# Is het waar dat gestandaardiseerde toetsen bijdragen aan het betrouwbaar vaststellen van het beheersingsniveau op het gebied van rekenen en beginnende geletterdheid in groep 1 en 2?

Beantwoord door: Annegien Langeloo (antwoordspecialist), José van der Hoeven en update Eveline Schoevers (kennismakelaars Kennisrotonde)

10 november 2017 - update 12 april 2023

KR. 272

## Kort antwoord

De vraag of het toetsen van kleuters zinvol is, wordt in de discussie hieromtrent niet eenduidig beantwoord. Uit onderzoek is gebleken dat het niet eenvoudig is om de betrouwbaarheid van de toetsafname bij jonge kinderen te garanderen. In het kader van validiteit is uit een aantal onderzoeken gebleken dat kleutertoetsen over het algemeen een kleine tot matige voorspellende waarde hebben. Daarbij kwam naar voren dat de voorspellende waarde voor rekenen groter is dan voor taalvaardigheid. Tenslotte bleek dat de voorspellende waarde van cognitieve toetsen bij jonge kinderen groter is als deze gebaseerd is op observaties of leerkrachtbeoordelingen. Dit betekent dat kleutertoetsen een bijdrage kunnen leveren aan het vaststellen van het beheersingsniveau van cognitieve vaardigheden, maar dat daarnaast het eigen oordeel van de leerkracht over de leerprestaties en de cognitieve ontwikkeling ook van belang lijkt. Vanaf augustus 2023 mogen scholen volgens de wet echter geen gebruik meer maken van schoolse kleutertoetsen met pen en papier. Vanaf dan moeten scholen aan de hand van observaties nagaan hoe hun kleuters er voor staan. Hiervoor kunnen scholen observatie instrumenten gebruiken die zijn onderzocht op betrouwbaarheid en validiteit, zoals ‘Kleuter in Beeld’.

## Antwoord

Een toets neem je af om te onderzoeken in welke mate iemand iets weet of kan. Dat kan op verschillende manieren, zoals met een observatie of met mondeling of schriftelijk toetsen. Bij toetsen kun je gebruik maken van methodeafhankelijke of methodeonafhankelijke toetsen. Methodeafhankelijke toetsen geven een beeld van de beklijving van de leerinhouden die in de lessen aan bod zijn geweest. Bij methodeonafhankelijke of gestandaardiseerde toetsen gaat het om het beheersingsniveau van kennis of vaardigheden. De resultaten kunnen worden vergeleken met de gemiddelde prestaties van een groep en er kan groei over een langere periode worden vastgesteld (Papenburg en Visser, 2013). De zogenaamde kleutertoets is een voorbeeld van een

methodeonafhankelijke toets om de vorderingen van kleuters op het gebied van beginnende geletterdheid en gecijferdheid in kaart te brengen.

## Betrouwbaarheid van toetsen

Als we een vaardigheidstoets afnemen, willen we zeker weten dat de toets iets zegt over de vaardigheden van de leerling. Een eerste aspect dat hierbij een rol speelt, is betrouwbaarheid. De betrouwbaarheid van een instrument is de mate van nauwkeurigheid van dat instrument. Dat wil zeggen dat als we dezelfde vaardigheid twee keer met hetzelfde instrument meten, we twee keer dezelfde uitkomst zullen krijgen (Slotboom, 1987).
Onderzoek heeft laten zien dat er bij meerdere afnames van een toets voor motorische en communicatievaardigheden bij jonge kinderen veel variatie in toetsscores was (Darrah, Hodge, Magill-Evans & Kembhavi, 2003). Ook ander onderzoek bevestigt dat de toetsscore van jonge kinderen per dag kan verschillen. Dit kan komen doordat jonge kinderen vaker snel zijn afgeleid, nog niet kunnen lezen en soms onverwacht kunnen reageren op toetsvragen door bijvoorbeeld te kiezen voor het mooiste plaatje in plaats van voor het juiste antwoord (Bordignon & Lam, 2004; Colpin et al., 2006; Snow & Van Hemel, 2008). Dit zou kunnen betekenen dat besluiten genomen op basis van één testmoment tot verschillende conclusies kunnen leiden en dat de toetsing dus niet betrouwbaar is. Pas vanaf een leeftijd van 6 jaar wordt het beter mogelijk om leerlingen betrouwbaar te toetsen (Colpin et al., 2006; Shepard et al., 1998). De betrouwbaarheid van een toets kan echter per instrument verschillen. Het is daarom bij het inzetten van toetsen van belang dat de toets is onderzocht op betrouwbaarheid.

## Validiteit van toetsen

Een tweede aspect bij het meten van vaardigheden betreft validiteit. Validiteit heeft betrekking op de geldigheid van de meting: meet je wat je bedoelt te meten (Slotboom, 1987)? Om te bepalen of toetsen bij kleuters valide zijn, wordt vaak gekeken naar de voorspellende waarde van een toets (ook wel predictieve validiteit). Oftewel, in hoeverre de uitkomsten op een toets bij de kleuters de prestaties in hogere klassen kunnen voorspellen.
Onderzoek naar de voorspellende waarde van kleutertoetsen laat geen eenduidig beeld zien: onderzoeksresultaten geven aan dat toetsen een kleine tot matige voorspellende waarde kunnen hebben.

In 2000 publiceerden La Paro en Pianta een overzichtsstudie van meer dan 60 wetenschappelijke onderzoeken die zich hebben gericht op de relatie tussen cognitieve of sociaal-emotionele toetsresultaten bij de kleuters (4 – 6 jaar) en in groep 3 of 4 (6 – 8 jaar). Over het algemeen vonden La Paro en Pianta matige effecten voor de voorspellende waarde van cognitieve resultaten en kleine effecten voor voorspellende waarde bij sociaal-emotionele vaardigheden.

Ook recentere studies (Duncan et al., 2007; Romano et al., 2010) hebben de voorspellende waarde van resultaten op kleutertoetsen voor latere prestaties op school onderzocht. In deze studies werden vergelijkbare resultaten gevonden als in de overzichtstudie van La Paro en Pianta (2000). Beide onderzoeken vonden daarnaast dat de voorspellende waarde van taalvaardigheid kleiner was dan die voor rekenen. Bovendien bleek dat toetsen die gebruik maken van subjectieve beoordelingen, zoals

observaties of leerkrachtbeoordelingen met een checklist, de grootste voorspellende waarde hebben.

Terwijl eerdergenoemde studies alleen onderzochten of toetsresultaten bij de kleuters latere prestaties konden voorspellen, is ook onderzoek gedaan naar de patronen in deze voorspellingen (Dollaghan & Campbell, 2009, Frans et al., 2017). Onderzoek naar de voorspellende waarde van kleutertoetsen voor risicoleerlingen liet zien dat een kleine groep leerlingen consequent tot de laagst scorende groep hoorde. Er is echter ook een grote groep leerlingen die een wisselender patroon liet zien: over de verschillende leerjaren hadden zij soms hoge en soms lage scores. Een deel van deze kinderen, scoorde bij de kleuters gemiddeld of bovengemiddeld, maar vielen in groep 3 en/of 4 uit op de toets (Frans et al., 2017). Deze kinderen werden bij de kleuters dus niet als zodanig herkend, terwijl dat wel de bedoeling is van een dergelijke toets.

## Het doel van toetsen

Bovendien was er vanuit het werkveld veel kritiek op de kleutertoetsen. Kleutertoetsen zouden gericht moeten zijn op het tijdig signaleren van mogelijke tekorten op reken- en/of taalgebied, zodat kleuters snel preventief extra ondersteuning kunnen krijgen. Daarom is er besloten dat scholen vanaf augustus 2023 geen gebruik meer maken van schoolse toetsen met opgavenboekje, pen en papier voor het volgen van de ontwikkeling van kleuters (Onderwijsinspectie, 2022). Vanaf dan moeten scholen aan de hand van observaties nagaan hoe hun kleuters er voor staan. Hiervoor kunnen zij observatie-instrumenten gebruiken die door de Expertgroep zijn goedgekeurd. Dit is vastgelegd in de wet die in februari 2022 door het parlement is aangenomen en zo is het ook geformuleerd door de inspectie. Op basis hiervan heeft CITO bijvoorbeeld het observatie-instrument Kleuter in beeld ontwikkeld, dat inmiddels is goedgekeurd door de Expertgroep Toetsen PO. Zowel de betrouwbaarheid als validiteit van dit instrument bleek voldoende om de beginnende geletterdheid en gecijferdheid van kleuters in kaart te brengen.

## Betrouwbaarheid van de toetsing bij kleuters lastig te garanderen

Bij het beoordelen van toetsen spelen betrouwbaarheid en validiteit een grote rol. Uit onderzoek is gebleken dat de betrouwbaarheid van de toetsing bij kleuters lastig te garanderen is. In het kader van validiteit is uit een substantieel aantal onderzoeken gebleken dat kleutertoetsen een kleine tot matige voorspellende waarde hebben. Daarbij kwam naar voren dat de voorspellende waarde voor rekenen groter is dan voor taalvaardigheid.

Ten slotte bleek dat de voorspellende waarde van cognitieve toetsen bij jonge kinderen groter is als deze gebaseerd is op observaties of leerkrachtbeoordelingen. Dat betekent dat methodeonafhankelijke toetsen een bijdrage kunnen leveren bij het vaststellen van het beheersingsniveau van kleuters, maar vanwege het matige voorspellende effect niet allesbepalend kunnen zijn. Vanaf augustus 2023 mogen scholen volgens de wet echter geen gebruik meer maken van kleutertoetsen. Vanaf dan moeten scholen aan de hand van observaties nagaan hoe hun kleuters er voor staan.

## Geraadpleegde bronnen

Bordignon, C. M., & Lam, T. (2004). The early assessment conundrum: Lessons from the past, implications for the future. Psychology in the Schools, 41(7), 737-749.

Colpin, M., Gysen, S., Jaspaert, K., Heymans, R., Van den Branden, K., & Verhelst, M. (2006). Studie naar de wenselijkheid en haalbaarheid van de invoering van centrale taaltoetsen in Vlaanderen in functie van gelijke onderwijskansen. Leuven: Centrum voor Taal en Onderwijs KU Leuven.

Dollaghan, C. A., & Campbell, T. F. (2009). How well do poor language scores at ages 3 and 4 predict poor language scores at age 6?. International Journal of Speech-Language Pathology, 11(5), 358-365.

Darrah, J., Hodge, M., Magill-Evans, J., & Kembhavi, G. (2003). Stability of serial assessments of motor and communication abilities in typically developing infants—implications for screening. Early human development, 72(2), 97-110.
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Sexton, H. (2007). School readiness and later achievement. Developmental psychology, 43(6), 1428.

Frans, N., Post, W. J., Huisman, M., Oenema-Mostert, I. C., Keegstra, A. L., & Minnaert, A. E. (2017). Early identification of children at risk for academic difficulties using standardized assessment: stability and predictive validity of preschool math and language scores. European Early Childhood Education Research Journal, 1-19.

Kim, J., & Suen, H. K. (2003). Predicting children's academic achievement from early assessment scores: A validity generalization study. Early Childhood Research Quarterly, 18(4), 547-566.

La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. Review of educational research, 70(4), 443-484.

Lansink, N. (2003). Toetsen van jonge kinderen. Arnhem CITO.

Papenburg & Visser (2013). Kleutertoetsen: een waardevolle aanvulling! Cito.

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: replication and extension using a nationwide Canadian survey. Developmental psychology,46(5), 995.

Shepard, L., Kagan, S. L., & Wurtz, E. (1998). Principles and Recommendations for Early Childhood Assessments.

Slotboom, A. (1987). Statistiek in woorden. De meest voorkomende termen en technieken. Groningen: Wolters Noordhoff.

Snow, K. L. (2006). Measuring school readiness: Conceptual and practical considerations. Early education and development, 17(1), 7-41.

Snow, C., & Van Hemel, S. (2008). Early childhood assessment: Why what and how? Report of the Committee on Developmental Outcomes and Assessments for Young Children.

## Over deze vraag

| | |
|---|---|
| Opgesteld door: | Annegien Langeloo (antwoordspecialist), José van der Hoeven en update Eveline Schoevers (kennismakelaars Kennisrotonde) |
| Vraagsteller: | directeur bestuurder basisonderwijs |
| Geraadpleegde expert: | Niek Frans (Rijksuniversiteit Groningen), Geraadpleegde expert update: Hans van Luit, Universiteit Utrecht |
| Onderwijssector: | Primair onderwijs |
| Trefwoorden: | basisonderwijs, toets, taal, rekenen, kleuters |
| Referentie: | Kennisrotonde (2023). *Is het waar dat gestandaardiseerde toetsen bijdragen aan het betrouwbaar vaststellen van het beheersingsniveau op het gebied van rekenen en beginnende geletterdheid in groep 1 en 2?* (1e update) (KR. 272). |

Dit antwoord is gepubliceerd op Kennisrotonde.nl. De Kennisrotonde is samen met NCO en Onderwijskennis een dienst van het NRO.