

Een interfacultaire vergelijking van probleemoplossende vaardigheden

L.W.T. Schuwirth, B.H. Verhoeven, A.J.J.A. Scherpbier, E.M.A. Mom, J. Cohen-Schotanus, H.J.M. van Rossum, C.P.M. van der Vleuten

Samenvatting

Vanuit de vraag of er verschillen te vinden zijn in probleemoplossende vaardigheden tussen studenten die in een Probleem Gestuurd Onderwijs (PGO) systeem studeren en studenten die in een traditioneel onderwijssysteem studeren is aan studenten uit alle zes jaargroepen van twee medische faculteiten (Groningen en Maastricht) een casusgerichte toets voorgelegd. De toets bestond uit 60 casus, gebaseerd op echte patiënten. Per casus werden vragen gesteld waarbij de studenten beslissingen moesten nemen over de casus. Als onderzoeksvariabelen zijn goedscores en responstijden gebruikt. In de eerste drie jaargroepen worden geen verschillen gevonden, in de laatste drie jaargroepen treden progressief verschillen op ten voordele van de Maastrichtse studenten. De meest voor de hand liggende verklaring zou zijn dat het hier een curriculumeffect betreft, mede gezien het feit dat het effect juist in die jaargroepen optreedt waarvan de Groningse studenten nog in het traditionele onderwijssysteem studeren, terwijl in de eerste drie jaargroepen waarin in die faculteit het 'Curriculum 2000' geïmplementeerd is, geen verschillen te vinden zijn. Er moet echter steeds bedacht worden dat dergelijke vergelijkingen moeilijk te ontdoen zijn van mogelijke methodologische bezwaren die de interpretatie bemoeilijken.

Inleiding

Bij het implementeren van een vernieuwing is het van onderwijskundig belang aan te tonen

dat deze vernieuwing ook daadwerkelijk een verbetering inhoudt. De ontwikkelaars van probleemgeoriënteerde onderwijscurricula (PGO) hebben daarom in het verleden vele vergelijkingen opgezet met andere, meer traditionele curricula. Deze vergelijkingen hebben vaak een dubbel doel gehad. In eerste instantie waren ze bedoeld om aan te tonen dat PGO in ieder geval niet slechter was dan het traditionele onderwijs, maar ook is geprobeerd aan te tonen dat PGO-curricula ten aanzien van sommige aspecten superieur zijn aan traditionele onderwijssystemen. De belangrijkste uitkomsten worden opgesomd in drie overzichtsartikelen.¹⁻³ Verschillende aspecten zijn in deze vergelijkingen betrokken: feitenkennis, vaardigheden, probleemoplossende vaardigheden, studentsatisfactie, kosten, et cetera, maar vanuit de meeste vergelijkingen handelen over kennis en probleemoplossende vaardigheden. Wat deze beide aspecten betreft wordt in de meerderheid van de studies noch een verschil in kennis, noch een verschil in probleemoplossende vaardigheden gevonden.^{1,3,4} Toch rapporteren sommige studies een verschil in feitenkennis ten nadele van de PGO-studenten. Als dit het geval is, dan is dit met name in het domein van de basisvakken.⁵⁻⁷ Anderzijds wordt soms gevonden dat het vermogen om klinische vragen op te lossen bij PGO-studenten wat nadrukkelijker aanwezig is dan bij hun collega's die meer traditioneel onderwijs volgen.^{1,3,7,8}

In hoeverre dit laatste een aanwijzing vormt voor het veronderstelde verschil in medisch probleemoplossende vaardigheden is echter nog niet duidelijk. Mede debet hieraan mag het

feit zijn dat het in het verleden gebruikte instrumentarium niet voldoende geschikt was om probleemoplossend vermogen te toetsen. Bij de ontwikkeling van deze instrumenten is men er in het verleden vaak van uitgegaan dat probleemoplossend vermogen een generieke vaardigheid is, die onafhankelijk van het aangeboden probleem is. Inzichten vanuit de cognitieve psychologie echter tonen dat een dergelijke expertise nauw samenhangt met het specifieke probleem dat aangeboden wordt en dat de mate van transfer van het ene probleem naar het andere zeer klein is.⁹⁻¹³ Hieruit vloeit voort dat probleemoplossend vermogen domeinspecifiek is, en dat bij toetsing dus een groot aantal verschillende problemen aan de kandidaat aangeboden moet worden om tot een generaliseerbare uitspraak te kunnen leiden.¹⁴ Dit heeft geleid tot de ontwikkeling die op dit moment het meest in zwang is, namelijk dat geprobeerd wordt tot uitspraken over probleemoplossende vaardigheden te komen door gebruik te maken van casus in de zogenaamde 'key-feature approach'.¹⁵ Hierbij worden alleen relevante contextuele en klinische symptomen en bevindingen gerapporteerd en slechts enkele vragen gesteld die gericht zijn op essentiële beslissingen. Zodoende is het mogelijk vele verschillende casus in een toets op te nemen. Verder wordt onderzoek gedaan naar de mogelijkheid andere parameters dan scores bij de beoordeling te gebruiken, zoals bijvoorbeeld responstijden.^{13 16}

Gebruikmakend van een toets bestaande uit casus in een 'key-feature approach', waarbij zowel score als responstijd gemeten werden, is twee jaar geleden een pilotstudy gedaan, waarbij vierde- en zesdejaars studenten van de medische faculteiten van Groningen en Maastricht met elkaar vergeleken werden (on gepubliceerd). De vraagstelling hierbij was of er verschillen meetbaar waren tussen beide faculteiten in gemiddelde scores en/of responstijden. Er werden toen in beide jaargroepen geen significante verschillen in gemiddelde scores

gevonden. De vierdejaars uit Groningen deden echter significant langer over hun toets dan hun Maastrichtse collega's. De steekproefgroottes toentertijd waren echter zeer klein ($n=11, 16, 8$ en 14) en de vraag deed zich voor of dit een toevallige bevinding was of dat er werkelijk geen verschillen meetbaar zouden zijn. In een andere vergelijking tussen beide faculteiten, waarbij gebruik is gemaakt van korte casuïstiek, werden wel significante verschillen gevonden.⁸ Hierbij werd de studenten gevraagd om bij 30 casus een (differentiaal) diagnose te geven. In deze studie werd een verschil gevonden tussen de studenten uit Groningen en Maastricht in de laatste drie jaargroepen ten voordele van de Maastrichtse studenten. Dit heeft geleid tot de eerste vraagstelling van de huidige studie of er wederom een verschil in gemiddelde scores en responstijden gevonden kan worden tussen de studenten uit Groningen en Maastricht. De hypothese is dat een verschil gevonden wordt in beide parameters ten voordele van de Maastrichtse studenten (hogere gemiddelde scores en lagere responstijden).

De tweede vraagstelling heeft te maken met het feit dat in Groningen drie jaar geleden een nieuw patiëntgeoriënteerd curriculum gestart is. Uitgaande van de aanname dat studenten in dergelijke curricula beter in staat zijn om problemen op te lossen dan hun meer traditioneel opgeleide collega's, doet zich de vraag voor of een curriculumeffect in Groningen waar te nemen is.⁸ Omdat het curriculumeffect niet los te zien is van een algemeen jaargroep-effect, wordt hiervoor de hypothese geformuleerd dat als er een verschil in scores tussen beide faculteiten te vinden is, dit voornamelijk in de laatste drie jaargroepen zichtbaar zal zijn.

Methode

Proefpersonen

Van beide faculteiten hebben proefpersonen

uit alle zes jaargroepen meegedaan. Per jaargroep en per faculteit is gestreefd naar groepen van 30 studenten. In het totaal hebben 355 studenten meegedaan. De aantallen per jaargroep en per faculteit zijn weergegeven in tabel 1. Slechts in twee jaargroepen heeft het aantal studenten dat zich aanmeldde voor deelname het streefaantal overtroffen. Alle studenten hebben een financiële vergoeding voor hun deelname gekregen, en binnen iedere faculteit en binnen iedere jaargroep is een extra geldbedrag uitgelooft voor die student die de hoogste score behaalt. Dit is gedaan om het prestatie-element van een werkelijke toets enigszins te emuleren.

Instrumenten

Als instrument is gekozen voor een toets die bestaat uit 60 casus uit de eerste lijn.¹⁷ Deze toets is aangeboden als computergestuurde casusgerichte toetsing (CCT). Hierbij is bij alle casus een "key-feature approach" gehanteerd.¹⁵ De casus zijn geschreven door verschillende huisartsen en gebaseerd op echte patiëntcontacten. Verschillende vraagvormen zijn gebruikt, variërend van gesloten vraagvormen tot open vraagvormen met de mogelijkheid tot een kort antwoord. Dit is gedaan omdat uit een groot aantal onderzoeken gebleken is dat de invloed van de inhoud van een vraag op datgene wat met de vraag gemeten wordt van groter belang is dan de vorm.¹⁸⁻²¹ Het gebruik maken van verschillende vraagvormen, waarbij de vraagvorm gekozen wordt op basis van de inhoud van de vraag is ons inziens daarom een reëlere en meer valide keuze.

De antwoordsleutels zijn in eerste instantie door de auteurs van de casus gegeven, maar daarna door een panel van 10 experts gevalideerd. Voor de in deze toets gebruikte casus geldt dat de experts het allen eens waren over de juiste antwoordsleutel. De eindscore werd via het computerprogramma berekend en ge-

definieerd als het percentage correcte antwoorden.

Per vraag is de vraagbeantwoordingstijd geregistreerd door het computerprogramma. Deze is voor iedere student gesommeerd over het totale aantal vragen. Van de Maastrichtse studenten is tevens bepaald of zij qua kennis vergelijkbaar zijn met de andere leden van hun jaargroep door de scores van hun reguliere toetsen te vergelijken met die van hun jaargenoten. Deze waren in alle jaargroepen niet significant van elkaar verschillend behalve in het eerste jaar. Alleen in deze jaargroep was de gemiddelde score van de deelnemers significant hoger dan die van hun jaargenoten. Voor Groningen bleek een dergelijke controle niet mogelijk.

Statistische analyse

In eerste instantie zijn van alle scores en de (gesommeerde) responstijden de gemiddelden per jaargroep en per universiteit met de bijbehorende standaarddeviaties bepaald. Vervolgens zijn binnen iedere faculteit de gemiddelde scores getest tegen de jaargroep met behulp van een one-way ANOVA met Tukey's HSD als post-hoc analyse. Verschillen tussen de faculteiten per jaargroep zijn in eerste instantie onderzocht door 95%-betrouwbaarheidsintervallen te bepalen. Eveneens is een two-way ANOVA (plaats x jaargroep tegen score) gedaan over alle gemiddeldes.

Resultaten

Tabel 1 en 2 geven de beschrijvende statistiek van respectievelijk de scores van de deelnemers uitgesplitst naar jaargroep en plaats, en van de responstijden. Uit tabel 1 blijkt dat er in de eerste vier jaargroepen geen verschillen tussen beide faculteiten zichtbaar zijn wat betreft scores. In het vijfde en zesde jaar zijn er wel verschillen. Uit de one-way ANOVA's blijkt dat in beide faculteiten een significant

Tabel 1. Beschrijvende statistiek van de scores (percentages correcte antwoorden).

| Jaargroep | Groningen | | | Maastricht | | |
|-----------|-----------|-----|-----|------------|-----|----|
| | N | Gem | Sd | N | Gem | Sd |
| 1 | 32 | 44% | 5% | 30 | 44% | 6% |
| 2 | 30 | 46% | 4% | 30 | 47% | 4% |
| 3 | 30 | 51% | 4% | 30 | 50% | 6% |
| 4 | 30 | 53% | 6% | 29 | 54% | 5% |
| 5 | 30 | 52% | 6% | 27 | 56% | 6% |
| 6 | 25 | 56% | 13% | 32 | 65% | 6% |

Tabel 2. Beschrijvende statistiek van de responstijden (in minuten).

| Jaargroep | Groningen | | | Maastricht | | |
|-----------|-----------|-------|-------|------------|-------|-------|
| | N | Gem | Sd | N | Gem | Sd |
| 1 | 32 | 24.95 | 9.37 | 30 | 29.67 | 11.92 |
| 2 | 30 | 25.06 | 7.85 | 30 | 30.05 | 10.22 |
| 3 | 30 | 26.03 | 7.17 | 30 | 28.37 | 7.62 |
| 4 | 30 | 26.63 | 8.35 | 29 | 26.97 | 5.90 |
| 5 | 30 | 27.02 | 7.92 | 27 | 30.07 | 10.77 |
| 6 | 25 | 25.92 | 11.33 | 32 | 25.23 | 8.07 |

hoofdeffect van jaargroep op score gevonden wordt (Groningen: $F(5,171)=12,50$ $p<.00001$; Maastricht: $F(5,172) = 58,75$ $p<.00001$). De post-hoc analyses geven aan dat in Maastricht het effect voornamelijk verklaard wordt door verschillen tussen de laatste twee jaren en de andere jaren, terwijl de verschillen in Groningen meer bepaald worden door de eerste twee jaren ten opzichte van de andere. Op de 2×2 ANOVA (jaargroep \times plaats tegen score) wordt naast de beide significante hoofdeffecten een significant interactie-effect gevonden ($F(1,5,354) = 5,50$, $p<.000$).

De responstijden vertonen een ander beeld. Uit tabel 2 wordt duidelijk dat er geen patroon in de verschillende jaargroepen binnen een faculteit te vinden is. De verschillen tussen jaargroepen bereiken nergens een significantie op een 5% niveau. Dit geldt tevens voor de verschillen tussen de faculteiten. Wel is het opvallend dat bijna in alle jaargroepen de

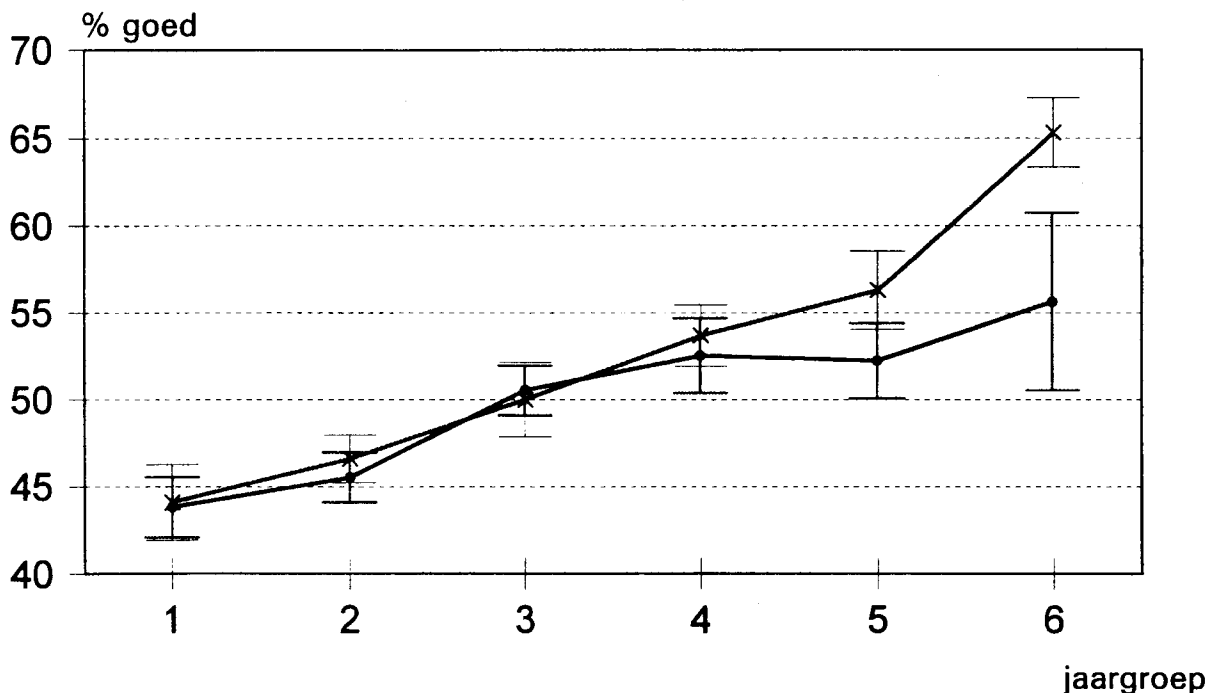
studenten uit Maastricht langer over de toets doen dan de studenten uit Groningen.

Om een duidelijker indruk te krijgen van de trend die de toetsscores door de jaargroepen heen vertonen, zijn ze, met hun 95%-betrouwbaarheidsintervallen, uitgezet in figuur 1. Hoewel het hier een transversale meting betreft, zijn de gegevens desondanks gerepresenteerd door middel van een lijngrafiek om de trends duidelijker zichtbaar te laten zijn. Het valt op dat de lijnen in de eerste drie jaargroepen samenlopen en vanaf jaargroep vier beginnen te divergeren ten voordele van de Maastrichtse studenten. Vanaf jaargroep vier zijn de Groningse studenten afkomstig uit het traditionele curriculum.

Discussie

De hypotheses worden voor een belangrijk deel ondersteund door de resultaten. Welis-

Gemiddelde scores van Groningen en Maastricht



Figuur 1. Gemiddelde met 95%-betrouwbaarheidsintervallen van Groningen (•) en Maastricht (x).

waar worden er geen significante verschillen in responstijden gevonden, maar er worden wel significante verschillen in gemiddelde scores tussen de faculteiten gevonden. Ook geldt dat deze verschillen alleen gevonden worden in die jaargroepen waarin studenten van het oude curriculum zitten.

Toch is enige voorzichtigheid met het trekken van conclusies geboden. Vergelijkingen van deze soort zijn niet eenvoudig omdat er nogal wat oncontroleerbare variabelen zijn.

Schmidt rapporteert een aantal belangrijke mogelijke storingsbronnen bij dergelijke onderzoeken.²² Zijn bezwaren vallen in drie categorieën uiteen, te weten populatie-effecten, steekproefeffecten en instrumentatie-effecten. De populaties zijn a priori al niet goed vergelijkbaar omdat studenten bewust voor een bepaalde faculteit kiezen. De verschillende uitvalspercentages maken dat de populaties in de loop van de studie alleen nog maar minder vergelijkbaar worden. De steekproeven kunnen meestal ook niet zuiver samenge-

steld worden: vaak wordt gewerkt met vrijwilligers en is het moeilijk om voldoende grote groepen te vinden. Randomisatie is dan niet mogelijk. De instrumenten die gehanteerd worden, zijn vaak bekender in de ene faculteit dan in de andere, wat de meting vertroebelt. In onze vergelijking zijn echter niet al deze storingsbronnen evenzeer aanwezig.

De populatie-effecten met betrekking tot de initiële studiekeuze en de verschillen in uitvalspercentages blijven waarschijnlijk wel onverminderd hun invloed uitoefenen. De invloed van curriculumherzieningen echter, kan hier niet per se als een storingsbron beschouwd worden. Een van de vraagstellingen is min of meer of deze invloed in de trend binnen de faculteit Groningen ook zichtbaar is. Het lijkt erop dat dit inderdaad het geval is.

Wat betreft de steekproefeffecten blijft staan dat gewerkt is met betaalde vrijwilligers zonder enige randomisatie. Toch heeft het feit dat de best scorende uit iedere jaargroep (en binnen iedere faculteit) een extra som geld

tegemeet kon zien mogelijk een reële toets-situatie meer geëmuleerd. Ook is gebleken dat de groepen (binnen de Maastrichtse studentengroep) in hun reguliere toetsresultaten nage-noeg vergelijkbaar zijn met hun jaargenoten.

De laatste mogelijke storing die op het eerste gezicht een forse invloed lijkt te hebben, maar bij nadere beschouwing veel minder belangrijk is, wordt gevormd door de lokale bekendheid met het meetinstrument. Studenten uit Maastricht zijn bekender met het fenomeen CCT, sommigen hebben zelfs de letterlijke toets 1 tot 2 jaar van tevoren gezien. Deze toets was toentertijd echter eenmalig, studenten hebben de toets niet meer mogen inzien, noch hebben zij materiaal mee naar buiten mogen nemen. Gezien de verdwijncurves die voor kennis gelden, mag met waarschijnlijkheid worden aangenomen, dat een herkenningseffect niet heeft bijgedragen aan het gevonden verschil.^{23 24} Er is geprobeerd de verschillen in bekendheid met de vorm van het instrument zo goed mogelijk te neutraliseren door middel van een gestandaardiseerde instructie, met een korte oefensessie voor alle deelnemers. Ervaringen met CCT binnen de reguliere toetsing in Maastricht leren dat het programma zodanig is, dat na deze korte instructie onbekendheid geen factor van belang meer is.

De eenvoudigste verklaring zou zijn, dat het gevonden verschil puur een PGO-effect is. Tenslotte treedt het pas op in die jaargroepen waarin het deel van de studenten uit Groningen traditioneel opgeleid is, terwijl de Maastrichtse studenten allen PGO-onderwijs hebben gehad. In de eerste drie jaren, waarin beide studentgroepen een PGO-opleiding hebben gevolgd, wordt geen verschil gezien.

Er zijn echter nog alternatieve verklaringen te bedenken. Zo kan het feit dat het hier een toets betreft die vrij nauwkeurig gericht is op het co-assistentenschap huisartsgeneeskunde of in ieder geval op de eerstelijnsgezondheidszorg een belangrijke alternatieve verklaring

zijn. De stageopbouw was ten tijde van dit experiment in beide faculteiten verschillend in die zin dat huisartsgeneeskunde een stage was die de Maastrichtse studenten als een van de eerste gevolgd hebben, terwijl het in Groningen de laatste stage is.

Een derde verklaring zou beter passen bij de gevonden resultaten van het geheel van doorgevoerde vergelijkingen. Tegelijkertijd met deze studie is een vergelijking gemaakt waarbij een kennistoets over vaardigheden gebruikt is. Hierbij werden verschillen vanaf het eerste jaar in gemiddelde scores gevonden in alle jaargroepen ten voordele van Maastricht.²⁵ Het is mogelijk dat het longitudinale skillslabprogramma aan de Maastrichtse faculteit studenten zodanig voorbereidt op hun co-assistentenschappen, dat zij zich in deze co-assistentenschappen niet bezig hoeven te houden met het aanleren van fysisch diagnostische vaardigheden. Theoretisch zouden zij daardoor meer aandacht kunnen besteden aan het oplossen van patiëntproblemen en dus in die zin meer rendement uit hun co-schappen kunnen halen. Dit zou dan zichtbaar moeten worden in een toets gebaseerd op patiëntproblemen, zoals beschreven in dit artikel.

Literatuur

1. Albanese MA, Mitchell S. Problem-based learning: a review of literature on its outcomes and implementation issues [published erratum appears in Acad Med 1993 Aug;68(8):615]. Acad Med 1993;68(1):52-81.
2. Schmidt HG, Dauphinee WD, Patel VL. Comparing the effects of problem-based and conventional curricula in an international sample. J Med Educ 1987;62(4):305-15.
3. Vernon DT, Blake RL. Does problem-based learning work? A meta-analysis of evaluative research. Acad Med 1993;68(7):550-63.
4. Verwijnen M, Vleuten CPM van der, Imbos T. A comparison of an innovative medical school with traditional schools: an analysis in the cognitive domain. In: Nooman ZM, Schmidt HG, Ezzat ES, eds. Innovation in medical education: an evaluation of its present status. New York: Springer Publishing Company, 1987:40-9.

5. Albano MG, Cavallo F, Hoogenboom R, Magni F, Majoor G, Manenti F. et al. An international comparison of knowledge levels of medical students: the Maastricht Progress Test. *Med Educ* 1996;30:239-45.
6. Saunders NA, McIntosh J, McPherson J, Engels CE. A comparison between University of Newcastle and University of Sydney final-year students, knowledge and competence. In: Nooman ZM, Schmidt HG, Ezzat ES, eds. *Innovation in medical education: an evaluation of its present status*. New York: Springer Publishing Company, 1987:50-4.
7. Baca E, Mennin SP, Kaufman A, Moore-West M. Comparison between a problem-based, community-oriented track and a traditional track within one medical school. In: Nooman ZM, Schmidt HG, Ezzat ES, eds. *Innovation in medical education: an evaluation of its present status*. New York: Springer Publishing Company, 1990:9-26.
8. Schmidt HG, Machiels-Bongaerts M, Hermans H, Cate TJ ten, Venekamp R, Boshuizen HPA. The development of diagnostic competence: comparison of a problem-based, and integrated, and a conventional medical curriculum. *Acad Med* 1996;71(6):658-64.
9. Regehr G, Norman GR. Issues in cognitive psychology: implications for professional education. *Acad Med* 1996;71(9):988-1001.
10. Boreham NC. The dangerous practice of thinking. *Med Educ* 1994;28:172-9.
11. Schmidt H, Norman G, Boshuizen H. A cognitive perspective on medical expertise: theory and implications. *Acad Med* 1990;65(10): 611- 22.
12. Norman G, Tugwell P, Feightner J, Muzzin L, Jacoby L. Knowledge and clinical problem-solving. *Med Educ* 1985;19:344-56.
13. Norman GR. Problem-solving skills, solving problems and problem-based learning. *Med Educ* 1988;22:270-86.
14. Elstein AS, Shulmann LS, Sprafka SA. *Medical problem-solving: an analysis of clinical reasoning*. Cambridge, MA: Harvard University Press, 1978.
15. Bordage G. An alternative approach to PMP's: the "key-features" concept. In: Hart IR, Harden R, eds. *Further developments in assessing clinical competence, Proceedings of the second Ottawa conference*. Montreal.: Can-Heal Publications Inc, 1987:59-75.
16. Vleuten CPM van der, Schuwirth LWT, Ronteltap CFM. A cognitive psychological interpretation of a few remarkable psychometric findings. In: Rothman AI, Cohen R, eds. *Proceedings of the Sixth Ottawa Conference on Medical Education*. Toronto: University of Toronto Bookstore Custom Publishing, 1995:506-8.
17. Schuwirth LWT, Vleuten CPM van der, Kock CA de, Peperkamp AGW, Donkers HHLM. Computerized case-based testing: a modern method to assess clinical decision making. *Med Teach* 1996;18(4):295-300.
18. Ward WC. A comparison of free-response and multiple-choice forms of verbal aptitude tests. *Applied Psychological Measurement* 1982;6(1):1-11.
19. Norman GR, Smith EKM, Powles AC, Rooney PJ, Henry NL, Dodd PE. Factors underlying performance on written tests of knowledge. *Med Educ* 1987;21:297-304.
20. Maatsch J, Huang R. An evaluation of the construct validity of four alternative theories of clinical competence. 25th ann. RIME conference. Chicago: AAMC, 1986:69-74.
21. McGuire C. Letter to the editor. *Teaching and Learning in Medicine* 1994;6(2):74.
22. Schmidt HG. Innovative and conventional curricula compared: what can be said about their effects? In: Nooman ZM, Schmidt HG, Ezzat ES, eds. *Innovation in medical education: an evaluation of its present status*. New York: Springer Publishing Company, 1990:1-7.
23. Semb GB, Ellis JA. Knowledge taught in school: what is remembered? *Review of Educational Research* 1996;64(2):253-86.
24. Sisson JC, Swartz RD, Wolf FM. Learning, retention and recall of clinical information. *Med Educ* 1992;26(6):454-61.
25. Scherpbier AJJA. *Kwaliteit van vaardigheidsonderwijs gemeten [proefschrift]*. Maastricht: Universiteit Maastricht, 1997.

DE AUTEURS

L.W.T. Schuwirth, arts, is als medisch-onderwijskundige verbonden aan de vakgroep Onderwijsontwikkeling en Onderwijsresearch van de Universiteit Maastricht.

B.H. Verhoeven, arts, docent skillslab, Universiteit Maastricht.

Dr. A.J.J.A. Scherpbier, arts, hoofd van het skillslab van de Faculteit der Geneeskunde, Universiteit Maastricht.

E.M.A. Mom, huisarts, verbonden aan de vakgroep Huisartsgeneeskunde van de Universiteit Maastricht.

Dr. J. Cohen-Schotanus, psycholoog, verbonden aan de Faculteit der Medische Wetenschappen, Rijksuniversiteit Groningen.

Prof. dr. H.J.M. van Rossum is als hoogleraar medische onderwijskunde verbonden aan de Faculteit der Medische Wetenschappen, Rijksuniversiteit Groningen.

Prof. dr. C.P.M. van der Vleuten, psycholoog, hoogleraar vakgroep Onderwijsontwikkeling en Onderwijsresearch, Universiteit Maastricht.

Correspondentieadres:

Drs. L. Schuwirth, vakgroep Onderwijsontwikkeling en Onderwijsresearch, Faculteit der Geneeskunde, Universiteit Maastricht, Postbus 616, 6200 MD Maastricht.