RQ. Convergence: an accelerated longitudinal approach. *Child Dev* 1953;**24** (2):145–52.
46 Schaie KW. *Longitudinal Studies of Adult Psychological Development*. New York, NY: Guilford Press 1983;8–12.
47 Enders CK. A primer on maximum likelihood algorithms available for use with missing data. *Struct Equ Modeling* 2001;**8** (1):128–41.
48 Little RJA, Rubin DB. *Statistical Analysis with Missing Data*, 2nd edn. Hoboken, NJ: John Wiley & Sons 2002;97–199.
49 Nouns ZM, Georg W. Progress testing in German-speaking countries. *Med Teach* 2010;**32** (6):467–70.
50 McArdle JJ, Hamagami F. Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data. In: Collins LM, Sayer A, eds. *New Methods for the Analysis of Change*, 1st edn. Washington, DC: American Psychological Association 2001;139–75.
51 McArdle JJ. Latent variable modelling of differences and changes with longitudinal data. *Annu Rev Psychol* 2009;**60** (1):577–605.
52 Rogosa D. A critique of cross-lagged correlation. *Psychol Bull* 1980;**88**:254–8.
53 Preacher KJ. *Latent Growth Curve Modeling*. Los Angeles, CA: Sage Publications 2008;22–53.
54 Muthén LK, Muthén BO. *Mplus: Statistical Analysis with Latent Variables: User's Guide*, 3rd edn. Los Angeles, CA: Muthén & Muthén 2011.
55 Kline RB. *Principles and Practice of Structural Equation Modeling*, 3rd edn. New York, NY: Guilford Press 2011;189–227.
56 Hu L, Bentler PM. Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999;**6** (1):1–55.

57 Mattick K, Knight L. High-quality learning: harder to achieve than we think? *Med Educ* 2007;**41** (7):638–44.
58 Lindblom-Ylänne S, Lonka K. Individual ways of interacting with the learning environment – are they related to study success? *Learn Instr* 1998;**9** (1):1–18.
59 Dempster FN, Corkill AJ. Interference and inhibition in cognition and behaviour: unifying themes for educational psychology. *Educ Psychol Rev* 1999;**11** (1):1–88.
60 Lustig C, Hasher L. Working memory span: the effect of prior learning. *Am J Psychol* 2002;**115** (1):89–101.
61 Unsworth N, Engle RW. The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol Rev* 2007;**114** (1):104–32.
62 Lustig C, May CP, Hasher L. Working memory span and the role of proactive interference. *J Exp Psychol Gen* 2001;**130** (2):199–207.
63 Duit R, Roth W, Komorek M, Wilbers J. Fostering conceptual change by analogies – between Scylla and Charybdis. *Learn Instr* 2001;**11** (4–5):283–303.
64 Novick LR, Holyoak KJ. Mathematical problem solving by analogy. *J Exp Psychol Learn Mem Cogn* 1991;**17** (3):398–415.
65 de Bruin ABH, Schmidt HG, Rikers RMJP. The role of basic science knowledge and clinical knowledge in diagnostic reasoning: a structural equation modelling approach. *Acad Med* 2005;**80** (8):765–73.
66 Nouns Z, Schauber S, Witt C, Kingreen H, Schuttpelz-Brauns K. Development of knowledge in basic sciences: a comparison of two medical curricula. *Med Educ* 2012;**46**:1206–14.
67 van derVeken J, Valcke M, deMaeseneer J, Schuwirth L, Derese A. Impact on knowledge acquisition of the transition from a conventional to an integrated contextual medical curriculum. *Med Educ* 2009;**43** (7):704–13.
68 Rubin DB. Inference and missing data. *Biometrika* 1976;**63** (3):581–92.
69 Enders C, Bandalos D. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Modeling* 2001;**8** (3):430–57.
70 Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods* 2002;**7** (2): 147–77.
71 Lövdén M, Ghisletta P, Lindenberger U. Social participation attenuates decline in perceptual speed in old and very old age. *Psychol Aging* 2005;**20** (3):423–34.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Methodological supplement and sample items.

Received 2 October 2012; editorial comments to author 10 December 2012, 20 February 2013; accepted for publication 13 March 2013