WEBGAME GEBASEERD REKENEN OP DE BASISSCHOOL

Rekentuin.nl: combinatie van oefenen en toetsen

Han van der Maas, Sharon Klinkenberg en Marthe Straatemeier

Met behulp van moderne ICT middelen, nieuwe psychometrie en op basis van recente inzichten in de ontwikkelingspsychologie hebben de auteurs een nieuw oefen-volgsysteem voor het onderwijs ontwikkeld. Dit systeem combineert het oefenen van vaardigheden en het monitoren (toetsen) van ontwikkeling. De eerste uitwerking van dit systeem is Rekentuin.nl, een webgame gebaseerd, adaptief oefen-volgsysteem voor het rekenen op de basisschool. Dit artikel legt uit hoe dit systeem werkt en welke mogelijkheden het biedt. Kernvraag is of dit type 'low stakes' testsystemen ook potentie heeft voor examens.

In 2009 heeft de Rekentuin op het NVE congres (26 november 2009) de Stichting Solberg-Verlinden/Wynand Wijnen aanmoedigingsprijs ontvangen. De Rekentuin is ontwikkeld aan de vakgroep Psychologische Methodenleer van de Universiteit van Amsterdam. De jury noemde het project psychometrisch vernieuwend en van belang voor de praktijk van het rekenonderwijs. Met name wees de jury erop dat bij de schatting van de rekenvaardigheid niet alleen rekening wordt gehouden met het al dan niet correct beantwoorden van de items, maar ook met de tijd die het kind heeft besteed aan ieder item. Verder wordt de schatting van de itemmoeilijkheden steeds verbeterd door gebruik te maken van een methode die is gebaseerd op het ELO-ratingsysteem uit de schaakwereld.
Op Rekentuin.nl kunt u een demo account aanmaken om de Rekentuin in werking te zien.

Innovaties achter Rekentuin.nl

Computerspelletjes

Mensen besteden enorm veel tijd aan spelletjes. Dat kunnen puzzels zijn, gezelschapsspelletjes, sport, en tegenwoordig vaak (online) gaming. Een interessant aspect is dat in de meeste games ook vaardigheden worden gemeten. Van telspelletjes voor kleuters tot geheugenspelletjes voor bejaarden. Blijkbaar zijn mensen bereid om vaardigheidstests voor hun plezier te doen als ze in een spelvorm worden gegoten. Spelletjes worden daarom door onderwijsinstellingen en bedrijven steeds vaker ingezet als lesmiddel. Het kan interessant zijn om vanuit een psychometrisch oogpunt computerspelen te analyseren. In zulke spelen worden vaardigheden gemeten en vergeleken. Het is bijvoorbeeld opvallend dat erg veel games het concept 'level' gebruiken waar in de psychometrie met continue scores gewerkt wordt. Het gebruik van 'levels' is zo algemeen dat we ons moeten afvragen waarom deze zo gewaardeerd worden. Misschien kan de psychometrie hier iets van leren of andersom bijdragen aan een optimaal gebruik van 'levels' in de scoring van games.

Online

Rekentuin.nl is in essentie een website waarop kinderen een tuintje met plantjes onderhouden door regelmatig rekenspelletjes te spelen (figuur1). Het niveau van de items in de rekenspelletjes wordt bepaald door een nieuw adaptief algoritme. Idealiter spelen kinderen wekelijks of zelfs dagelijks ter vervanging van de sommen die ze in hun schriftjes maken. De responsen van de kinderen worden automatisch geanalyseerd en teruggekoppeld naar de leerkrachten in de vorm van foutenanalyses, groeicurven en andere vergelijkingen. Het adaptieve algoritme zorgt dus tegelijkertijd voor adequaat oefen- en testmateriaal. Bovendien dienen de data als bron voor wetenschappelijk onderzoek.

Responsetijden

Een vernieuwing is dat de responsetijden van

Spelletjes worden steeds vaker ingezet als lesmiddel

Figuur 1 Screenshot Rekentuin

kinderen worden meegewogen in de scores en de schatting van de rekenvaardigheid. Dit gebeurt door kinderen een scoringsregel op te leggen. Deze 'high speed, high stakes' regel houdt in dat kinderen hoger scoren als ze snel een correct antwoord geven. Een snel incorrect antwoord leidt echter tot een hele lage score. Gokken is dus riskant. Deze regel geeft het oplossen van rekenitems al meer het karakter van een spel. In standaard adaptieve tests worden items aangeboden waarop de kans op correct steeds ongeveer 50% is. Voor kinderen, en vaak ook voor volwassenen, is dat demotiverend. Voordeel van de scoringsregel is dat door het meenemen van reactietijden ook makkelijke items kunnen worden afgenomen, omdat juist op makkelijke items antwoordsnelheid informatie verschaft over vaardigheid.

Elo rating systeem

Een andere vernieuwing is de toepassing van het Elo rating systeem uit het schaken voor het schatten van zowel de vaardigheid van de kinderen als de moeilijkheid van de items. Schakers krijgen een rating (vaardigheidsschatting) op basis van de uitslagen van onderlinge schaakpartijen. In de Rekentuin spelen kinderen tegen items als waren het schaakpartijen. In gewone computer adaptieve tests moeten eerst de itemmoeilijkheden in een dure pretest fase bepaald worden. In dit systeem is dat niet meer nodig. Zowel de vaardigheidsschattingen als de itemmoeilijkheden worden voortdurend aangepast op basis van de antwoorden van spelers op items. Bij dit systeem kunnen startwaarden worden meegegeven aan de kinderen en de items, maar dit is niet noodzakelijk. In de Rekentuin krijgen kinderen een initiële vaardigheidsschatting op basis van hun leeftijd. De initiële itemmoeilijkheden worden bepaald op basis van somkenmerken.

Ontwikkeling van kinderen

Met de Rekentuin is het mogelijk om hoogfrequente data over de ontwikkeling van kinderen te verzamelen, bijvoorbeeld hoe een kind een bepaald item over de weken heen beantwoordt. Figuur 2 geeft een voor-

Kinderen spelen de Rekentuin veel op school maar ook thuis

beeld. In de grafiek van figuur 2 staat op de x-as de week waarin het item is gemaakt en op de y-as de reactietijd. Het antwoord '?' betekent dat het kind op het vraagteken heeft geklikt. In eerste instantie antwoordt dit kind snel maar kiest het bijna altijd voor het vraagteken. Halverwege week 42 leert het kind het correcte antwoord en antwoordt hij of zij met wisselende reactietijden. Dit wordt gevolgd door een periode van snelle en grotendeels correcte antwoorden.

Werkt Rekentuin.nl?

Er maken momenteel al duizenden kinderen gebruik van de Rekentuin. Per dag komen ongeveer 50.000 responsen binnen, in totaal al meer dan 18 miljoen. Kinderen spelen veel op school maar ook thuis. Ruim 1/3 van de responsen worden verzameld buiten schooltijd. Er zijn nu 12 spelletjes ontwikkeld waaronder de hoofddomeinen van het hoofdrekenen (optellen, aftrekken, vermenigvuldigen en delen). De betrouwbaarheid van de testen is hoog. De correlatie tussen domeinen als optellen en aftrekken is bijvoorbeeld .91. De correlatie met de scores van het Cito leerlingvolgsysteem is ook hoog (≈.8).

De toekomst

Individuele leerlijnen

Op basis van technologische ontwikkelingen is een spoedige doorbraak te voorzien als het gaat om het individueel gebruik van computers in het onderwijs. De schriftjes, die leerkrachten na schooltijd elke dag moeten corrigeren, gaan verdwijnen. Het is goed denkbaar dat basisschoolleerlingen over enkele jaren de Ipad uit hun laatje pakken om, onder meer, taal en rekenen te oefenen. Computer adaptieve monitoring systemen als Rekentuin komen dan goed van pas. Leerkrachten hoeven zich niet meer bezig te houden met toetsen, oefenen en nakijken maar kunnen zich richten op individuele instructie op basis van de informatie de zij verkrijgen van dergelijke systemen. Leerlingen kunnen veel meer dan nu, op hun eigen niveau werken en individuele leerlijnen volgen. Uit de data, verzameld met de Rekentuin, blijkt dat de individuele verschillen in rekenvaardigheid groot zijn. Op het domein optellen scoren bijvoorbeeld 19% van de kinderen in groep 4 boven het gemiddelde van groep 5 en 7% zelfs boven het gemiddelde van groep 6.

Blijven vernieuwen

De Rekentuin zal zich in verschillende opzichten blijven vernieuwen. Uiteraard ligt het voor de hand het aantal rekendomeinen uit te breiden. Binnenkort zal bijvoorbeeld klokkijken beschikbaar worden. Een andere vernieuwing is een betere verbinding met instructie via de foutenanalyses. Leerkrachten en leerlingen krijgen dan directe toegang tot bijvoorbeeld rekeninstructievideo's gekoppeld aan bepaalde typen fouten. Zo zal een leerling die veel fouten maakt in sommen met nullen (bv. 7000 - 3000 = 6700) direct toegang krijgen tot een scala aan filmpjes waarin deze fouten worden aangepakt. Dit levert tevens weer interessante onderzoeksdata op over de effectiviteit van instructiemethodes. Weer een andere vernieuwing betreft het onderhoud van de itembank door gebruikers. Leerkrachten en misschien zelfs kinderen mogen dan nieuwe items toevoegen en beheren. Het idee is om een zelforganiserend adaptief oefen- en volgsysteem te creëren waarin gebruikers zelf voor inhoud zorgen en de psychometrische parameters zich voortdurend

Figuur 2 Voorbeeld van dataverzameling over de ontwikkeling van een kind

De individuele verschillen in rekenvaardigheid in de Rekentuin zijn groot

automatisch aanpassen. Uiteraard onder continue monitoring ter bewaking van de kwaliteit.

Een taal equivalent

Uiteraard is de computer adaptieve monitoring technologie niet alleen bruikbaar voor rekenen en voor kinderen. Zo werkt het team momenteel aan een taal equivalent van de Rekentuin. Maar uiteindelijk is deze technologie bruikbaar voor alle gevallen waarin veel geoefend moet worden. Een voorbeeld is het statistiekonderwijs in de sociale wetenschappen. Met een itembank van enkele honderden opgaven is een 'statistiektuin' redelijk snel te realiseren, vooral als docenten en studenten zelf bereid zijn items in te voeren.

Low en high stakes testing

Vergelijking van prestaties in Rekentuin

Hoe verhoudt deze technologie zich nu met examens en andere tests waaraan belangrijke beslissingen zijn verbonden (*high stakes tests*)? Rekentuin is in eerste instantie een typisch voorbeeld van *low stakes testing*. Er zijn geen normen die zijn gebaseerd op een zorgvuldig uitgevoerd normeringsonderzoek. Maar omdat er inmiddels honderden gebruikers per leeftijdsgroep zijn, is het mogelijk een indicatie te geven van een percentielscore. Een leerkracht kan eenvoudig zien dat een leerling uit de klas op bijvoorbeeld vermenigvuldigen tot de beste 10% leerlingen behoort maar op aftrekken onder het gemiddelde scoort. Er is echter geen controle op de samenstelling van de steekproef. Het is denkbaar dat de ruim honderd scholen die nu gebruikmaken van Rekentuin voorop lopen in het rekenonderwijs, maar even goed kan beweerd worden dat juist zwakke scholen een dergelijk product afnemen.

Ook controle is een probleem. Kinderen spelen in de Rekentuin in de meest uiteenlopende omstandigheden. Soms speelt een kind zelfstandig in de klas waar op dat moment toevallig veel afleiding is, een andere keer speelt het thuis waar een vader of moeder mogelijk hulp kan bieden.

Uiteindelijk zullen deze nadelen wel meevallen. Hoe groter het aantal deelnemers wordt hoe betrouwbaarder de percentielscores worden. Bovendien wordt ook gebruik gemaakt van andere scoringssystemen waarin de prestatie van kinderen wordt uitgedrukt in het verwachte scoringspercentage op de hele itembank. Men hoeft niet zo bang te zijn voor vals spel. Een kind dat op een avond sterk in rating stijgt dankzij de hulp van een familielid wordt de eerstvolgende keer vanzelf, dankzij het adaptieve toetsingsmodel, geconfronteerd met veel moeilijkere items. Zijn of haar rating zal dan net zo snel weer dalen. Aan de andere kant kunnen kinderen die een keer slecht scoren, hun score de volgende dag alweer goed maken. Faalangst speelt een veel minder grote rol dan bij examens. Als kinderen jarenlang regelmatig oefeningen maken in dit adaptieve monitoringsysteem, zullen incidentele uitschieters naar boven of naar beneden geen rol spelen.

Combineren met examens

Een incidentele meting met een *low stakes test* zal natuurlijk nooit een vervanging zijn voor een goed opgezet examen. Indien op een gekozen tijdstip de vaardigheid van een persoon moet worden vastgesteld, geeft een examen met een zorgvuldig uitgevoerd normeringsonderzoek een beter beeld. Maar het is wel de vraag of het belang van examens niet zal afnemen als er meerjarige ontwikkelingsprofielen beschikbaar zijn gebaseerd op *low stakes* monitoringsystemen. De Cito eindtest op de basisschool heeft ook in belang ingeboet door het succes van het Cito leerlingvolgsysteem. Na 7 jaar volgen kan de eindtest eigenlijk niet meer voor grote verrassingen zorgen. Ook voor het voortgezet en hoger onderwijs is zo'n rol voor monitoringsystemen denkbaar. Op deze manier kunnen knelpunten eerder worden vastgesteld en aangepakt waardoor hogere slagingspercentages worden bereikt op de examens.

Voorlopig lijkt een combinatie van hoogfrequente *low stakes* volgsystemen en incidentele *high stakes* examens de beste oplossing. Dan heeft men zowel de beschikking over nauwkeurige groeiprofielen van kinderen als over goed genormeerde toetscores.

☐ De heer prof.dr. H.L.J. van der Maas, de heer drs. S. Klinkenberg en mw. drs. M. Straatemeier zijn verbonden aan de universiteit van Amsterdam (UvA) en Oefenweb.nl, het spin-off bedrijf van de UvA waar de Rekentuin nu onder valt. E-mail: h.l.j.vandermaas@uva.nl.