

Ontwikkeling van een computergestuurde test voor Reflectiviteit-Impulsiviteit

M.M. Liew-On en P.R.J. Simons

Verschillen tussen reflectieven en impulsieven

Kagan (1965) ontwikkelde de oorspronkelijke Matching Familiar Figures Test (MFFT) op basis van de hypothese dat tijd besteed aan reflectie en analyse tot minder fouten leidt, terwijl snelle reacties in een ambigue taak met opgaven waarbij het antwoord niet direct voor de hand ligt veel fouten opleveren. Hij is van mening, dat reflectiviteit-impulsiviteit een cognitieve stijldimensie is. Reflectieven nemen ook in andere testsituaties meer tijd en zouden volgens Kagan uit zijn op nauwkeurigheid; zij zouden dus hogere prestatienormen hebben dan impulsieven. In longitudinaal onderzoek vond hij, dat kleuters die later als reflectief werden geclassificeerd, veelvuldig lachen nadat zij een probleem hebben opgelost en moeilijker problemen prefereren dan kleuters die later worden geclassificeerd als impulsieven. Reflectieven tonen ook meer teleurstelling na het maken van fouten op de MFFT. Gedragsobservaties van Cameron (1977) wezen uit, dat reflectieve kinderen stoppen als ze een fout gemaakt hebben en nadenken, terwijl impulsieven gewoon doorgaan. Revelle (1986) vond dat impulsieven sneller beginnen aan een taak, maar het niet zo lang volhouden als reflectieven. De laatsten zijn ook minder gevoelig voor beloningen.

Revelle meent dat reflectieven weten dat zij efficiënt kunnen werken onder tijdsdruk en daarom geneigd zijn het kalm aan te doen tot de deadline bijna bereikt is. In onderzoek van Meichenbaum, Burland, Gruson & Cameron (1985), waarin de kinderen hardop moesten denken bij het uitvoeren van taken, bleken reflectieven zichzelf beter te instrueren en betere oplossingsstrategieën te formuleren.

Deze en andere onderzoeken schetsen een beeld van impulsieven als inefficiënte probleemoplossers. Zij verwerken informatie minder systematisch en passen minder adequate strategieën toe dan reflectieven. Uit verschillend onderzoek bleek dat strategie-trainingen kunnen leiden tot betere prestaties van impulsieve kinderen op andere taken (Meichenbaum, 1977). Soppe (1979) geeft een overzicht van onderzoek waaruit de modificeerbaarheid van impulsiviteit blijkt en hij concludeert dat directe strategie-trainingen bij 8 tot 12 jarigen tot verbeteringen leiden, maar alleen op korte termijn.

De originele Matching Familiar Figures Test

De Matching Familiar Figures Test (MFFT) is een individuele, schriftelijke test, waarmee de dimensie reflectiviteit-impulsiviteit -ook wel conceptueel tempo genoemd- wordt gemeten. Hierbij noteert de testafnemer de fouten en de reactietijden, die met behulp van een stopwatch worden gemeten. De testpersonen worden, afhankelijk van de

reactietijd en het aantal fouten, ingedeeld in vier groepen: langzaam en accuraat (reflectief), langzaam en inaccuraat, snel en inaccuraat (impulsief) en snel en accuraat. Ongeveer 70% van de testpopulatie blijkt als reflectief dan wel impulsief gecategoriseerd te worden; de resterende 30% is ongeveer gelijk verdeeld over de groepen snel-accuraat en langzaam-inaccuraat.

De betrouwbaarheid van de foutenscores is gering; bij test-hertest onderzoek werden correlaties tussen .20 en .70 voor de reactietijden en tussen .20 en .50 voor de foutenscores gevonden. De interne consistentie is tussen .40 en .60 (Soppe, 1979). Ook de standaardisering is inadequaat. Problemen met de standaardisering worden deels veroorzaakt door de individuele afname-procedure, deels door de wijze van scoring. Door de dubbele mediaansplit zijn de scores niet normaal verdeeld, waardoor er statistische moeilijkheden ontstaan bij de vergelijking van de testresultaten met andere tests (Ault, Mitchell en Hartmann, 1976). De test is niet genormeerd en dientengevolge is er geen criterium om testresultaten te vergelijken. Merriënboer en Jelsma (1988) lieten als tussenoplossing de antwoorden op de MFFT op de computer intypen; ze geven aan, dat een geheel computer-gestuurde test de standaardisering zou optimaliseren. Verder wijst Van de Vijver (1987) erop, dat computergestuurde afname ook een aanzienlijke arbeidsbesparing oplevert. In het kader van onderzoek naar zelfstandig leren is daarom de computergestuurde test van reflectiviteit-impulsiviteit (C-MFFT) ontwikkeld, voornamelijk met het oog op de tijds- en arbeidsbesparing die computergestuurde afname oplevert (Simons en Liew-On, 1989; Liew-On, 1990).

De computergestuurde Matching Familiar Figures Test

De ontwikkeling van de C-MFFT omvatte 3 fasen.

1e fase: Programma.

Na het ontwerp van het programma werd de software-kwaliteit geëvalueerd aan de hand van eerder geïnventariseerde criteria (Daal en Liew-On, 1988). Tegelijk werd ook de testkwaliteit, moeilijkheidsgraad van de items en de itemalternatieven door deskundigen en een groep studenten (die als naïeve gebruikers fungeerden) beoordeeld.

2e fase: Testafname.

Bij 155 leerlingen uit de 3e klas van de HAVO (gemiddelde leeftijd 15 jaar) van 3 verschillende scholengemeenschappen (Waalwijk en Tilburg).

3e fase: Normering.

Uit de verkregen resultaten is een criterium gedestilleerd, waarmee indeling van de testpersonen meteen plaatsvindt.

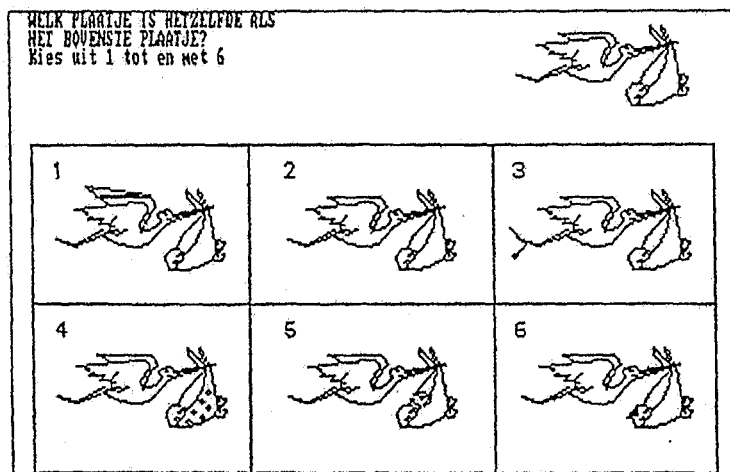
De C-MFFT bestaat uit 13 items waarvan de eerste twee oefenopgaven zijn. De opgaven bestaan uit één normfiguur en 6 tot 8 keuzemogelijkheden. Uit de laatste moet die figuur geselecteerd worden die identiek is aan de normfiguur. Deze wordt ook wel de standaard genoemd. Testafname duurt ca. 20 minuten. De test kan op elke IBM-compatibele computer met minimaal 512 kB Ram gebruikt worden. Voor presentatie van de

plaatjes is een CGA-grafische kaart nodig. Bij gebruik van een hulpprogramma waarmee een CGA kaart wordt gesimuleerd, kunnen in de plaatjes afwijkingen optreden.

Van de Vijver (1987) wijst erop, dat bekendheid met computers van belang kan zijn voor prestaties op computergestuurde tests. Om te voorkomen, dat de test een beroep doet op deze vaardigheden, wordt vooraf een computergestuurd instructieprogramma afgenomen waarin de relevante toetsen worden getoond en uitgelegd.

Ontwerp en Constructvaliditeit

Enerzijds moest de C-MFFT als parallelversie van de Matching Familiar Figures Test kunnen fungeren, anderzijds moest rekening worden gehouden met de grenzen en mogelijkheden die het gebruik van de computer impliceert. Bij de constructie van de C-MFFT is, in navolging van de oorspronkelijke MFFT, gestreefd naar een ambigue probleemsituatie met hoge antwoordonzekerheid. De afbeeldingen moesten daarom zo gekozen worden, dat er voldoende mogelijkheden waren relevante onderdelen te variëren. De antwoordonzekerheid wordt ook bepaald door saillantie van de stimulusdimensies; kenmerken van stimuli mogen niet teveel opvallen. De positie van de alternatieven t.o.v. de standaard bleek ook van belang; het moest niet in één oogopslag duidelijk zijn wat het juiste alternatief was. Bij de evaluatie bleek, dat de rechterbovenhoek de beste plaats was voor de standaard; de tekst kon anders interfereren met de normfiguur. Een bijkomend voordeel van plaatsing in de rechterbovenhoek is dat het uitnodigt tot systematisch vergelijken van de normfiguur met de alternatieven. De juiste zoekstrategie omvat het elimineren van de mogelijkheden die niet correct zijn. Ter illustratie van deze criteria is een testplaatje opgenomen:



Figuur 1: Voorbeeld van een C-MFFT item

Het stimulusmateriaal werd ontwikkeld en geëvalueerd op Olivetti M-19-computers met twee diskdrives, 512 Kb intern geheugen en een CGA grafische kaart. Het programma is geschreven in de auteurstaal PILOT. PILOT verschaft een flexibele schermcontrole, waarbij diagrammen, tekst, feedback en antwoorden op hetzelfde scherm kunnen worden geplaatst. Per item wordt geregistreerd welk alternatief de leerling kiest en hoeveel tijd er verloopt tussen presentatie en intoetsing van de antwoorden. De tijd wordt geregistreerd in seconden. Een eenvoudig tekenprogramma PCPG van Eugene Ying uit de Public Domain Software bibliotheek werd gebruikt om de plaatjes te maken. De afbeeldingen werden m.b.v. twee hulpprogramma's van PILOT geconverteerd en gecompimeerd.

Procedure

De drie deelnemende scholen hadden verschillende IBM-compatibele computers beschikbaar. De leerlingen werden tijdens de testafname gefilmd om later na te kunnen gaan of storende factoren een rol hadden gespeeld bij de afname. Het was immers van belang, dat de reactietijd alleen verband hield met het oplossingsproces. De test werd afgenomen bij 155 leerlingen uit de 3e klas Havo: 57 jongens en 98 meisjes.

Resultaten.

De score-indeling

Bij de oorspronkelijke MFFT werd geen numeriek scoringssysteem gehanteerd; bij de C-MFFT wordt wel een score toegekend. Door een andere scoreverdeling te hanteren is getracht de vergelijkbaarheid de resultaten van verschillende testafnames te verhogen.

Bij test-hertest onderzoek van de schriftelijke MFFT bleek dat leerlingen vaak verschillend werden geclassificeerd (Soppe,1979), bijvoorbeeld eerst als impulsief en bij hertest als snel-accuraat. Om dit probleem te ondervangen worden in de scoreverdeling bij de C-MFFT twee groepen van zgn. "twijfelgevallen" onderscheiden. De scores van deze groepen vallen in het gebied rond de mediaan van de responsetijd.

TEMPO	AANTAL FOUTEN	
	Veel (> 6)	Weinig (≥ 6)
Snel (RT < 450)	impulsieven	snel en accuraat
Gemiddeld (RT 450-564)	twijfelgevallen	twijfelgevallen
Langzaam (RT > 564)	langzaam en inaccuraat	reflectieven

Tabel 1: indeling op de C-MFFT aan de hand van fouten en reactietijden (RT = totale responsetijd in seconden, waarbij de responsetijd de tijd tot het eerste antwoord is).

Itemanalyse van de C-MFFT: moeilijkheidsgraad van de items, item-test-correlatie, betrouwbaarheid en homogeniteit.

Om de betrouwbaarheid van een test te verhogen wordt vaak testverlenging aanbevolen. Dat is ook gebeurd bij de C-MFFT (vgl. Cairns & Cammock, 1978). Bij de evaluatie van de C-MFFT bleek echter dat na presentatie van 16 afbeeldingen vermoeidheidsverschijnselen optreden. Voor het samenstellen van de definitieve versie is Nunnally (1967) gevolgd: items die het hoogst correleren met de totale testscore moeten worden geselecteerd, omdat deze waarschijnlijk minder ambigue zijn, niet extreem moeilijk of gemakkelijk zijn, meer variantie gemeen hebben met de gemeenschappelijke factor van de items en daardoor de betrouwbaarheid van de test verhogen.

Eerst werden de p-waarden berekend om na te gaan of de items de juiste moeilijkheidsgraad hebben. De p-waarden is de proportie leerlingen met nul fouten op een item: $p=P/N$, waarbij P het aantal leerlingen met nul fouten is en N het totaal aantal responsen op dat item (de Groot, 1973). Omdat bij de C-MFFT meer fouten op een item gemaakt konden worden, werd niet het aantal leerlingen, maar het aantal responsen per item gebruikt. Uit een p-waarde van 1 of bijna 1 kan geconcludeerd worden dat bijna alle leerlingen het goede antwoord geven. Hoe kleiner p, hoe moeilijker het item; als p kleiner is dan 0.50 maakt de meerderheid het item fout. In tabel 2 zijn de p-waarden opgenomen.

Vervolgens werden de item-rest correlaties berekend om na te gaan of elk item differentieert tussen "betere" en "minder goede" leerlingen. De eerste twee opgaven zijn oefenopgaven, die mogen dus niet differentieren. Tevens werden de betrouwbaarheid en homogeniteit van de test berekend.

ITEMS	1	2	3	4	5	6	7	8	9	10	11	12	13
p=P/N	.74	.64	.53	.53	.50	.60	.78	.72	.79	.64	.40	.75	.77
r(item-rest)	.04	.06	.39	.50	.40	.32	.19	.29	.41	.32	.40	.29	.16

Tabel 2: De p-waarden en de correlaties van de items met de test-score

Uit tabel 2 blijkt dat de hoogste p-waarde 0.79 is. Voor geen enkel items was p gelijk aan de gokkans. Alleen de p-waarde van item 11 ($p=0.40$) geeft aan dat relatief veel leerlingen dit item fout maken. Omdat de gemiddelde p-waarde 0.65 is en de gemiddelde moeilijkheidsgraad ligt tussen 0.52 en 0.64, is dit item gehandhaafd.

De Groot (1973) geeft als vuistregel, dat een item-rest-correlatie kleiner dan 0.15 er op duidt dat het item niet differentieert. In tabel 2 is te zien dat deze coëfficiënt voor de testopgaven (item 3 t/m 13) in alle gevallen groter is. Bij de eerste oefenopgave is deze 0.04 en bij de tweede is dat 0.06; deze items correleren niet met de totaalscore zoals ook

de bedoeling was. De correlaties van item 7 en item 13 zijn slechts iets hoger dan .15 en zouden eventueel voor verbetering in aanmerking komen.

De homogeniteit bleek 0.89 te zijn, een aanmerkelijke verbetering ten opzichte van de MFFT. Voor de reactietijden afzonderlijk werd een coëfficiënt alfa van 0.9 gevonden en voor de fouten was alfa 0.62. Zoals te verwachten viel, bleken de fouten significant negatief te correleren met de responsetijden. ($r=.31$; $n=155$; $p=.00$).

Conclusie

Het bleek mogelijk een computergestuurde test van reflectiviteit-impulsiviteit te ontwikkelen die een aantal bezwaren opheft die aan de oorspronkelijke MFFT kleven.

Met behulp van de C-MFFT is verder onderzoek naar de theoretische betekenis van reflectiviteit-impulsiviteit mogelijk. Gegevens, die tot nu toe niet in onderzoek betrokken werden -zoals die van langzame, inaccurate leerlingen- kunnen ook opgeslagen en verwerkt worden voor onderzoeksdoeleinden. Ook wordt hiermee een aanzet gegeven de test voor andere populaties te normeren.

De test blijkt goed te differentiëren; de interne consistentie is hoog.

De verbeteringen t.o.v. de MFFT, behalve de eerder genoemde, zijn:

1. De standaardisering werd verbeterd door een nieuwe score-indeling te hanteren.
2. Ook voor meting van de stabiliteit is de huidige score-indeling meer geëigend. Bij de scoring van de originele MFFT ontstond door de dubbele mediaansplit een indeling in vier groepen. De relatief lage stabiliteit kan veroorzaakt worden doordat een aantal grensgevallen de ene keer als reflectief, de andere keer als behorend tot een andere groep wordt geclassificeerd. Daarom worden in de nieuwe score-indeling twee groepen van zogenaamde "twijfelgevallen" onderscheiden.
3. De relatief geringe kosten en de verbeterde standaardisering van de afname maken normeringsonderzoek nu mogelijk.

Met de C-MFFT kan op praktische wijze een bepaald type deficiëntie in de studie-aanpak van leerlingen gediagnosticeerd worden, namelijk te snel en onsystematisch werken waardoor teveel fouten worden gemaakt. Voor toepassing bij andere groepen dan de hier onderzochte populatie is verder normeringsonderzoek noodzakelijk.

Literatuur.

- Ault, R.L., Mitchel, C. & Hartmann, D.P. (1976). Some Methodological Problems in Reflection-Impulsivity Research. *Child Development*, 47, 227-231.
- Cairns, E. & Cammock, T. (1978). Development of a more Reliable Version of the Matching Familiar Figures Test. *Developmental Psychology*, 14, 5, 555-560.

- Cameron, R. (1977). *Conceptual tempo and children's problem solving behavior: A Developmental Task Analysis*. Unpublished doctoral dissertation, University of Waterloo, Canada.
- Daal, M. & Liew-On, M. (1988) *De ontwikkelingsgang van een computergestuurd metacognitief trainingsprogramma, gericht op oriëntering*. Tilburg: Katholieke Universiteit Brabant.
- Groot, A. de & Naerssen, R.F. van (1973). *Studietoetsen construeren, afnemen, analyseren, deel II*. Den Haag: Mouton.
- Kagan, J. (1965). Impulsive and Reflective Children. Significance of Conceptual Tempo. In: Krumboltz, J.D. (Ed.) *Learning and the educative process*. Chicago: Rand McNally.
- Liew-On, M.M. (1990). *De effecten van een computergestuurd metacognitief trainingsprogramma*. Tilburg: Katholieke Universiteit Brabant.
- Meichenbaum, D. (1977). *Cognitive Behavior Modification*. New York: Plenum Press.
- Merriënboer, J.J.G. van & Jelsma, O. (1988). The Matching Familiar Figures Test: Computer or Experimenter Controlled Administration? *Educational and Psychological Measurement*, 48.
- Nunnally, J.C. (1967). *Psychometric Theory*. New York: McGraw-Hill.
- Revelle, W. (1987). Personality and Motivation: Sources of Inefficiency in Cognitive Performance. *Journal of Research in Personality*, 21, 436-452.
- Simons, P.R.J. & Liew-On, M. (1989). *Breadth of Orientation: Individual difference and Training*. Paper presented at the Conference of the European Association on Learning and Instruction (EARLI), Madrid.
- Soppe, H.J.G. (1979). Impulsiviteit-reflexiviteit: denken en doen in ontwikkeling. In: W.Koops & J.J. van der Werf (Red.). *Overzicht van de Ontwikkelingspsychologie*. Groningen: Wolters Noordhoff.
- Vijver, F. van de (1987). Computerondersteunde tests in de diagnostische praktijk. *De Psycholoog*, 22, 10-14.

Adres auteurs:

M.M. Liew-On
 P.R.J. Simons
 Katholieke Universiteit Brabant
 Postbus 90153
 5000 LE TILBURG