



Opgesteld door: Leonie Brummer (antwoordspecialist) en Sjerp van der Ploeg (Kennismakelaar Kennisrotonde)
Vraagsteller: onderwijsstichting - directeur-bestuurder

Referentie: Kennisrotonde. (2018). *Op welke manieren draagt toetsing maximaal bij aan leerwinst in het voortgezet onderwijs?* (KR. 419). Den Haag: Kennisrotonde.

29 augustus 2018

Vraag

Op welke manieren draagt toetsing maximaal bij aan leerwinst in het voortgezet onderwijs?

Kort antwoord

Toetsen draagt het meeste bij aan leerprestaties en leerwinst als de nadruk ligt op formatief toetsen (in tegenstelling tot summatief toetsen). Dat betekent toetsen gericht op het in kaart brengen van de voortgang van het leerproces en de (informatierijke) feedback die daarop wordt gebaseerd. Verder blijkt voor leerprestaties relevant dat verschillende soort vragen in een formatieve toets worden gecombineerd en dat er frequent een beperkte hoeveelheid leerstof wordt getoetst (waarbij er sprake is van een niet universeel vast te stellen plafond in het aantal toetsmomenten per week). Het verschilt per vak of digitale dan wel schriftelijke toetsen beter zijn voor leerprestaties.

Toelichting antwoord

Toetsing vormt een essentieel onderdeel van ons onderwijs en is voornamelijk vakgericht ingedeeld (Joosten-ten Brinke, 2011). Het doel is om te bepalen hoe een leerling ervoor staat en/of het onderwijs aan een leerling toe te spitsen op specifieke behoeften. VO-scholen bepalen veelal aan het begin van een schooljaar welke toetsingsmomenten worden ingepland. Vakken in het voortgezet onderwijs kennen voornamelijk een leerproces waarbij leerlingen een paar weken met de lesstof bezig zijn en daarna getoetst worden (Van den Berg, 2016). In de literatuur komt een aantal kenmerken van toetsing steeds terug. Die kenmerken gebruiken we hier als ordening voor het

themagewijs presenteren van ons antwoord. We onderscheiden de volgende thema's: toetsingsaard, toetsingsfrequentie en toetsingsvorm.

Toetsingsaard

Bij toetsen wordt standaard het onderscheid gemaakt tussen summatieve en formatieve toetsen. Summatieve toetsen zijn gericht op het afsluiten van een leerperiode en op het beoordelen van leerprestaties, veelal in de vorm van een cijfer. Het zijn momentopnames en de uitkomst van de toets wordt bijvoorbeeld gebruikt voor toelating tot vervolgonderwijs. Formatieve toetsen zijn gericht op het in kaart brengen van de voortgang van het leerproces. De uitkomsten zijn bedoeld als input voor feedback voor zowel leerling als docent. Het levert voor de docent diagnostische informatie om het instructie- en leerproces bij te sturen en te verbeteren en het onderwijs toe te spitsen op de behoeften van de individuele leerling (Schildkamp e.a. 2104; Van den Berg, 2016).

Formatieve toetsen geven daarnaast leerlingen inzicht in hun eigen leerproces en bevordert hun motivatie aan (Kennisrotonde, 2017; Sluijsmans et al., 2013).

Summatief toetsen heeft als doel leerprestaties te *meten*. Mogelijk levert een puur summatieve toets een bijdrage aan leerprestaties wanneer er voor de leerling veel van die toets afhangt (bijvoorbeeld een schoolexamen). De leerling zou dan erg gemotiveerd zijn om de leerstof te beheersen om daarmee het gewenste resultaat te boeken. Onderzoek naar dit mogelijk effect is tamelijk beperkt en richt zich vooral op de zogenaamde high stake toetsen in het kader van de No Child Left Behind Act in de VS waarin niet zozeer voor de leerlingen maar vooral voor de scholen en leraren veel afhangt van de toetsuitkomsten. Daaruit blijkt dat high stake (summatief) toetsen niet te leiden tot een hogere motivatie en niet tot betere leerprestaties (Amrein & Berliner, 2003).

Formatief toetsen heeft als doel de leerprestaties te *verbeteren* (oftewel leerwinst te boeken). Dat betekent dat we hier dus vooral op literatuur rondom formatief toetsen focussen. Toetsen die in eerste instantie ontwikkeld zijn om summatief te toetsen kunnen overigens ook effectief gebruikt worden om leren te stimuleren (Ayala, et al. 2008; Bennett, 2011; Black & Wiliam, 1998, 2009; Hattie & Timperley, 2007; Lee, 2006; Nicol & Macfarlane-Dick, 2006; Stiggins, 2005; Swaffield, 2011; Taras, 2007; 2009). Doordat deze toetsen ontwikkeld zijn om aan te sluiten bij de standaarden, maakt het formatief gebruik ervan meteen duidelijk welke standaarden nog niet beheerst worden of welke leerlingen achter blijven in hun ontwikkeling (Stiggins, 2005).

Over effectieve formatieve toetsing is het nodige bekend (bijv. Kennisrotonde, 2017). Formatieve toetsing kenmerkt zich door

- a) rijke feedback die leerlingen helpt beter te leren,
- b) feedback die wordt ingezet in verschillende fasen van het leerproces en
- c) stimuleert tot zelfregulatie en actieve betrokkenheid (Sluijsmans et al., 2013).

We bespreken hieronder kort hoe deze drie kenmerken kunnen bijdragen aan hogere motivatie en betere leerprestaties van leerlingen (ontleend aan Kennisrotonde, 2017).

Rijke feedback die helpt beter te leren

Feedback is een cruciaal onderdeel van formatief toetsen en het meest effectief wanneer ze kort en krachtig is, gebaseerd op heldere criteria en verbeterpunten (Schildkamp et al., 2014). In hun reviewstudie presenteren Sluijsmans et al. (2013) een compleet overzicht met tips voor goede feedback. Voorbeelden zijn: 1) richt de feedback op de taak en niet op de leerling, 2) geef de feedback in overzichtelijke eenheden om cognitieve overbelasting te voorkomen, en 3) reduceer onduidelijkheid tussen de prestatie en de leerdoelen. Samenvattend kan rijke feedback bijdragen aan een hogere motivatie en betere leerprestaties, doordat voor leerlingen duidelijker is wat de leerdoelen zijn, wat ze wel en (nog) niet goed doen, en wat ze kunnen doen om wel op het gewenste niveau te komen.

Feedback die wordt ingezet in verschillende fases van het leerproces

Het is belangrijk dat feedback niet alleen na prestaties wordt gegeven, maar juist ook bij aanvang en tijdens het leerproces. Dit sluit aan bij het cyclische karakter van het leerproces, en bevordert de (leer)doelgerichtheid. In de literatuur worden drie typen feedback onderscheiden: feed-up, feedback en feed-forward:

1. Feed-up: Waar gaat de leerling naartoe? Oftewel: wat zijn de beoordelingscriteria en standaarden?
2. Feedback: Hoe heeft de leerling de taak uitgevoerd? Oftewel: welke vooruitgang wordt geboekt ten aanzien van de beoordelingscriteria en standaarden?
3. Feed-forward: Hoe gaat de leerling verder? Oftewel: welke aanpak is nodig om tot groei te komen?

Stimuleren van zelfregulatie en actieve betrokkenheid

Formatief evalueren vraagt om en stimuleert een actieve rol van leerlingen, wat kan leiden tot hogere motivatie en betere leerprestaties. Bij formatieve evaluatie zijn waarden als vertrouwen in jezelf, het erkennen van verschillen tussen leerlingen en motivatie het uitgangspunt. Dit sluit bijvoorbeeld aan bij de zelfdeterminatietheorie van Ryan & Deci (2000), waarin competentie, autonomie en sociale verbondenheid basisbehoeften zijn en bijdragen aan welbevinden.

Toetsingsfrequentie

Er is veel onderzoek gedaan naar de toetsfrequentie (e.g. Adescope, Trecisan & Sundararajan, 2017; Adkins & Linville, 2017), voornamelijk in summatieve toets situaties. De toetsfrequentie hangt nauw samen met de toetsomvang. Frequenter toetsen betekent minder leerstof per toets en dat wil zeggen dat er minder te leren valt voor een toets (Bangert-Drowns, et al., 1991, zoals geciteerd in Adkins & Linville, 2017). Tevens blijken leerlingen zich beter voor te bereiden op de toets als zij frequenter worden getoetst en krijgen leerlingen meer feedback krijgen als er meer toetsmomenten zijn (Mines, 2014). Uitkomsten van onderzoek naar toetsfrequentie zijn niet eenduidig; er wordt zowel positieve als negatieve als geen invloed op leerprestaties gerapporteerd.

Eenzijds lijkt frequenter toetsen een gunstige invloed te hebben op leerprestaties (zie Kling et al., 2005; Leeming, 2002, beide geciteerd in Adkins & Linville, 2017).

Vermoedelijk spelen hierbij twee mechanismen een rol. Ten eerste wordt frequent toetsen gezien als een middel voor leerlingen met uitstelgedrag (procrastinerende leerlingen) om alsnog meer te leren. De toetsen dienen hierbij gericht te zijn op een klein aantal leerconcepten (Cid de Orta, Cabrera & Bernatzky, 2018). Ten tweede zorgt bij frequenter toetsen ervoor dat resultaten snel teruggekoppeld worden om voortgang weer te geven evenals de gebieden die nog meer aandacht behoeven (Cid de Orta, Cabrera & Bernatzky, 2018).

Anderzijds blijken er ook onderzoeken waarin een minder duidelijke samenhang bestaat tussen toetsingsfrequentie en leerprestaties (Mines, 2014). Cid de Orta e.a. (2018) vonden een positieve invloed van toetsingsfrequentie op leerprestaties, maar ze vonden geen verschil in scores tussen wekelijks en tweewekelijks toetsen. Vermoedelijk is er sprake van een plafondeffect met betrekking tot de toetsingsfrequentie. Op een gegeven moment is het zinloos om nóg meer te toetsen.

Toetsingsvorm

Toetsen kunnen verschillende typen vragen bevatten, zoals meerkeuzevragen of open vragen. De meta-analyse van Adesope, Trevisan en Sundararajan (2017) liet zien dat gemengde toetsen, dus verschillende soorten vragen, een positieve invloed hebben op de leerresultaten en daarmee bijdragen aan leerwinst. Met oog op het leerproces is het belangrijk dat leerlingen een meerkeuzevraag beantwoorden en aanvullen met argumenten. Hiermee krijgt een meerkeuzevraag een formatief karakter. Voorbeeldvragen zijn: Waarom heb je dit antwoord gekozen? Waar in de tekst vind je verwijzingen voor dit antwoord?

Verder kunnen toetsen op verschillende manieren worden afgenomen: bijvoorbeeld digitaal, schriftelijk en mondeling. Een deel van de literatuur meldt dat er geen verschillen bestaan in toetsresultaten tussen digitale en schriftelijke toetsen (zie bijvoorbeeld Boo & Vispoel, 2012; Mojarrad, Hemmati, JafariGohar, & Sadeghi, 2013; Tsai & Shin, 2013). Een meta-analyse van Kingston (2009) laat echter zien dat het effect van digitaal toetsen op de leerprestaties afhankelijk was van het vak. Het bleek dat digitale toetsresultaten hoger uitvielen voor Engels en *Social studies* (bijvoorbeeld geschiedenis, aardrijkskunde) dan voor Wiskunde. Voor Wiskunde vielen de toetsresultaten hoger uit voor schriftelijke toetsen (Kingston, 2009, zoals geciteerd in Joosten-Ten Brinke, Weber & Ljubin Saveski, 2016). Een digitale toetswijze heeft verder voordelen omdat het efficiënt(er) is, bijdraagt aan de leesbaarheid en snellere correctie voor docenten en in sommige gevallen mogelijkheden geeft om maatwerk te bieden aan leerlingen (Pino-Silva, 2008).

Er is nog weinig vergelijkend onderzoek gedaan naar adaptief en lineaire (niet-adaptieve) digitale toetsen. Volgens Joosten-Ten Brinke, Joosten-Ten Brinke, Weber en Ljubin Saveski (2016) is het van belang dat hier meer onderzoek naar gedaan wordt zodat er duidelijkere conclusies getrokken kunnen worden over de invloed van adaptiviteit op leerprestatie en leerwinst.

Een manier om digitale toetsen in te zetten zijn stemkastjes of mobiele telefoons, waarop leerlingen hun antwoorden op meerkeuzevragen kunnen doorgeven (Matuk et al., 2015). De inzet hier kan summatief of formatief van aard zijn. Echter wordt aangeraden om deze middelen eerder formatief in te zetten dan summatief vanwege verhoogde angstgevoelens bij summatieve toetsen. De afnamevorm heeft invloed op de motivatie en angstgevoelens en daarmee kan het indirect invloed hebben op de toetsresultaten (Joosten-Ten Brinke, Weber & Ljubin Saveski, 2016). Het is natuurlijk zo dat

angstgevoelens kunnen opspelen bij zowel digitale als schriftelijk toetsen. Bij enkele toetsingswijzen worden de resultaten echter direct zichtbaar zonder volledig de anonimiteit van de leerling te waarborgen (denk bijvoorbeeld aan Kahoot), wat kan leiden tot verhoogde angstgevoelens.

Daarentegen draagt het inzetten van deze middelen positief bij aan de betrokkenheid en motivatie van leerlingen. Het is mogelijk dat deze betrokkenheid en motivatie ook bijdraagt aan leerprestaties. Het blijft nog onduidelijk of deze betrokkenheid en motivatie verklaard kunnen worden door (a) de *game-like* setting, (b) de nieuwheid van de middelen of (c) een verhoogde cognitieve betrokkenheid (Kay & Knaack, 2009). Leerlingen geven aan een voorkeur te hebben voor digitaal toetsen, maar hun motivatie wordt niet beïnvloed door de manier waarop resultaten teruggekoppeld worden (Joosten-Ten Brinke, Weber & Ljubin Saveski, 2016).

Samenvattend

Op grond van bovenstaande concluderen we dat toetsen het meeste bijdraagt aan leerprestaties en leerwinst als de nadruk ligt op formatief toetsen. Dat betekent toetsen gericht op het in kaart brengen van de voortgang van het leerproces en de (informatierijke) feedback die daarop wordt gebaseerd. Verder blijkt voor leerprestaties relevant dat verschillende soort vragen in een formatieve toets worden gecombineerd en dat er frequent een beperkte hoeveelheid leerstof wordt getoetst (waarbij er sprake is van een niet universeel vast te stellen plafond in het aantal toetsmomenten per week). Het verschilt per vak of digitale dan wel schriftelijke toetsen beter zijn voor de leerprestaties.

Geraadpleegde bronnen

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, 87(3), 659-701. doi: 10.3102/0034654316689306

Adkins, J. K., & Linville, D. (2017). Testing frequency in an introductory computer programming course. *Information Systems Education Journal*, 15(3), 22.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting students achievement with the Classroom Assessment Scoring System. *School Psychology Review*, 42(1), 76-98.

Amrein, A. L., Berliner, D. C. (2003) The Effects of High-Stakes Testing on Student Motivation and Learning. *Educational Leadership*, 60 (5) 32-38.

Berg, E. van den (2016). Formatieve toetsing met onmiddellijke feedback. In D. Sluismans en R. Kneyber (Red.), *Toetsrevolutie. Naar een feedbackcultuur in het voortgezet onderwijs* (pp. 17-34). Culemborg, Nederland: Phronese.

- Boo, J. & Vispoel, W. (2012). Computer versus paper-and-pencil assessment of educational development: A comparison of psychometric features and examinee preferences. *Psychological Reports, 111*, 443-460. doi: 10.2466/10.03.11.PR0.111.5.443-460
- Boud, D. (2005). *Enhancing learning through self assessment*. NY, United States: Routledge
- Bukowski, W. M., Laursen, B., & Rubin, K. H. (2018). *Handbook of peer interactions, relationships, and groups*. NY, United States: The Guildford Press.
- Chen, C.-H. (2010). The implementation and evaluation of a mobile self- and peer-assessment system. *Computers & Education, 55*(1), 229-236. doi: 10.1016/j.compedu.2010.01.008
- Hwang, G.-J., Hung, C.-M., & Chen, N.-S. (2014). Improving learning achievements, motivations and problem-solving skills through a peer assessment-based game development approach. *Educational Technology Research and Development, 62*, 129-145. doi: 10.1007/s11423-013-9320-7
- Jenkins, M. (2004). Unfulfilled promise: Formative assessment using computer-aided assessment. *Learning and Teaching in Higher Education, 1*, 67-80.
- Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen* [Oratie]. Heerlen: Open Universiteit. Verkregen via http://dspace.ou.nl/bitstream/1820/3781/1/Joosten-ten%20Brinke_2011_Eigentijds%20toetsen%20en%20beoordelen.pdf op 15 juli 2018.
- Joosten-ten Brinke, D., Weber, A., & Ljubin Saveski, G. (2016). *Digitaal grootschalig summatief toetsen in het primair onderwijs*. Heerlen: Open universiteit.
- Kay, R., & Knaack, L. (2009). Exploring the use of audience response systems in secondary school science classrooms. *Journal of Science Education and Technology, 18*(5), 382-392. doi: 10.1007/s10956-009-9153-7
- Kerssen-Griep, J., & Witt, P. L. (2012). Instructional Feedback II: How Do Instructor Immediacy Cues and Facework Tactics Interact to Predict Student Motivation and Fairness Perceptions? *Communication Studies, 63*(4), 498-517. doi: 10.1080/10510974.2011.632660
- Van der Kleij, F., Vermeulen, J., Eggen, T., & Veldkamp, B. (2017). *Leren van toetsen: een cyclisch proces*. (herziene versie, oorspronkelijke uitgave 2013), Arnhem: Cito/RCEC.
- Matuk, C. F., Linn, M. C., & Eylon, B. (2015). Technology to support teachers using evidence from student work to customize technology-enhanced inquiry units. *Instructional Science, 43*(2), 229-257. doi: 10.1007/s11251-014-9338-1
- Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118-141. doi: 10.1016/j.jslw.2006.01.003

Mines Jr., R. O. (2014). The impact of testing frequency and final exams on student performance. *American Society for Engineering Education Southeast Section Conference*. Verkregen via <http://www.asee-se.org/proceedings/ASEE2014/Papers2014/19.pdf>.

Mojarrad, H., Hemmati, F., JafariGohar, M., & Sadeghi, A. (2013). Computerbased assessment (CBA) vs. Paper/pencilbased assessment (PPBA): An investigation into the performance and attitude of Iranian EFL learners' reading comprehension. *International Journal of Language Learning and Applied Linguistics World*, 4, 418-428.

Morris, J. (2001). Peer assessment: a missing link between teaching and learning? A review of the literature. *Nurse Education Today*, 21(7), 507-515. doi: 10.1054/nedt.2001.0661

Narciss, S. (2008). Feedback strategies for interactive learning tasks. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum. (pp. 125-144 proof-version)

Pino-Silva, J. (2008). Student perceptions of computerized tests. *ELT Journal*, 62(2), 148-156. doi:10.1093/elt/ccl056

Roeleveld, J., Oomens, M., & Coenen, A. (2015). *Leerwinst en toegevoegde waarde in het voortgezet onderwijs: Eindrapportage*. Den Haag/Utrecht/Amsterdam: Inspectie van het onderwijs/Oberon/Kohnstamm Instituut. Verkregen van: <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/publicaties/2015/08/20/leerwinst-en-toegevoegde-waarde-in-het-voortgezet-onderwijs/leerwinst-en-toegevoegde-waarde-in-het-voortgezet-onderwijs.pdf>.

Schildkamp, K., Heitink, M., Kleij, F. van der, Hoogland, I., Dijkstra, A., Kippers, W. & Veldkamp, B. (2014). *Voorwaarden voor effectieve formatieve toetsing. Een praktische review*. Enschede: Univeriteit Twente.

Sluijsmans, D., Joosten-ten Brinke, D., & Van der Vleuten, C. (2013). *Toetsen met leerwaarde: een reviewstudie naar de effectieve kenmerken van formatief toetsen*. Maastricht: Universiteit Maastricht.

Tsai, T. H., & Shin, C. D. (2013). A Score Comparability Study for the NBDHE Paper-Pencil Versus Computer Versions. *Evaluation & the health professions*, 36(2), 228-239. doi: 10.1177/0163278712445203

Welsh, M. (2007). Engaging with peer assessment in post-registration nurse education. *Nurse Education in Practice*, 7(2), 75-81. doi: 10.1016/j.nepr.2006.04.006

Williams, D. (2013). *Cijfers geven werkt niet*. Meppel, Nederland: Ten Brink.

Zimmerman, B. J. (2008). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk, & B. J. Zimmerman (Eds). *Motivation and Self-Regulated Learning: Theory, research, and applications*. New York, NY: Taylor & Francis.

Meer weten?

[Toets! Magazine over toetsen in het voortgezet onderwijs](#)

[Kennisrotonde \(2017\): Wat is er bekend over de effecten van formatief evalueren bij leerlingen en docenten?](#)

[Kennisrotonde \(2016\): Heeft het nakijken van schriften zin?](#)

Onderwijssector

Voortgezet onderwijs

Trefwoorden

Formatief toetsen, summatief toetsen, leerwinst, feedback