



Universiteit
Leiden
The Netherlands

Leren van toetsen: hoe je met goede toetsvragen het leren kunt stimuleren

Wijk, E.V.; Janse, R.J.; Langers, A.M.J.

Citation

Wijk, E. V., Janse, R. J., & Langers, A. M. J. (2024). Leren van toetsen: hoe je met goede toetsvragen het leren kunt stimuleren. *Nederlands Tijdschrift Voor Geneeskunde*, 168. Retrieved from <https://hdl.handle.net/1887/4103774>

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/4103774>

Note: To cite this publication please use the final published version (if applicable).

DISCLAIMER



Onafhankelijke informatie is niet gratis. Het NTvG investeert veel geld om het hoge niveau van haar artikelen te waarborgen, door een proces van peer-review en redactievoering. Het NTvG kan alleen bestaan als er voldoende betaalde abonnementen zijn. Het is niet de bedoeling dat onze artikelen worden verspreid zonder betaling. Wij rekenen op uw medewerking.

Leren van toetsen

Hoe je met goede toetsvragen het leren kunt stimuleren

Elise V. van Wijk, Roemer J. Janse en Alexandra M.J. Langers

Samenvatting

Veel zorgverleners zijn betrokken bij het opleiden van toekomstige zorgverleners. Het ontwikkelen van toetsen is een belangrijk onderdeel van het opleiden, maar het maken van goede toetsvragen is nog niet zo eenvoudig. In dit artikel worden praktische tips gegeven voor het opstellen van goede toetsvragen. Daarnaast wordt de veelgebruikte meerkeuzevraag ter discussie gesteld en wordt er een veelbelovend alternatief geïntroduceerd: de Very Short Answer Question (VSAQ). Deze open vraag met een gelimiteerd antwoord van 1-4 woorden kan de kwaliteit van toetsen aanzienlijk verbeteren en het leren stimuleren. De wetenschappelijk onderbouwde voordelen worden besproken en mogelijke obstakels om VSAQ's toe te passen worden weerlegd, zodat niets meer in de weg staat om meer VSAQ's op te nemen in toetsen.

Iedereen herinnert zich de tentamens tijdens de studie. Ook in de latere loopbaan moeten er nog de nodige toetsen en examens gemaakt worden, die zowel kennis als praktische vaardigheden meten. In dit artikel ligt de focus op kennistoetsen.

Veel van deze kennistoetsen zijn bedoeld om te meten of er voldoende kennis is opgedaan over een onderwerp, waarbij het kennisniveau veelal beoordeeld wordt in de vorm van een cijfer, certificeringen of studiepunten. Dit noemen we summatief toetsen of 'assessment of learning'.

Een toets kan echter ook als doel hebben om van te leren of het leerproces te sturen, vaak door het gebruik van persoonlijke feedback. Deze vorm van toetsen wordt formatief toetsen of 'assessment for learning' genoemd. Voor een effectieve formatieve functie is het van belang dat de nadruk van de feedback ligt op het begeleiden en niet op het summatief beoordelen, zodat de leerling uit de 'examenstand' in de 'leerstand' komt.¹

Daarnaast kunnen toetsen ook ingezet worden tijdens het leerproces om te zorgen dat kennis langer beklijft middels 'retrieval practice': het actief terughalen van kennis uit het geheugen en het opnieuw toepassen daarvan. Uit onderzoek blijkt dat herhaaldelijk toetsen – een vorm van retrieval practice – zorgt voor een significant hogere kennisretentie op de lange termijn bij studenten in vergelijking met studenten die evenveel tijd besteden aan het bestuderen van een onderwerp zonder getoetst te worden.²

Deze verschillende vormen van toetsen komen aan bod in de bachelor-, master- en vervolgopleiding. Het is echter ook na deze periode nog belangrijk om jezelf te blijven ontwikkelen en te blijven leren. Leren gaat immers je leven lang door. De artikelen voor nascholing en opleiding die het NTvG op zijn website aanbiedt, kunnen daar bijvoorbeeld goed in ondersteunen en ervoor zorgen dat kennis up-to-date blijft.

Betrouwbare en valide toetsen

Veel zorgverleners zijn betrokken bij het opleiden van de zorgverleners van de toekomst. Een onderdeel van het opleiden is het ontwikkelen van toetsen. Een goede toets draagt op een betrouwbare manier bij aan meten van de leerdoelen en stimuleert het leerproces. Het ontwikkelen van een toets begint bij het bedenken en opstellen van vragen die de betreffende leerdoelen op een betrouwbare en valide manier toetsen. Een betrouwbare toets betekent dat de toets consistent en nauwkeurig meet en bij herhaaldelijke afname vergelijkbare resultaten oplevert. Een toets is valide wanneer deze in overeenstemming is met de leerdoelen en daadwerkelijk toetst wat getoetst moet worden. Hierbij is het dus van belang dat leerdoelen ondersteund worden door passende leermiddelen – leeractiviteiten – en dat de toets aansluit bij het leerdoel en de leermiddelen. Dit wordt ook wel 'constructive alignment' genoemd.³

In de tabellen zijn nuttige tips te vinden voor het schrijven van goede toetsvragen. Naast algemene tips die van toepassing zijn op alle toetsvragen (tabel 1), zijn hier ook tips te vinden die specifiek betrekking hebben op meerkeuzevragen (tabel 2).

Tip	Toelichting
De leerdoelen bepalen welke (soort) vragen worden gesteld.	
De vraag toetst de toepassing van praktisch relevante kennis.	
Korte casusbeschrijvingen ('klinische vignetten') vormen een goede basis voor een vraag.	Begin met het probleem van de patiënt, gevolgd door relevante voorgeschiedenis, bevindingen van het lichamelijk onderzoek, eventueel diagnostische resultaten, eerdere behandelingen en andere bevindingen.
De vraag bevat één duidelijk probleem of vraag.	
De vraag is correct, helder en bondig geformuleerd.	
De vraag bevat geen ontkenning.	
Vragen zoals: 'Wat is geen oorzaak van...' zijn niet toegestaan.	Dit kan de lezer in verwarring brengen. Hetzelfde geldt voor ontkenningen in antwoordopties van meerkeuzevragen.
De vraag bevat geen absolute (nooit, altijd) of vage (soms, vaak) woorden.	
Zelden is 'altijd' of 'nooit' van toepassing. Bij meerkeuzevragen zijn antwoordopties met dit soort termen erin meestal fout.	
De vraag is beoordeeld door een beoordelingscommissie of collega, het liefst met een andere achtergrond.	
Er wordt een psychometrische analyse van de vragen na de toets in de digitale toetsomgeving uitgevoerd.	
Een psychometrische analyse geeft waardevolle informatie over de kwaliteit van de vragen, zodat deze zo nodig verbeterd kunnen worden voor een volgend afnamemoment.	Bijvoorbeeld: informatie over frequent gegeven foute antwoorden geeft inzicht in het kennisniveau van de groep en eventuele misconcepties. Als studenten die een toets gemiddeld genomen goed maken, bij een bepaalde vraag in groten getale een verkeerd antwoord geven, is er mogelijk sprake van een sleutelfout (verkeerd antwoord in antwoordmodel) of onduidelijkheid in de vraag.

Tabel 1
Algemene tips voor het schrijven van een goede toetsvraag

Meerkeuzevragen

De meerkeuzevraag is een veelgebruikte vraag binnen het medisch onderwijs. De vraag is breed inzetbaar, heeft een hoge betrouwbaarheid en is makkelijk na te kijken voor grote groepen deelnemers.⁴ In de literatuur is er echter toenemend bewijs dat de meerkeuzevraag onderhevig is aan een aantal beperkingen. Zo kunnen studenten gebruikmaken van aanwijzingen in de antwoordopties om tot het juiste antwoord te komen ('cueing') zonder inhoudelijke kennis nodig te hebben.⁵ Dit wordt mede gefaciliteerd doordat het verzinnen van plausibele antwoordopties één van de grootste uitdagingen is bij het maken van een meerkeuzevraag.

Tip	Toelichting
De vraag moet voldoen aan de afdekproef.	De vraag moet nog steeds beantwoord kunnen worden als de antwoordopties zijn afgedekt. Zo wordt de student aangemoedigd na het lezen van de vraag eerst na te denken over het juiste antwoord, alvorens naar de opties te kijken ('forward reasoning'). De lead-in moet dus een echte vraag zijn, en <i>niet</i> : 'welke van onderstaande antwoorden is juist?' of 'deze ziekte is een voorbeeld van: ...'. Als gebruik wordt gemaakt van een plaatje of grafiek, let er dan op dat deze ook echt nodig is om de vraag te kunnen beantwoorden. Hetzelfde geldt ook voor het inleidende tekstje ('de stam') van de vraag.
Laat de antwoordopties overeenkomen met de vraag, zowel inhoudelijk als grammaticaal.	Als er in de vraag enkelvoud wordt gebruikt, moet dit ook gehandhaafd worden in de antwoordopties. Daarnaast moeten de antwoordopties inhoudelijk mogelijk zijn bij de vraag, dus bij vragen als 'welke intraveneuze injectie' niet de optie 'CT-scan' geven.
Zorg dat maar één antwoordoptie correct of duidelijk de beste optie is.	
De afleiders – foute antwoordopties – moeten redelijk zijn, gebaseerd op veelvoorkomende kennishiaten of misconcepties.	
De antwoordopties mogen maar één onderwerp of probleem bevatten, net als de vraag.	Probeer de volgende stellingen te voorkomen: 'doe X en Y'; 'X is verhoogd, Y is verlaagd'; 'doe X omdat Y'.
De antwoordopties zijn homogeen, dat wil zeggen: vergelijkbaar in inhoud, lengte en formulering.	
Rangschik de antwoordopties in logische volgorde: alfabetisch of numeriek.	
Zorg dat de antwoordopties genoeg van elkaar verschillen: er mag geen sprake zijn van overlap.	
Gebruik geen zinnen als 'geen van bovenstaande alternatieven' of 'alle bovenstaande alternatieven'.	

Tabel 2

Tips voor het maken van meerkeuzevragen

Door de aanwezigheid van plausibele foutieve antwoordopties bij een meerkeuzevraag worden deelnemers ook opzettelijk blootgesteld aan misinformatie. Dit vergroot de kans op het herinneren van onjuiste informatie op een later moment ('misinformation effect') als er niet direct feedback wordt gegeven – wat in de praktijk vaak het geval is.

Tot slot stimuleert dit vraagtype een leerstrategie die gericht is op herkenning.⁶ Deze leerstrategie sluit niet goed aan bij de klinische praktijk, waarin (jonge) dokters de geleerde stof moeten kunnen toepassen en bijna nooit de mogelijkheid hebben te kiezen uit een lijst met mogelijke antwoorden. Kortom, het wordt hoog tijd dat we wat vaker afstappen van de meerkeuzevraag en ons openstellen voor andere vraagtypen.

De 'Very Short Answer Question'

De transitie naar digitaal toetsen is in een stroomversnelling geraakt door de COVID-19-pandemie en daarbij is het gebruik van andere vraagtypen naast de klassieke meerkeuzevraag, zoals open vragen, een stuk makkelijker geworden. Zo is er meer mogelijkheid gekomen voor het gebruik van een vraagtype dat goed geschikt is voor toetsing binnen het medisch onderwijs: de 'Very Short Answer Question' (VSAQ). Dit is een open vraag waarop een kort antwoord van 1-4 woorden moet worden gegeven, soms ook 'invulvraag' genoemd.

Dit vraagtype is anders dan een open vraag waar geen, of een hogere, woordenlimiet op zit (tabel 3). In dit vraagtype is cueing niet mogelijk en correct gokken een stuk lastiger, wat beter onderscheidt tussen studenten die de stof wel en niet beheersen.^{7,8}

Daarnaast heeft de VSAQ ook een hogere betrouwbaarheid dan de meerkeuzevraag.^{8,9}

VSAQ	Meerkeuzevraag
Een 65-jarige man presenteert zich met frequente waterdunne diarree, gepaard gaande met koorts en buikpijn. Hij is bekend met COPD Gold-stadium III en maakte recent twee maal een pneumonie door, die beide keren antibiotisch werd behandeld. In afwachting van het fecesonderzoek wordt een sigmoidoscopie verricht, waarbij het typische beeld van een pseudomembraneuze colitis wordt gezien. Wat is de meest waarschijnlijke verwekker van dit ziektebeeld?	Een 65-jarige man presenteert zich met frequente waterdunne diarree, gepaard gaande met koorts en buikpijn. Hij is bekend met COPD Gold-stadium III en maakte recent twee maal een pneumonie door, die beide keren antibiotisch werd behandeld. In afwachting van het fecesonderzoek wordt een sigmoidoscopie verricht, waarbij het typische beeld van een pseudomembraneuze colitis wordt gezien. Wat is de meest waarschijnlijke verwekker van dit ziektebeeld? A. <i>Campylobacter jejuni</i> B. <i>Clostridium difficile</i> C. <i>Giardia lamblia</i> D. <i>Staphylococcus aureus</i>

Tabel 3
Voorbeeld VSAQ en corresponderende meerkeuzevraag

Een groot deel van studenten geeft aan dat ze op een andere manier gaan studeren als ze weten dat ze met VSAQ's getoetst zullen worden, waardoor er minder focus zal komen op het louter herkennen van de leerstof.^{8,9} Deze actievorm van studeren zal waarschijnlijk ook leiden tot een beter behoud van de opgedane kennis op de lange termijn. Daarnaast resulteren oefentoetsen ('retrieval practice') met VSAQ's tot een beter behoud van kennis. Dit komt doordat er voor het beantwoorden van deze vragen meer inspanning nodig is om de stof terug te halen uit het geheugen ('retrieval effort') dan bij meerkeuzevragen.² Tot slot sluit dit vraagtype ook beter aan bij de klinische praktijk dan de meerkeuzevraag.

Wat staat toepassing van VSAQ's nog in de weg?

Qua gebruik leent de VSAQ zich goed voor het stellen van kennisvragen en klinische redeneervragen, die worden ingeleid met een korte klinische casusbeschrijving. Een veelgehoord bezwaar onder docenten tegen dit vraagtype is dat het nakijken langer zou duren. Meerdere onderzoeken tonen aan dat het nakijken van VSAQ's efficiënt en relatief snel kan worden gedaan in digitale toetssystemen, met gemiddeld genomen twee minuten nakijktijd per vraag voor een volledig cohort van rond de 350 geneeskundestudenten.^{8,9}

Inmiddels wordt in alle Nederlandse geneeskundefaculteiten gebruikgemaakt van een digitaal toetssysteem dat de mogelijkheid biedt om de correcte antwoorden van tevoren in te voeren, zodat deze direct worden goed gerekend. Zo hoeven alleen de incorrecte antwoorden nog te worden beoordeeld. Mochten daar nog correcte antwoorden tussen zitten, dan kan dit antwoord bij één student alsnog als correct worden aangemerkt in het toetssysteem, waarna dit betreffende antwoord direct wordt goed gerekend bij alle studenten die hetzelfde antwoord hebben gegeven. Zo gaat het nakijkproces steeds sneller en efficiënter en is de extra nakijktijd dus ook zeker geen reden om af te zien van het toepassen van dit vraagtype. Bovendien hoeven er voor dit vraagtype geen plausibele afleiders (foute antwoordopties) verzonnen te worden, waardoor er flink wat tijd bespaard wordt bij het maken van de vraag. Er lijken dus weinig obstakels te zijn om meer VSAQ's te gebruiken in toetsen. Tegelijkertijd kunnen VSAQ's de kwaliteit van toetsen en daarbij het onderwijs aanzienlijk verbeteren. Waarom worden VSAQ's dan toch nog weinig gebruikt? Een van de redenen hiervoor kan de onbekendheid zijn met dit vraagtype en met de mogelijkheden van de nieuwe digitale toetssystemen. Onbekend maakt vaak onbeminde. Verder gaat verandering of vernieuwing vaak gepaard met weerstand. Dit zal zowel bij de docent die de vragen moet maken als bij de deelnemer die de vragen moet beantwoorden het geval zijn.

Een van de belangrijkste bezwaren tegen VSAQ's, de extra nakijktijd, is al weggenomen, zoals we hiervoor reeds hebben uitgelegd. Daarnaast vormt een korte workshop met uitleg en de mogelijkheid om te oefenen een goede voorbereiding voor docenten, en is ondersteuning vanuit de faculteit belangrijk.⁹

Uiteraard is het ook belangrijk om deelnemers goed te informeren en voor te bereiden alvorens een nieuw vraagtype te introduceren. Studenten ervaren de VSAQ als moeilijker dan de meerkeuzevraag,^{8,9} maar uiteindelijk zullen ze er veel meer profijt van hebben. Door goede voorlichting en het aanbieden van oefenmogelijkheden, kan er iets van de weerstand worden weggenomen (tabel 4). Ook zijn er in de literatuur handige tips te vinden over hoe je VSAQ's het beste in het curriculum kunt implementeren.¹⁰

Tip	Toelichting
De vragen zijn idealiter nieuw en geen omgeschreven bestaande meerkeuzevragen.	Om tijd te besparen en goede meerkeuzevragen te hergebruiken, kunnen deze worden omgeschreven door de antwoordopties te elimineren. Herschreven vragen blijken echter toch minder specifiek te zijn, met een grotere variatie aan antwoorden als gevolg. ¹⁰ Hoewel het maken van nieuwe vragen eerst meer tijd zal kosten, zorgt het uiteindelijk voor tijdswinst bij het nakijken en – nog veel belangrijker – vragen van betere kwaliteit.
De vraagstelling is specifiek genoeg.	Aangezien het antwoord kort moet zijn, is het essentieel om zo specifiek mogelijk in de vraagstelling aan te geven wat voor soort antwoord verwacht wordt. Als de vraag niet specifiek genoeg is, kan een deelnemer die de kennis wel heeft de vraag toch verkeerd beantwoorden. De vraag 'Wat is de behandeling van eerste keus?' bij een patiënt met hypertensie is te breed gesteld. Het is onduidelijk of leefregels of medicatie wordt bedoeld. Dit kan specifieker worden gemaakt door bijvoorbeeld te vragen naar de 'meest aangewezen medicamenteuze behandeling'. ⁹
Docenten en toetsdeelnemers zijn goed voorbereid op het gebruik van VASQ's.	Verminder ongerustheid bij de deelnemer door inzicht te geven in het nakijkproces. Maak duidelijk dat alle antwoorden worden nagekeken en dat de student niet wordt afgerekend op spelfouten.. Laat deelnemers ook oefenen met VSAQ's. Door het oefenen met het vraagtype krijgen de deelnemers en docenten ook gelijk inzicht in hun kennishiaten.
De nakijkstrategie is duidelijk besproken om het nakijkproces efficiënt te laten verlopen.	

Tabel 4
Tips voor het maken en implementeren van VSAQ's

Conclusie

Toetsing is een belangrijk en terugkerend onderdeel in de loopbaan van zorgverleners en kan verschillende doelen hebben, afhankelijk van de context. Gezien de impact van toetsing op het leerproces is het van groot belang dat de vragen in de toets van goede kwaliteit zijn, aansluiten bij de dagelijkse praktijk, aanzetten tot een actieve manier van studeren en motiveren om de leerstof goed te doorgronden. Het wordt tijd dat we uit onze comfortzone stappen, de meerkeuzevraag durven los te laten en meer VSAQ's opnemen in toetsen. Op basis van de argumenten in dit artikel en de bijgevoegde tips zijn er geen redenen meer om dit niet te doen. Zo kunnen we met goede, representatieve toetsvragen het leren stimuleren, in plaats van het leren te frustreren.

- Online artikel en reageren op ntvg.nl/D8012
- Leiden UMC, afd. Onderwijs Expertisecentrum, Leiden: drs. E.V. van Wijk, promovendus; afd. Klinische Epidemiologie: R.J. Janse, promovendus; afd. Maag- Darm- Leverziekten: dr. A.M.J. Langers.
- Contact: E.V. van Wijk (e.v.van_wijk@lumc.nl)
- Belangenconflict en financiële ondersteuning: geen gemeld.
- Aanvaard op 27 november 2023
- Citeer als: Ned Tijdschr Geneesk. 2024;168:D8012

Literatuur

1. Brand PLP, Scheele F. [Feedback in de medische opleiding](#). Ned Tijdschr Geneesk. 2022;166:D6708
2. Butler AC, Roediger HL III. Testing improves long-term retention in a simulated classroom setting. Eur J Cogn Psychol. 2007;19:514-27. [doi:10.1080/09541440701326097](https://doi.org/10.1080/09541440701326097).
3. Biggs J. Enhancing teaching through constructive alignment. High Educ. 1996;32:347-64. [doi:10.1007/BF00138871](https://doi.org/10.1007/BF00138871).
4. Al-Rukban MO. Guidelines for the construction of multiple choice questions tests. J Family Community Med. 2006;13:125-33. [doi:10.4103/2230-8229.97543](https://doi.org/10.4103/2230-8229.97543). [Medline](#)
5. Al-Kadri HM, Al-Moamary MS, Roberts C, Van der Vleuten CPM. Exploring assessment factors contributing to students' study strategies: Literature review. Med Teach. 2012;34(Suppl 1):S42-50. [Medline](#)
6. Veloski JJ, Rabinowitz HK, Robeson MR, Young PR. Patients don't present with five choices: an alternative to multiple-choice tests in assessing physicians' competence. Acad Med. 1999;74:539-46. [doi:10.1097/00001888-199905000-00022](https://doi.org/10.1097/00001888-199905000-00022). [Medline](#)

7. Schuwirth LWT, van der Vleuten CP, Donkers HHLM. A closer look at cueing effects in multiple-choice questions. Med Educ. 1996;30:44-9. [doi:10.1111/j.1365-2923.1996.tb00716.x](https://doi.org/10.1111/j.1365-2923.1996.tb00716.x). [Medline](#)
8. Sam AH, Field SM, Collares CF, et al. Very-short-answer questions: reliability, discrimination and acceptability. Med Educ. 2018;52:447-55. [doi:10.1111/medu.13504](https://doi.org/10.1111/medu.13504). [Medline](#)
9. van Wijk EV, Janse RJ, Ruijter BN, et al. Use of very short answer questions compared to multiple choice questions in undergraduate medical students: An external validation study. PLoS One. 2023;18:e0288558. [doi:10.1371/journal.pone.0288558](https://doi.org/10.1371/journal.pone.0288558). [Medline](#)
10. Bala L, Westacott RJ, Brown C, Sam AH. Twelve tips for introducing very short answer questions (VSAQs) into your medical curriculum. Med Teach. 2022; 1-8 [Medline](#).