

Hoe redden we het onderwijs wèl?

Ewald Vervaeke¹

Op 30 mei 2023 was er in De Balie (Amsterdam) een debat over het oplossen van de onderwijscrisis.² Vóór de aanvang van het debat heeft Stichting Histos het pamflet 'Hoe redden we het onderwijs wèl?' onder het gehoor verspreid.³

Het pamflet opent met een schets van de onderwijscrisis:

- * Blijvend dalende leesvaardigheid en leesmotivatie, blijvend dalende rekenvaardigheid en blijvend dalende spellingvaardigheid ('hij vindt het vreemd'; groep 8!).
- * Blijvend oplopend lerarentekort: de pabo's leveren te weinig leerkrachten af; leerkrachten stappen voortijdig op, omdat er niet naar hen wordt geluisterd en ze alsmaar moeten trekken aan kinderen die nog niet aan bepaalde leerstof toe zijn.

Het pamflet noemt twee sleutels om de crisis op te lossen:

- * Het kind en zijn psychologische ontwikkeling volgens de lijn '... → oudere peuter → kleuter → jong schoolkind → ...' worden de kern van het onderwijs: aan de pabo's, in de scholen en bij de onderwijsinspectie.
- * Zaken als *Directe instructie* en *Monitor Leskwaliteit* verdwijnen uit het onderwijs.

Die sleutels tover ik niet uit een hoge hoed, maar hebben een wetenschappelijke grondslag.

Zienswijzen op het kind

Om de onderwijscrisis te plaatsen onderscheiden de meeste sprekers op 30 mei twee zienswijzen op het kind, zijn ontwikkeling en zijn onderwijs: de romantische en de empirische/wetenschappelijke.

De romantische zienswijze zou ten grondslag liggen aan de onderwijscrisis en de oplossing zou van de empirische/wetenschappelijke zienswijze moeten komen. Om de crisis op te lossen zou het hele onderwijs gebaseerd dienen te zijn op de uitkomsten van 'op evidentie gebaseerd of door evidentie geïnformeerd (voortaan: 'evidentief'⁴) onderzoek'. Onder 'evidentie' wordt verstaan: een statistisch significante uitkomst van een inferentieel-statistische⁵ berekening op getalsmaten zoals kalenderleeftijden, scores op puntschalen, vragenlijsten en meerkeuzetoetsen, aantal bezoeken per maand aan een bibliotheek, en zo meer.

Volgens mij zijn er echter geen twee maar drie zienswijzen en hoort evidentief onderzoek tot de kern van de onderwijscrisis.

1. De *romantische zienswijze* draait om een optimistisch mensbeeld, zoals Rousseaus theorie, Vygotskij's leer met de zone van de naaste ontwikkeling en Bruners 'ontdekkend leren' waarin geen aandacht is voor de rijpheid van het kind voor de te ontdekken leerstof. Met de sprekers van 30 mei wijs ik het toepassen van romantische zienswijzen in het onderwijs af.

2. De *empirische/wetenschappelijke zienswijze* draait om verschijnselen en om hun feitelijk nagetrokken verklaringen. Net als in de natuurwetenschappen en in de taalkunde verstaan empirische wetenschappers in de ontwikkelingspsychologie en onderwijskunde (voortaan 'O/O') zich met verschijnselen. Daar trachten ze beschrijvende verklaringen voor te vinden, die aan nieuwe feiten zijn nagetrokken. Een houdbare verklaring is een feit dat kan worden toegepast.⁶

Voorbeeld 1. Piagets theorie verklaart verschijnselen als non-permanentie (een zuigeling zoekt een voorwerp dat het heeft zien verdwijnen, niet op de verdwijplaats terug) en non-conservatie (een kleuter meent dat in { * * * * * } minder sterren zitten dan in { * * * * * }, ook als hij heeft geteld dat er in beide vijf zitten).

Voorbeeld 2. Mijn schrijf- en leestheorie verklaart verschijnselen als Eriks spiegelbeeldig schrijven van zijn naam als ÆRIK en Ilse's louter-hakkend lezen van LIP als 'L, i, p' als ze ILSE en PAPA kan schrijven.

Non-conservatie, spiegelbeeldig schrijven en louter hakken worden verklaard met de psychologische structuur van de kleuter. Die structuur bestaat uit onomkeerbare, abstract-logische verbanden. Vanwege die onomkeerbaarheid functioneert de kleuter in één richting. Als jong schoolkind zal Ilse LIP hakken-en-plakkend lezen als 'L, i, li; p, lip' of als 'L, i, p; lip'. Vooral de eerste leeswijze laat zien wat er bij het jonge schoolkind gebeurt:

- Ilse ziet L, herkent die uit ILSE en zet L om in de klank /l/;
- ze gaat *naar rechts*, ziet I, herkent die eveneens uit ILSE en zet I om in de klank /i/;
- * ze gaat *in haar hoofd terug* naar de klank /l/ en vormt met /i/ het klankgeheel /li/;
- ze gaat *naar rechts*, ziet P, herkent die uit PAPA en zet P om in de klank /p/;
- * ze gaat *in haar hoofd terug* naar het klankgeheel /li/ en vormt met /p/ het klankgeheel /lip/.

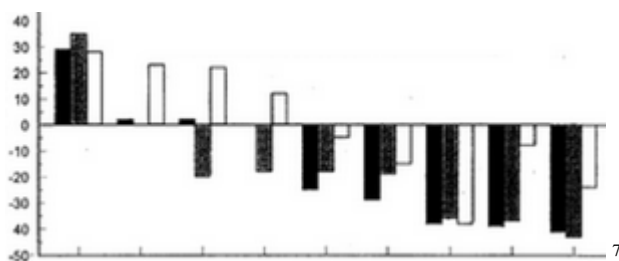
Het leren lezen van een woord van drie letters bestaat dus niet uit drie stappen, zoals men wellicht zou denken, maar uit vijf stappen. De drie stappen met • kon Ilse al als kleuter zetten, maar de twee stappen met * niet omdat haar operaties onomkeerbaar waren. Als jong schoolkind kan ze dat wel omdat haar operaties dan omkeerbaar zijn.

3. De *empiristische zienswijze* draait om data en inferentiële statistiek. De sprekers van 30 mei noemen de empiristische zienswijze ten onrechte 'empirisch', want er komt geen beschrijvend verklaren van verschijnselen en natrekken van verklaringen bij te pas. Deze zienswijze deelt uitgangspunten met de meeste mens-, maatschappij- en beleidswetenschappen. Zonder O/O-reden kent men aan gedragingen getalsmatige waarden toe en voert men daar inferentieel-statistische berekeningen mee uit zoals gemiddeldes, standaardafwijkingen, correlatiecoëfficiënten en significantieniveaus.

Voorbeeld 1. IQ-tests.

Voorbeeld 2. Meerkeuzetoetsen van het Cito en andere leveranciers van psychometrische meetinstrumenten.

Voorbeeld 3. Directe Instructie waar geen empirisch (in de zin van 2) bewijs en slechts psychometrisch-inferentieel-statistische evidentie voor is; zie de meetinstrumenten de *Metropolitan Achievement Test* (MAT), de *Wide Range Achievement Test* (WRAT), de *Raven's Coloured Progressive Matrices* (RCPM), de *Intellectual Achievement Responsibility Scale* (IARS) en de *Coopersmith Self-Esteem Inventory* (CSEI) en zie uitkomsten als



Twee twistappels

In het licht van de driedeling romantisch-empirisch-empiristisch zijn er in O/O twee twistappels: het al dan niet bestaan van de ontwikkelingsfasen en de status van psychometrische meetinstrumenten.

Ontwikkelingsfasen

Volgens de Zwitserse bioloog, epistemoloog en ontwikkelingspsycholoog Jean Piaget (1896-1980) en volgens Stichting Histos (opgericht in 1987) bestaan er ontwikkelingsfasen. Aan die fasen worden globale kalenderleeftijden gehangen om aan te geven aan wat voor kinderen we in de ongestoorde ontwikkeling dienen te denken. Het gaat echter niet om die kalenderleeftijden maar om de psychologische structuren van die fasen en om het feit dat ze in een vaste volgorde uit en op elkaar ontstaan. Zie hierboven voor de onomkeerbare psychologische verbanden van de kleuter en de omkeerbare psychologische verbanden van het jonge schoolkind.

Sedert ongeveer 1980 wordt in O/O in toenemende mate aan het bestaan van deze psychologische structuren getwijfeld. Dat is echter nooit omdat men dat empirisch aangetoond zou hebben maar altijd uit methodologische overwegingen. Rond ongeveer 2000 is dat twijfelen uitgemond in een vrijwel volledige marginalisering van de empirische psychologie. Testuitslagen en statistische berekeningen zouden laten zien dat psychologische structuren niet bestaan. Vooral de Amerikaanse psycholoog Charles Brainerd (geboren in 1944) heeft hierin een belangrijke rol gespeeld. Weliswaar erkent hij op vele plaatsen dat Piagets proeven replicerbaar zijn, maar toch wijst hij ze af omdat ze niet voldoen aan de methodologische eisen van onderzoek met getalsmatige data en inferentiële statistiek. Twee voorbeelden van Brainerds waardering voor Piagets theorie in zijn boek van 1987⁸:

‘Verschillende onderzoekers [...] hebben gevonden dat heel jonge kinderen inderdaad grafische verzamelingen construeren als hen een verzameling voorwerpen wordt gegeven om te sorteren’ (p.179).

‘Het onderzoek naar de oordelen van kinderen over evenwijdige rijen voorwerpen, en er is flink wat onderzoek beschikbaar, geeft consistente steun voor [Piagets drie-fase-verklaring]’ (p.186).

Zijn afwijzing verwoordt Brainerd onder meer zo: ‘[Piagets] toetsprocedures zijn niet wat we *gestandaardiseerd* noemen. [...] Voor degenen voor wie het wezen van wetenschappelijk onderzoek zijn onpersoonlijkheid en openbare communiceerbaarheid is, riekt dit feit onvermijdelijk naar priesterambt en obscurantisme’ (p.39v; Brainerds cursivering).

Brainerds stelling ‘empirisch replicerbaar; empiristisch niet-replicerbaar’ is volgens mij het kortst en het duidelijkst verwoord door de Nederlandse psycholoog Nico Frijda (1927-2015): ‘Methodologisch zijn Piagets experimenten vrijwel allen zwak te noemen; nauwelijks controle over de proefsituatie, vrijwel geen statistische verwerking der gegevens. Maar steeds zijn de experimenten origineel en de conclusies intrigerend (de meeste resultaten worden over door nauwkeurige replicaties door anderen bevestigd)’.⁹

Psychometrische meetinstrumenten

De stelling ‘empirisch replicerbaar; empiristisch niet-replicerbaar’ roept onder meer de vraag op naar de status van empiristische instrumenten als puntsschalen. Evidentieve onderzoekers noemen ze empirisch, maar zijn ze dat?

Een voorbeeld van een empirische schaal is de Beaufort-windschaal in de weerkunde. Getallen die daarbij horen, zoals in ‘windkracht 2 Beaufort’ van een weerdienst, zijn volledig door iets inhoudelijks (*i.c.* weerkundigs) gedekt; Beauforts allereerste empirische schaal (van 1807) bijvoorbeeld ordent winden naar de optuiging van een oorlogsschip met zeilen.¹⁰

Het antwoord op de vraag of die volledige inhoudelijke dekking er ook is bij psychometrische toetsen in de O/O, is helaas negatief. Nemen we een 7-puntsschaal. Op 20 beweringen kan men bijvoorbeeld antwoorden:

Nee! Nee. Nee? ? Ja? Ja. Ja!

Voor elk antwoord krijgt men een aantal punten, bijvoorbeeld op de bewering ‘Dit kind is goed in rekenen’ 1, 2, 3, 4, 5, 6 respectievelijk 7 punten.¹¹

Echter, deze puntentoeiding is slechts zinvol als er onder zo’n schaal een O/O-reden is. Dat wil zeggen, is het antwoord dat in rekenkundig opzicht 6 punten oplevert,

- * in O/O-opzicht zes keer zoveel waard is als het antwoord dat in rekenkundig opzicht 1 punt oplevert?
- * in O/O-opzicht drie keer zoveel waard is als het antwoord dat in rekenkundig opzicht 2 punten oplevert?
- * in O/O-opzicht twee keer zoveel waard is als het antwoord dat in rekenkundig opzicht 3 punten oplevert?

Zolang er geen O/O-theorie is die laat zien dat die puntentoedeling een O/O-basis heeft, verkrijgen we hier slechts rekenkundige feiten mee als ‘Dit kind heeft op die schooltoets 134 punten’ en ‘Dat kind heeft er 81 punten op’. Wat die rekenkundige feiten in O/O-opzicht betekenen, weet voorsnóg niemand.¹² Toch worden ze in evidentief onderzoek en onderwijs als O/O-feiten beschouwd en dus bijvoorbeeld ook als pleitend vóór Directe Instructie.¹³

Kortom, in empirisch opzicht hoeft aan het bestaan van psychologische structuren, kortweg ‘ontwikkelingsfasen’, niet te worden getwijfeld want zelfs empiristen als Brainerd en Frijda erkennen de (nauwkeurige) empirische replicerbaarheid van Piagetiaanse proeven. En psychometrische meetinstrumenten hebben vanwege de speculatieve puntentoedelingen geen volledige empirische grondslag zodat het wetenschappelijk gezien voorsnóg geen empirische toetsen zijn.

De onderwijscrisis

Op 30 mei werd de onderwijscrisis volledig toegeschreven aan de romantische zienswijze op het kind, zijn ontwikkeling en zijn onderwijs. Ook werd alle hoop gesteld op de niet-romantische zienswijze zoals in Directe Instructie. We hebben zojuist echter gezien dat Directe Instructie allerminst empirisch is onderbouw met O/O-feiten en dat er slechts psychometrische en inferentieel-statistische ‘evidentie’ voor pleit. Die ‘evidentie’ is echter empiristisch en niet empirisch van aard zoals we zojuist voor puntsschalen hebben gezien.

Iets dergelijks geldt voor het ontkennen of negeren van de psychologische ontwikkeling. De geestelijke vader van Directe Instructie, de Amerikaanse psycholoog Siegfried Engelmann (1931-2020) stelt onverbloemd dat Directe Instructie ‘ontwikkelingsvoortgang en -theorie schuwt’.¹⁴ Onrijpe leerlingen zouden met Directe Instructie ‘gesofisticeerde wiskundige operaties en logische schema’s’ leren op een manier die de gemiddelde tienjarige overtreft. In 1967 zou zijn onderzoek dat hebben uitgewezen.¹⁵

Ik heb Engelmanns onderzoek bij 19 kinderen gerepliceerd – 8 vierjarige peuters en 11 vijfjarige kleuters. Het blijkt dat de kleuters Engelmanns taak geheel op eigen kracht, dus zonder Directe Instructie of wat voor training ook, binnen enkele seconden oplossen en de peuters niet.¹⁶ Die taak vergt dus allerminst ‘gesofisticeerde wiskundige operaties en logische schema’s’. En andermaal is bewezen dat een kind dat over de benodigde psychologische structuur beschikt, de juiste reacties zelf kan ontdekken: de peuter beschikt niet over die structuur en de kleuter wel.

In Directe Instructie wordt de psychologische ontwikkeling zoals die in empirisch psychologisch onderzoek is gevonden, dus ten onrechte ontkend. Dit geldt voor alle benaderingen die niet op empirisch bewijs berusten maar waar slechts statistische evidentie voor is.

In het algemeen geldt dat het officiële onderwijsbeleid de afgelopen 20-25 jaar de onderwijscrisis heeft doen ontstaan en nu die crisis aan het verdiepen is tot een regelrechte onderwijsramp, door twee dingen ten onrechte te doen:

- ontkennen of negeren van de psychologische ontwikkeling van het kind volgens de lijn ‘... → oudere peuter → kleuter → jong schoolkind → ...’;
- opvolgen van de adviezen uit evidentief onderzoek door waarde te hechten aan significante uitkomsten van inferentieel-statistische berekeningen alsof die empirisch bewijs over de werkelijkheid waarin het Nederlandse onderwijs zich bevindt, zouden geven, terwijl er slechts sprake is van empiristische evidentie over een papieren werkelijkheid.

Hier volgen acht voorbeelden:

Voorbeeld 1. De invloed van het Cito is in de periode 2000-2023 versterkt, onder meer doordat de centrale eindtoets voor het basisonderwijs vanaf 2015 verplicht is.¹⁷

Voorbeeld 2. In 2006 bracht de Inspectie van het onderwijs de brochure *Iedereen kan leren lezen* uit, met daarin de aanbeveling om kinderen in groep 2 (en als het even kon ook in groep 1) letters aan te leren.¹⁸ Ik heb laten zien dat onderzoek waarin bewezen zou zijn dat dat een goede zaak is, niet op empirisch bewijs is gebaseerd maar slechts op empiristische en dus evidentief onderzoek berust.¹⁹

Voorbeeld 3. Sedert 2014 heeft het ministerie van OCW op toetsgebied naar een adviesorgaan. Daar is de Expertgroep uit voortgekomen, die niet naar de psychologische ontwikkeling van het kind kijkt en slechts naar psychometrische kwaliteiten van toetsen.²⁰

Voorbeeld 4. Gedurende enkele jaren was de aanduiding ‘COTAN-gecertificeerd’ bij onderwijsmateriaal een positief punt. De Inspectie van het onderwijs had er voor zover ons bekend geen bezwaar tegen en staat er volgens ons ook in 2023 positief tegenover. Welnu, COTAN staat voor ‘Commissie TestAangelegenheden Nederland’ en wie ‘test’ zegt, zegt niet ‘psychologische ontwikkeling van het kind’ en wel ‘psychometrie en inferentiële statistiek’.

Voorbeeld 5. *Directe Instructie* wordt al vele jaren door veel instellingen bij schoolbesturen aangeprezen, met zo veel succes dat er in februari 2021 al meer dan 50.000 exemplaren van een boek over Expliciete Directe Instructie waren verkocht. Het heeft de onderwijscrisis voorsnóg niet opgelost en volgens mij om bovenstaande redenen wel verdiept.

Voorbeeld 6. Het huidige regeerakkoord wil slechts ‘effectief bewezen lesmethodes (lees: met psychometrische en andere data en inferentiële statistiek)’.²¹

Voorbeeld 7. *Monitor Leskwaliteit* van de Inspectie van het Onderwijs, waarin met een 4-puntsschaal wordt gewerkt en leerkrachten worden beoordeeld vanuit normen die aan Directe Instructie zijn ontleend.²²

Voorbeeld 8. De motie over het terugdringen van evidentief bewezen ineffektieve lesmethodes en stimuleren van evidentief bewezen effectieve lesmethodes is op 30 mei 2023 maar liefst 129 stemmen vóór door de Tweede Kamer aangenomen.²³

Wat te doen?

Op grond van dit betoog dat op empirische feiten is gebaseerd, dienen er ten minste twee dingen te gebeuren als we de onderwijscrisis willen oplossen.

Onderzoek naar de waarde van evidentief onderzoek

Op dit ogenblik zijn evidentief onderzoek en onderwijs dominant onder O/O wetenschappers. In beginsel dient de overheid daarom hun adviezen op te volgen en die van de minderheid die empirische O/O beoefent, naast zich neer te leggen. Indien, wat de WSK, Stichting Histos en ik voorzien, de onderwijscrisis zich gaat verdiepen omdat de overheid sterk inzet op evidentief onderzoek, dan zijn de regering en de Tweede Kamer daar uiteindelijk wel verantwoordelijk voor. Het lijkt me daarom zaak om – als de politiek enige geloofwaardigheid schenkt aan dit artikel – een grondig onderzoek in te stellen dan wel te laten instellen naar evidentief onderzoek in het algemeen en naar de empirische dan wel louter rekenkundige aard van de uitkomsten van de inferentiële statistiek in het bijzonder.

Beoefenaren van de inferentieel-statistische methode zullen zo'n onderzoek uit zichzelf niet instellen want dat doen ze al niet na hun eigen replicatiecrisis in 2015.²⁴ Die replicatiecrisis is overigens een reden te meer om grote vraagtekens te zetten bij de waarde van evidentief onderzoek, ondanks zijn getalsmatige uiterlijk. Immers, wat in het ene evidentieve onderzoek statistisch significant is, is dat in een ander evidentief onderzoek wellicht of – gezien de titel 'Why most published research findings are false' van een artikel in noot 24 – zelfs waarschijnlijk niet. Ook achten empiristische onderzoekers als Brainerd en Frijda empirisch psychologisch onderzoek wél replicerbaar (zie onder het kopje 'Ontwikkelingsfasen'), wat vanuit empirisch-wetenschappelijk standpunt bekeken terecht is.

Uitbreiding van het aantal sleutels

In de inleiding heb ik twee sleutels genoemd om de onderwijscrisis op te lossen. Die kan ik nu toelichten en onderbouwen.

* Pabo's leiden theoretisch en praktisch op vanuit de ontwikkeling van het kind volgens de lijn '... → oudere peuter → kleuter → jong schoolkind → ...'. Voor die lijn is immers volop empirisch bewijs en wantrouwen tegen het bestaan van die ontwikkelingslijn vanuit evidentief onderzoek is niet terecht omdat evidentief onderzoek zelf geen grondslag heeft en geen empirisch bewijs kan leveren laat staan levert.

* De onderwijsinspectie gaat na of scholen leerlingen inderdaad in hun ontwikkeling volgen.

* Evidentief onderzoek en onderwijs, meerkeuzetoetsen, *Directe instructie* en *Monitor Leskwaliteit* verdwijnen uit het onderwijs en de Expertgroep aan het ministerie van OCW wordt opgeheven. Zie onder de kopjes 'De empiristische zienswijze' en 'Psychometrische meetinstrumenten' en de voorbeelden 3, 5 en 7 onder het kopje 'De onderwijscrisis'.

Met dank aan Rob Kooijman, secretaris van Stichting Histos

Noten

- 1 Dr. Ewald Vervaeke is wis- en natuurkundige, ontwikkelings- en leespsycholoog, oprichter en medewerker van Stichting Histos (www.stichtinghistos.nl/) en mede-oprichter en kerngroep lid van de Werk- en Steungroep Kleuteronderwijs (WSK) (<https://wsk-kleuteronderwijs.nl/>).
- 2 Video van het debat: www.debalie.nl/debalie-tv/red-het-onderwijs/. Bespreking van de debatavond: 'Inspectie mengt zich in debat' (Didactief-on-line: www.didactiefonline.nl/artikel/inspectie-mengt-zich-in-debat).
- 3 'Hoe redden we het onderwijs wél?' (www.stichtinghistos.nl/wp-content/uploads/2023/06/2023_05_30_pamflet-30-mei_DEFINITIEF.doc). Het pamflet is met medeweten en instemming van De Balie verspreid.
- 4 Het woord 'evidentief' munt ik naar analogie van 'inventie-inventief' en 'preventie-preventief'. Het alternatief 'evidentieel' heeft sedert Franz Boas (1911) en Roman Jakobson (1957) in de taalkunde een betekenis die sterk afwijkt van wat in 'op evidentie gebaseerd of door evidentie geïnformeerd' wordt bedoeld
- 5 De term 'statistiek' heeft ten minste zes betekenissen:
 - a. Beschrijvende statistiek. Dit is het tellen en turven van allerlei zaken ten behoeve van de staatsinrichting. Hier komt het woord 'statistiek' vandaan; het Italiaanse woord 'statistico' betekende oorspronkelijk 'de staat ('lo stato') betreffende.
 - b. Kans- of waarschijnlijkheidsrekening.
 - c. Verzekerings- of actuariële wiskunde.
 - d. Meetfouttheorie in de exacte wetenschappen.
 - e. Inferentiële statistiek in de mens-, maatschappij- en beleidswetenschappen. Deze is begonnen als een omkering van de meetfouttheorie in de exacte wetenschappen naar een meettheorie. Die meetfouttheoretische oorsprong is nog steeds te zien in aan het feit dat haar uitkomsten altijd afwijkingen van gemiddeldes zijn.
 - f. Stochastische natuurkunde.

Zie E. Vervaeke, 'Statistiek en de statistieken', *Struktuur en genese*, 2004, vol.17, p.26-54 en E. Vervaeke, 'Statistisch supplement – I', *Struktuur en genese*, 2005, vol.18, p.7-24.
- 6 Vier natuurwetenschappelijke en een taalkundig voorbeeld zijn:

Voorbeeld 1. Newtons zwaartekrachttheorie verklaart verschijnselen als de vrije val, de getijden en de maanbeweging.

Voorbeeld 2. De theorie van het periodieke systeem verklaart verschijnselen als de eigenschappen van edelgassen en de eigenschappen van de halogenen als fluor en chloor.

Voorbeeld 3. '[D]e quantumveldentheorie en de algemene relativiteitstheorie [...] beschrijven en verklaren een groot aantal experimenteel waargenomen verschijnselen' (minister Dijkgraaf, *A geometrical approach to two-dimensional conformal field theory* (proefschrift), 1989, p.194; <https://dspace.library.uu.nl/handle/1874/210872>).

Voorbeeld 4. Pasteurs kiemtheorie verklaart verschijnselen bij bedorven voedsel met microscopische georganiseerde wezens (en niet, zoals voordien, met spontane generatie); die wezens worden nu 'bacteriën' genoemd, een term die al in 1838 is gemunt door de Duitse bioloog en geoloog Ehrenberg (1795-1876).

Voorbeeld 5. Bentley's theorie verklaart metrische afwijkingen in Homerus met de 'wau', een uit het Grieks verdwenen letter die als de Engelse 'w' moet hebben geklonken.
- 7 G.L. Adams & S. Engelmann, *Research on Direct Instruction; 25 years beyond DISTAR*, Seattle, Educational Achievement Systems, 1996, p.71v. E. Vervaeke, 'Directe Instructie (DI)', *Struktuur en genese*, 2021, vol.33, p.4-53, met name p.25-27 (www.stichtinghistos.nl/wp-content/uploads/2021/10/SG2021_directe-instructie-di_p.4-53.pdf).
- 8 C.J. Brainerd, *Piaget's theory of intelligence*, Englewood Cliffs, Prentice-Hall, 1978.

- 9 N.H. Frijda, 'Inleiding', in J. Piaget, *De psychologie van de intelligentie*, Amsterdam, De Bussy, 1970 (eerste, Franstalige uitgave van 1947), met name p.9.
- 10 E. Vervae, 'Operationaliseren', *Struktuur en genese*, 2022, vol.34, p.4-50, met name p.33-37. Bijvoorbeeld, windkracht 2, 'lichte bries' geheten, is 'wind die een oorlogsschip dat met alle zeilen is bijgezet, 3 of 4 knopen zal dragen' en windkracht 10, 'zware storm' geheten, is 'wind waarin een oorlogsschip geen ander zeildoek kan laten zien dan stormstagzeilen'.
- 11 Of op de bewering 'Dit kind is slecht in rekenen' 7, 6, 5, 4, 3, 2 respectievelijk 1 punt.
- 12 In empiristisch onderzoek tracht men een punttoedeling als 1-2-3-4-5-6-7 te nuanceren met de zogeheten item-response-theorie (IRT); zie bijvoorbeeld https://en.wikipedia.org/wiki/Item_response_theory. De punttoedeling wordt dan bijvoorbeeld 1,3-2,4-2,8-3,9-5,6-6,4-6,9.
- In de empirische wetenschappen verwijst een theorie altijd naar verschijnselen die men met die theorie tracht te verklaren en is ze nagetrokken aan nieuwe feiten. Zie de voorbeelden in noot 6. Met de IRT tracht men geen psychologische verschijnselen te verklaren zodat ze alleen al om die reden geen theorie is maar een leer.
- De IRT (of eigenlijk IRL, met L van 'leer') doet geen recht aan het feit dat een punttoedeling in O/O gebaseerd dient te zijn op een O/O inzicht. Er is echter geen O/O theorie waaruit volgt dat het antwoord dat 3,9 punten krijgt, in O/O opzicht drie keer zo veel waard is als het antwoord dat 1,3 punten oplevert en dat het antwoord dat 5,6 punten krijgt, in O/O opzicht twee keer zo veel waard is als het antwoord dat 5,6 punten oplevert. Met andere woorden, met of zonder IRT leveren de psychometrie en de inferentiële statistiek slechts rekenkundige feiten op en geen O/O feiten, terwijl ze wel als O/O feiten worden gebracht en worden opgevat.
- 13 Iets dergelijks geldt voor vragenlijsten en meerkeuzetoetsen. Als daarin 40 beweringen of vragen staan met elk 4 antwoordmogelijkheden, dan neemt men voor de totaalscore S_{40} en de 40 afzonderlijke scores s_i ($i = 1, \dots, 40$) = 0 of 1, feitelijk de formule $S_{40} = s_1 + s_2 + \dots + s_{39} + s_{40}$, aan. Echter, voor zo'n formule is geen grondslag in enige O/O-theorie.
- In al deze gevallen maken beoefenaren van evidentief onderzoek op begripsmatig vlak een speculatieve sprong die vergelijkbaar is met de speculatieve sprong in de alchemie. Precies zoals alchemisten ten onrechte dachten dat ze met scheikundige proeven van lood goud konden maken, zo denken beoefenaren van evidentief onderzoek dat ze met inferentieel-statistische berekeningen van rekenkundige feiten O/O-feiten zouden kunnen maken. Echter, precies zoals alchemisten slechts klatergoud bleken te hebben als ze succes meenden te hebben, zo hebben beoefenaren van evidentief onderzoek slechts schijnresultaten.
- 14 G.L. Adams & S. Engelmann, *Research on Direct Instruction; 25 years beyond DISTAR*, Seattle, Educational Achievement Systems, 1996, hoofdstuk 3, met name p.27 (www.nifdi.org/docman/suggested-reading/book-excerpts/research-on-direct-instruction-25-years-beyond-distar-engelmann-adams-1996/178-chapter-3-myths-about-di-research-on-di-25-years-beyond-distar-1996/file.html); van internet geplukt op 7 juni 2023).
- 15 S. Engelmann, 'Teaching formal operations to preschool advantaged and disadvantaged children', *The Ontario journal of educational research*, 1967, vol.9, p.193-207.
- 16 'Directe Instructie (DI)' (op.cit), p.19v.
- 17 'Cito: de lange tentakels van een toetsfabriek' (www.didactiefonline.nl/artikel/actueel-dossier-cito).
- 18 Inspectie van het onderwijs, *Iedereen kan leren lezen*, september 2006, met name p.3 en p.5 (www.makkelijklezenplein.nl/index.php?cmd=file&action=download&file=1366#:~:text=Iedereen%20kan%20leren%20lezen'%20is,de%20Inspectie%20van%20het%20Onderwijs.&text=Hierbij%20ontvangt%20u%20de%20brochure.op%20het%20gebied%20van%20leesonderwijs); van internet geplukt op 7 juni 2023).
- 19 E. Vervae, *Naar school; psychologie van 3 tot 8 jaar*, Delft, Elmar, 2012 (eerste druk van 2007), p.186-191.
- 20 'Maak kennis met de Expertgroep Toetsen Primair Onderwijs' (www.tijdschriftdepsycholoog.nl/wp-content/uploads/2019/01/De-Psycholoog-12-TOTAAL-DEF-LR-31-34.pdf); van internet geplukt op 7 juni 2023).
- 21 'Omzien naar elkaar, vooruitkijken naar de toekomst', p.23 (<https://open.overheid.nl/documenten/ronl-f3cb0d9c-878b-4608-9f6a-8a2f6e24a410/pdf>).
- 22 'Monitor leskwaliteit po – technisch rapport', met name p.8 (4-puntsschaal) en p.9v (verband met Directe Instructie); op p.32-50 staan vele tientallen getallen die inferentieel-statistisch significant zijn (zoek bijvoorbeeld op 'signif') (www.onderwijsinspectie.nl/documenten/rapporten/2023/03/06/monitor-leskwaliteit). Zie ook 'Factsheet Monitor leskwaliteit po' (www.onderwijsinspectie.nl/documenten/publicaties/2023/03/06/factsheet-monitor-leskwaliteit-po).
- 23 'Motie van het lid Paul c.s. over bewezen ineffectieve lesmethodes terugdringen en bewezen effectieve lesmethodes stimuleren' (www.tweedekamer.nl/kamerstukken/moties/detail?id=2023Z08915&did=2023D21369).
- 24 'Replication crisis' (www.psychologytoday.com/us/basics/replication-crisis#:~:text=The%20term%2C%20which%20originated%20in,that%20%20study%20using%20similar%20procedures).
 'Why most published research findings are false'
https://www.researchgate.net/publication/7686290_Why_Most_Published_Research_Findings_Are_False
 'Estimating the reproducibility of psychological science'
http://www.researchgate.net/publication/281286234_Estimating_the_reproducibility_of_psychological_%20science.