



**Het toetsen van
vaardigheden is een
essentieel onderdeel
van onderwijs.**

Van checklist naar context

Vaardigheden toetsen met oog voor de praktijk

■ Peter Renden

Dr. P.G. Renden is universitair docent bij de vakgroep Onderwijsontwikkeling en Onderwijsonderzoek aan de Universiteit Maastricht. E-mail: peter.renden@maastrichtuniversity.nl

■ Sanne Rovers

Dr. S.F.E. Rovers is universitair docent bij de vakgroep Onderwijsontwikkeling en Onderwijsonderzoek aan de Universiteit Maastricht. E-mail: s.rovers@maastrichtuniversity.nl

■ Jimmy Frèrejean

Dr. J. Frèrejean is universitair docent bij de vakgroep Onderwijsontwikkeling en Onderwijsonderzoek aan de Universiteit Maastricht. E-mail: j.frerejean@maastrichtuniversity.nl

Het toetsen van vaardigheden is een essentieel onderdeel van onderwijs in verschillende domeinen. Denk bij vaardigheden bijvoorbeeld aan het voeren van gesprekken als HR-medewerker, maar ook aan technische handelingen, zoals het vervangen van remblokken als automonteur of het inbrengen van een infuus als verpleegkundige. Een belangrijke functie van het toetsen van vaardigheden is het vaststellen of een student bekwaam is om in de praktijk professioneel, veilig en effectief te handelen.

Dit vereist dat de toetsing representatief is voor de complexe en geïntegreerde aard van dat professioneel handelen (Van der Vleuten & Schuwirth, 2005). Daarmee wordt bijvoorbeeld bedoeld dat een verpleegkundige niet alleen technisch correct een infuusnaald moet inbrengen, maar tegelijk ook klinische signalen moet interpreteren, het welzijn van de patiënt moet bewaken en het handelen moet afstemmen op veranderende situaties, zoals bij een patiënt die tijdens de handeling bleek en klam wordt, mogelijk als teken dat diegene gaat flauwvallen.

Het is echter de vraag of veel vaardigheidstoetsen aan de criteria van validiteit voldoen. Indien dat niet het geval is, zou dat betekenen dat het gedrag van de student tijdens een toets slechts beperkt voorspellend is voor het functioneren in de beroepspraktijk (Van Merriënboer et al., 2025). In dit artikel bespreken we twee valkuilen in gangbare manieren van

vaardigheidstoetsing. De eerste valkuil is dat vaardigheden vaak geïsoleerd worden beoordeeld, buiten de functionele context waarin ze moeten worden toegepast (Renden & Dikken, 2023). De tweede valkuil is het gebruik van strikte beoordelingscriteria, zoals gedetailleerde checklists die de uitvoering opdelen in afzonderlijke stappen (Koerner & Staller, 2018). Beide benaderingen doen geen recht aan de complexiteit van professioneel handelen. Aan de hand van een sterk onderbouwd onderwijsontwerpmodel, het 4C/ID-model (zie Kader 1) (Van Merriënboer et al., 2025), zullen we deze valkuilen beschouwen en bediscussieren en bieden we alternatieven die beter aansluiten bij realistische taakuitvoering en transfer. Daarbij gebruiken we een uitgewerkt voorbeeld uit de fysiotherapie (zie Frèrejean et al., 2025) om te illustreren hoe toetsing wel valide en praktijkgericht vormgegeven kan worden.

Kader 1. Een korte uitleg van het 4C/ID-model

Wat is het 4C/ID-model?

Het 4C/ID-model is een ontwerpmodel voor onderwijs en toetsing van complexe vaardigheden. Centraal staat het principe dat beroepscompetenties niet los, maar in samenhang binnen realistische taken moeten worden aangeleerd en beoordeeld. Het model biedt richtlijnen voor het opzetten van samenhangende leerprogramma's met leertaken die variëren in aard en context, die stapsgewijs toenemen in complexiteit en zijn voorzien van passende niveaus van ondersteuning. Deze taken worden aangevuld met een doelgericht aanbod van theorie en feedback. Ook voor toetsprogramma's onderstreept het model het belang van toetsing in realistische en betekenisvolle contexten, waarbij vaardigheden worden beoordeeld als geïntegreerd professioneel handelen in plaats van als geïsoleerde elementen.

Valkuil 1: vaardigheden geïsoleerd toetsen

In de beroepspraktijk komt het nauwelijks voor dat iemand geïsoleerde deeltaken uitvoert. Een fysiotherapeut voert niet alleen onderzoek uit, bijvoorbeeld naar knieklachten, maar baseert de keuze en uitvoering daarvan op wat de patiënt tijdens het anamnesegeprek (de intake) heeft verteld en op de daaruit afgeleide hypothesen. Daarnaast informeert de fysiotherapeut de patiënt over de stappen van het onderzoek en vraagt actief naar hoe bepaalde bewegingen aanvoelen. Toch komt het in het onderwijs nog vaak voor dat vaardigheden als geïsoleerde deeltaken (los van de beroepscontext) worden getoetst (Renden & Dikken, 2023). De onderliggende aanname is dat beheersing van alle losse onderdelen automatisch leidt tot succesvolle uitvoering van de gehele taak. Onderzoek laat echter

zien dat deze aanname niet klopt: vaardigheden die buiten hun context zijn geleerd en getoetst, blijken vaak slecht overdraagbaar naar volledige, realistische situaties (Koerner & Staller, 2021).

In tegenstelling tot een gefragmenteerde benadering bij leren en toetsen, benadrukken hedendaagse onderwijsmodellen juist het belang van het uitvoeren van hele taken die lijken op de beroepspraktijk (Frèrejean et al., 2019). Bij het toetsen van gedrag in zulke taken kan beoordeeld worden of de student alle aspecten geïntegreerd kan toepassen — precies zoals dat in het werkveld nodig is (Van Merriënboer et al., 2025). Het toetsgedrag van de student wordt daarmee een realistischer indicatie van de transfer naar de beroepspraktijk (zie bijvoorbeeld Andersen et al., 2016; Hutter et al., 2023). Om het verschil tussen beide manieren van toetsen te illustreren beschrijft Kader 2 twee fictieve toetsvormen van een fysiotherapeutische onderzoekstechniek: één waarin de vaardigheid geïsoleerd wordt getoetst (Casus 1) en één waarin de vaardigheid is ingebed in een hele taak (Casus 2).

De twee toetsvormen in Kader 2 laten zien hoe de toetsopzet het gedrag van de student kan beïnvloeden. Hoewel er op het eerste gezicht wellicht weinig mis lijkt met de toets in Casus 1, zegt deze weinig over hoe de student zou handelen in de beroepspraktijk. Het is immers onwaarschijnlijk dat een patiënt uit zichzelf vraagt om een specifieke test, zoals de Jobe-test. Casus 2 is meer representatief voor de beroepspraktijk en vraagt dan ook om ander gedrag van de student. Uiteraard zijn deze voorbeelden bewust zwart-wit neergezet om het contrast te benadrukken en is de werkelijkheid in de onderwijspraktijk genuanceerder. In de praktijk bestaat er een continuüm tussen volledig geïsoleerde en volledig geïntegreerde toetsing. Toch maken de voorbeelden duidelijk waarom toetsing binnen hele taken een betrouwbaarder beeld geeft van de beroepsbevoegdheid dan geïsoleerde toetsing.

Kader 2. Twee voorbeeld van het toetsen van onderzoekstechnieken binnen de fysiotherapie

Casus 1.

Tom, een student fysiotherapie, bevindt zich in een stationstoets. De stationstoets is zo opgezet dat hij bij elk station een specifieke vaardigheid moet laten zien. Bij een van de stations heeft hij een squat-oefening (kniebuigingen) voorgedaan en uitgelegd aan een patiënt. Bij een ander station heeft hij aan een patiënt uitleg gegeven over een leefstijlinterventie gericht op meer bewegen. Hij is nu bij een station aangekomen waarin hij wordt getoetst op het uitvoeren van de Jobe-test (onderzoek van de schouder) bij een patiënt. Hij krijgt een korte opdracht: "Voer de Jobe-test uit." Hij let er goed op dat hij de armen van de patiënt in de juiste positie plaatst, de juiste weerstand geeft en de test fysiek correct uitvoert.

Casus 2.

Sophie, een student fysiotherapie, wordt getoetst op het uitvoeren van lichamelijk onderzoek bij een patiënt met schouderklachten. Voordat zij begint, heeft ze een anamnesegeprek gevoerd waarin de patiënt aangaf vooral pijn te ervaren bij het reiken boven schouderhoogte. Op basis van deze informatie besluit Sophie om specifieke testen uit te voeren die gericht zijn op de rotator cuff (groep van spieren en pezen in de schouder). Tijdens het onderzoek informeert zij de patiënt over wat hij kan verwachten en stelt tussendoor vragen wanneer ze onverwachte pijnreacties ziet. Ze past haar aanpak aan op basis van de reacties van de patiënt en de hypothesen die zij tijdens het klinisch redeneren bijstelt.

Valkuil 2: het gebruik van gedetailleerde checklists

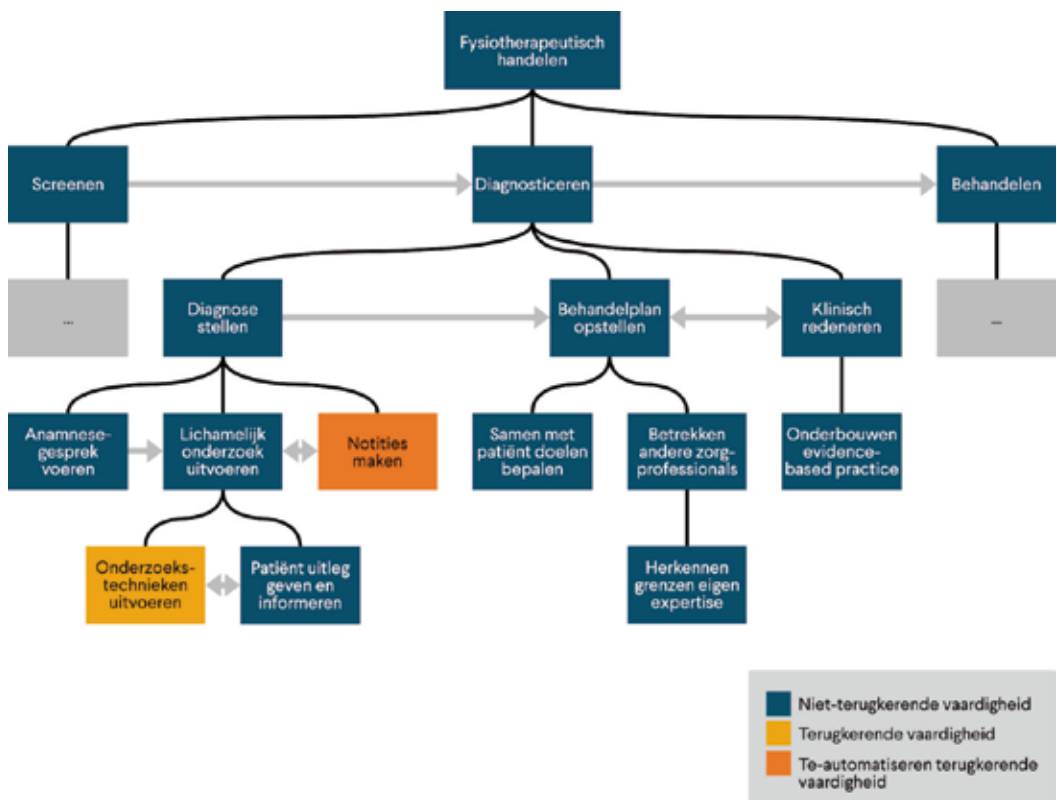
Het ligt voor de hand dat in Casus 1 en 2 verschillende toetsinstrumenten worden gebruikt. Bij toetsing is het van belang dat het uiteindelijke toetsinstrument aansluit op het te toetsen gedrag (Ten Cate & Scheele, 2007). In Casus 1 ligt de focus op de technische uitvoering volgens een checklist: als de student een stap overslaat of als een positionering net afwijkt van het protocol, wordt dat waarschijnlijk fout gerekend. In Casus 2 ligt de nadruk op hoe de student de informatie uit het anamnesegeprek gebruikt om passende onderzoekstechnieken te selecteren en uit te voeren. Daarnaast wordt beoordeeld of de uitvoering van de onderzoekstechnieken kan leiden tot een juiste diagnose en hoe zij tijdens het onderzoek communiceert met de patiënt. Hoewel de afstemming van toetsinstrument op eindgedrag eenvoudig lijkt, blijkt de praktijk

vaak weerbarstiger. Vaardigheden die op het eerste gezicht een constant eindgedrag suggereren (zoals in Casus 1), vragen binnen geïntegreerde taken vaak om meer flexibiliteit (zoals in Casus 2). Dat komt doordat factoren zoals patiëntkenmerken, onverwachte reacties, complicaties of contextuele omstandigheden het handelen beïnvloeden (Verdoes et al., 2025; Wulf & Shea, 2002). In zulke situaties schieten checklists tekort: ze waarderen professionele variatie niet en classificeren adaptief gedrag soms onterecht als fout (Van Merriënboer et al., 2025). Een belangrijke valkuil is dat beoordelaars, zelfs bij hele taken, de taak alsnog ontleden in losse handelingen die apart worden getoetst. De onderlinge samenhang van deelvaardigheden raakt dan uit beeld. Het gevolg is dat het geheel van de taak wordt teruggebracht tot een optelsom van stappen, terwijl juist de integratie van die stappen bepalend is voor professioneel handelen.

Denken vanuit een vaardighedenhiërarchie

De eerder beschreven valkuilen in vaardigheidstoetsing kunnen deels worden voorkomen door te werken met een vaardighedenhiërarchie (Figuur 1). In het 4C/ID-model worden samenhangende vaardigheden niet als losse, op zichzelf staande eenheden beschouwd, maar als onderdelen van een hiërarchisch gestructureerd netwerk (Van Merriënboer et al., 2025). Een dergelijk netwerk maakt zichtbaar welke vaardigheden voorwaardelijk zijn voor andere, complexere handelingen, en hoe ze gezamenlijk bijdragen aan een succesvolle taakuitvoering (Frèrejean et al.,

2025). Door die relaties expliciet te maken, wordt inzichtelijk hoe vaardigheden functioneel met elkaar samenhangen. Een hiërarchie doet daarmee meer dan een lijst: in plaats van een eenvoudige opsomming, laat een hiërarchie duidelijk zien welke vaardigheden elkaar beïnvloeden, in welke volgorde ze moeten worden uitgevoerd en welke mate van integratie vereist is. Dit ondersteunt niet alleen het onderwijsontwerp, maar ook het selecteren van toetsvormen die passen bij het beoogde leerproces. Zonder zo'n structuur dreigt het risico dat deelvaardigheden getraind of getoetst worden zonder oog voor hun rol binnen het geheel.



Figuur 1. Een vereenvoudigde vaardighedenhiërarchie voor fysiotherapeutisch handelen. Overgenomen uit: Frèrejean et al., 2025. Omdat we in de uitleg alleen focussen op diagnosticeren hebben we, om de omvang van de figuur beperkt te houden, de uitwerking van screenen en behandelen in deze figuur niet meegenomen.

Figuur 1 toont een voorbeeld van een vereenvoudigde vaardighedenhiërarchie voor fysiotherapeutisch handelen, met als onderdelen screenen, diagnosticeren en behandelen. In deze toelichting ligt de nadruk op het onderdeel diagnose stellen. In de figuur wordt onderscheid wordt gemaakt tussen drie typen vaardigheden (Van Merriënboer et al., 2025):

- Niet-terugkerende vaardigheden, zoals het stellen van een diagnose, zijn vaardigheden waarbij geen uitvoering exact hetzelfde is en die telkens nieuwe afstemming vragen.
- Terugkerende vaardigheden, zoals het uitvoeren van onderzoekstechnieken, zijn vaardigheden die volgens een standaard-procedure worden uitgevoerd en waarbij de uitvoering vaak hetzelfde (of ten minste vergelijkbaar) is.
- Te automatiseren terugkerende vaardigheden, zoals het noteren van bevindingen, zijn vaardigheden die vrijwel zonder bewuste aansturing verlopen.

Om een diagnose te kunnen stellen voert een fysiotherapeut onder meer onderzoekstechnieken uit, zoals beschreven in de twee casussen. Deze technieken kunnen, vanuit hun aard als terugkerende vaardigheden die vaak op vergelijkbare manier worden uitgevoerd, in theorie geïsoleerd worden getoetst. Maar zodra hun functionele rol binnen de gehele taak wordt meegenomen, zoals zichtbaar wordt in de hiërarchie, ligt geïntegreerde toetsing veel meer voor de hand. Het expliciteren van een vaardighedenhiërarchie helpt dus bij het selecteren van toetsvormen en bijbehorende beoordelingsinstrumenten die de praktijk weerspiegelen. Daarmee wordt een belangrijke stap gezet richting valide, contextgevoelige toetsing van beroepsvaardigheden.

Enige nuances voor de toetspraktijk

Hoewel toetsing binnen hele, authentieke taken het ideaal blijft, is het in de praktijk niet altijd haalbaar om een volledige casus in één toetsmoment te vangen. In zulke gevallen kan

een stationstoets een waardevol alternatief zijn, mits die zorgvuldig wordt ontworpen. Een bruikbare metafoer is die van een treinreis: de student stapt halverwege in en reist een stuk mee. Dat betekent dat de eerdere input (zoals gegevens uit een anamnese of screenings-gesprek) beschikbaar is, actief wordt benut en dat de student de bevindingen zodanig vastlegt dat de overdracht naar de volgende zorgprofessional mogelijk is. Hoewel slechts een deel van het zorgproces wordt getoetst, kan een stationstoets op deze manier toch realistisch en representatief zijn, mits de volledige vaardigheid — zoals uitgewerkt in de vaardighedenhiërarchie — als beoordelingskader dient. De toets blijft dan niet beperkt tot een technische handeling, maar omvat ook klinisch redeneren, communicatie en contextuele afstemming. Zo levert de stationstoets wél een valide inschatting op van beroepsbekwaamheid.

Summatieve toetsing van hele taken sluit formatieve beoordeling van deeltaken of afzonderlijke vaardigheden niet uit. Het kan zelfs waardevol zijn om tussentijds na te gaan of bijvoorbeeld een technische handeling of gespreksvaardigheid voldoende beheerst wordt vóór deze binnen een complexere taak worden toegepast. Maar ook bij formatieve toetsing geldt: verlies het grotere geheel niet uit het oog. Vaardigheden worden in de praktijk zelden in isolatie toegepast. Afhankelijk van de casus en de situatie kunnen uitvoering en interpretatie sterk variëren (Jensen et al., 2022). Dat vraagt om ruimte voor functionele variatie, zelfs in de setting van een deeltaak. Bovendien laat onderzoek zien dat een te sterke focus op technische instructie en gedetailleerde feedback de transfer naar complexere, geïntegreerde situaties juist kan belemmeren (bijv. Masters, 1992; Van Brakel-Van Lobenstein et al., 2025). Als studenten overspoeld worden met technische beoordelingscriteria, verschuift hun aandacht naar het correct uitvoeren van details – ten koste van adaptief en professioneel handelen. Natuurlijk is duidelijke

instructie zinvol, zeker bij het aanleren van cruciale vaardigheden (zie bijvoorbeeld vaardigheidskompas.nl, De Haagse Hogeschool, 2024), maar overdaad schaadt. Technische instructies en feedback moeten ondersteunend zijn aan het beroepsmatige doel van de vaardigheid, bijvoorbeeld het stellen van een juiste diagnose, niet leidend.

Conclusie

Toetsing van beroepsvaardigheden vereist een doordachte afstemming tussen toetsvorm, beoordelingscriteria en het doel van de toets. Hoewel de toetsing van deeltaken en het gebruik van checklists een zekere objectiviteit bieden, schieten ze tekort bij het beoordelen van professioneel, geïntegreerd handelen in realistische situaties. In lijn met het 4C/ID-model worden vaardigheden idealiter onderwezen en getoetst in samenhang binnen hele, betekenisvolle taken. Een vaardighedenhiërarchie helpt daarbij: ze maakt zichtbaar hoe vaardigheden samenhangen en in welke context ze functioneel tot hun recht komen. Dit schept ruimte voor variatie, afstemming op de beroepspraktijk en een valide beoordeling van bekwaamheid. Ook bij toetsing van deeltaken, zoals bij stationstoetsen of bij formatieve beoordelingen, blijft het essentieel om het grotere geheel in zicht te houden. Alleen dan wordt duidelijk of studenten het geleerde weten toe te passen op een manier die recht doet aan de complexiteit van de praktijk. ■

Referenties

- Andersen, J. P., Pitel, M., Weerasinghe, A., & Papazoglou, K. (2016). Highly realistic scenario based training simulates the psychophysiology of real world use of force encounters: Implications for improved police officer performance. *Journal of Law Enforcement*, 5, 1–13.
- Button, C., Seifert, L., Chow, J. Y., Araújo, D., & Davids, K. (2021). *Dynamics of skill acquisition. An ecological dynamics approach*. Champaign, IL: Human Kinetics.
- De Haagse Hogeschool. (2024). *Vaardigheidskompas*. <https://vaardigheidskompas.nl/>
- Frèrejean, J., van Merriënboer, J. J. G., Kirschner, P. A., Roex, A., Aertgeerts, B., & Marcellis, M. (2019). Designing instruction for complex learning: 4C/ID in higher education. *European Journal of Education*, 54, 513–524. <https://doi.org/10.1111/ejed.12363>
- Frèrejean, J., Wopereis, I., Hoogveld, B., Janssen-Noordman, A., & Van Merriënboer, J. (2025). *Innovatief onderwijs ontwerpen. De ontwerpprincipes van het 4C/ID-model*. Groningen/Utrecht, Nederland: Noordhoff Uitgevers bv.
- Hutter, R. I., Renden, P. G., Kok, M., Oudejans, R., Koedijk, M., & Kleygrewe, L. (2023). *Criteria for the high quality training of police officers*. In: Staller, M.S., Koerner, S., Zaiser, B. (eds) *Police conflict management*, Volume II. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-41100-7_2
- Jensen, R. D., Brydges, R., & Grierson, L. (2022). Re-examining the integration of routine and adaptive expertise: there is no such thing as routine from a motor control perspective. *Advances in Health Sciences Education*, 27, 1283–1291. <https://doi.org/10.1007/s10459-022-10163-1>
- Koerner, S., & Staller, M. S. (2018). From system to pedagogy: Towards a nonlinear pedagogy of self-defense training in the police and the civilian domain. *Security Journal*, 31, 645–659. <https://doi.org/10.1057/s41284-017-0122-1>
- Koerner, S., & Staller, M. S. (2021). Police training revisited—Meeting the demands of conflict training in police with an alternative pedagogical approach. *Policing: A Journal of Policy and Practice*, 15(2), 927–938. <https://doi.org/10.1093/policing/paaa080>
- Masters, R. S. W. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in

the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, 83(3), 343-358. <https://doi.org/10.1111/j.2044-8295.1992.tb02446.x>

- Renden, P. G., & Dikken, J. (2023). Introducing the constraints-led approach in nurse education: An innovative perspective on skill acquisition. *Nurse education today*, 121, 105672. <https://doi.org/10.1016/j.nedt.2022.105672>
- Ten Cate, O., & Scheele, F. (2007). Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Academic Medicine*, 82, 542-547.
- Van Brakel-Van Lobenstein, N., van Leeuwen-Prins, S. T., Verdoes, L., de Waal, G. M., Kamsteeg, M. C. J., Oudejans, R. R. D., Devereaux, J., van Zaalen, Y., Dikken, J. & Renden, P. G. (2025). Teaching nursing skills without detailed protocols: Effects of an implicit learning strategy in nursing education. *Advances in Health Sciences Education*. <https://doi.org/10.1007/s10459-025-10421-y>
- Van der Vleuten, C. P. M., & Schuwirth, L. W. T. (2005). Assessing professional competence: from methods to programmes. *Medical Education*, 39(3), 309-317. <https://doi.org/10.1111/j.1365-2929.2005.02094.x>
- Van Merriënboer, J. J. G., Kirschner, P. A., & Frèrejean, J. (2025). *Ten steps to complex learning. a systematic approach to four-component instructional design* (4th edition). New York, NY: Routledge.
- Verdoes, L., Dikken, J., Van Brakel-Van Lobenstein, N., Van Leeuwen-Prins, S. T., De Waal, G. M., & Renden, P. G. (2025). Embracing a constraints-led approach for skills acquisition in nursing education: Insights from a focus group study. *International Journal of Educational Research Open*, 9, 100459. <https://doi.org/10.1016/j.ijedro.2025.100459>
- Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, 9(2), 185-211.

Gesignaleerd

Simone Paauw 27 maart 2025

Loopbaan buitenlands gediplomeerde artsen on hold door wachten op toets

Artsen die van buiten de Europese Unie naar Nederland komen, willen hier graag hun beroep als arts zo snel mogelijk oppakken. Maar de wachttijd voor de toets die dat mogelijk maakt, is met bijna twee jaar wel érg lang, vinden zij.

De loopbaan van ruim 150 buitenlands gediplomeerde artsen staat on hold en dat frustriert. Deze artsen hebben altijd al een lange adem nodig gehad, voordat zij zich konden inschrijven in het BIG-register. Ze moeten de Nederlandse en de Engelse taal voldoende beheersen en vakinhoudelijke toetsen afleggen. En daarna soms een periode coschappen lopen, waar zij – net als veel geneeskundestudenten – vaak op moeten wachten.

Lees verder op: <https://www.medischcontact.nl/actueel/laatste-nieuws/artikel/loopbaan-buitenlands-gediplomeerde-artsen-on-hold-door-wachten-op-toets>