

Over de zin van replicatie-onderzoek

Commentaar op het artikel 'Het oplossen van redactie-opgaven: een experimenteel programma' van J. N. van den Berge-Scheijgrond en M. W. Bleek-Way

E. ASSINK,

*Psychologisch Laboratorium, Rijksuniversiteit
Utrecht*

N. VERLOOP

*CITO, afd. Onderzoek en Psychometrische
Dienstverlening, Arnhem*

Het geeft ons voldoening dat een door ons in 1977 gepubliceerd artikel geleid heeft tot een replicatie-onderzoek door Van den Berge en Bleek (VdB & B). Het artikel waarin over dit onderzoek wordt gerapporteerd is voor ons echter aanleiding tot het maken van een aantal kanttekeningen. Wij willen ons daarbij beperken tot essentiële zaken en zullen niet vervallen in allerlei mugeziffige op- en aanmerkingen over schoonheidsfoutjes in dit replicatie-onderzoek.

In de inleiding wordt, onder verwijzing naar Lykken (1968), door de auteurs gesteld dat het hier om een *operationele* replicatie gaat: Lykken zelf (1968, p. 155) definieert dit als een replicatie-onderzoek waarin 'one strives to duplicate exactly just the sampling and experimental procedures given in the first author's report of his research'. Wanneer we nu van de auteurs horen (par. 1.1.) dat zij veranderingen hebben aangebracht in 'de grootte van de onderzoeksgroep, de klassesituatie, het experimentele programma en de meetinstrumenten', dan is het zonder meer duidelijk dat we hier *niet* met een operationele replicatie te maken hebben; verwachtingen die men redelijkerwijs van een operationele replicatie kan hebben zijn in dit geval dan ook ongegrond. Door de omvang van de wijzigingen die de auteurs hebben aangebracht komen zij veel meer in de buurt van wat Lykken (p. 155 e.v.) een *constructieve* replicatie noemt. Kort gezegd houdt dit in dat de onderzoeker die de replicatie uitvoert veel meer vrijheid heeft om naar eigen inzicht verbeteringen in de treatment, meetprocedures en dataverwerking aan te brengen.

Op zichzelf is de keuze voor een dergelijke vorm van replicatie uitstekend te verdedigen;

aan het eind van ons eigen onderzoeksverslag signaleerden wij in onze opzet een aantal belangrijke tekortkomingen, waardoor het zelfs weinig zinvol zou zijn het onderzoek nog eens te repliceren op de wijze zoals bij een operationele replicatie gebeurt. Genoemde tekortkomingen zouden immers in een constructieve replicatie voorkomen kunnen worden. Bij de bespreking van onze resultaten stelden wij destijds dat het door ons ontwikkelde experimentele programma op vier punten zou moeten worden bijgesteld:

1. De overgang van het concreet aangeboden probleem naar een abstracte representatie daarvan zou zorgvuldiger en uitvoeriger in de lessen moeten worden uitgewerkt.
2. Er zou meer aandacht moeten worden besteed aan het verbaal leren omgaan met de algemene denkschema's die worden aangeboden, zodat kinderen door deze verbale explicitering zich de eigenschappen van deel-geheel relaties beter bewust worden en deze leren koppelen aan het opstellen van vergelijkingen.
3. Het programma zou, mede als gevolg van de bovenstaande punten, een langere tijdsperiode moeten beslaan.
4. Het programma zou op een grotere groep leerlingen moeten worden uitgeprobeerd.

Het verbaast ons nu dat VdB & B, die toch een fors aantal veranderingen in het onderzoek hebben aangebracht, van deze aanbevelingen alleen de laatste ter harte hebben genomen. (De overgang van concreet naar abstract blijkt nog steeds onvoldoende zorgvuldig te zijn uitgewerkt). Dat is jammer, niet alleen omdat VdB & B (evenals wij destijds) geen statistisch aantoonbare effecten vinden, maar vooral ook omdat hierdoor de discussie over het kernidee dat aan de eerste twee aanbevelingen ten grondslag ligt uit het vizier verdwijnt. Wij doen hier op het idee van 'ontwikkeld onderwijs'. Beter onderwijs moet vooral gericht zijn op het verhogen van de leerpotenties bij de leerlingen door hen waardevolle, breed toepasbare, algemene denkschema's bij te brengen, kortom systemen van theoretische begrippen. Het

deel-geheel schema voor het oplossen van redactieopgaven is zo'n algemeen cognitief schema.

Bij het schrijven van een programma om dit aan te leren zal de grootste zorg moeten worden besteed aan opbouw en uitwerking, omdat de *aard van de sturing van het onderwijsleerproces* naar onze mening dé doorslaggevende voorwaarde is voor het bereiken van goede resultaten. Het is immers bekend dat met name de zwakke leerlingen zelden belangrijke fundamentele verbanden in leerstofinhouden zelf ontdekken. Helaas hebben VdB & B juist dit punt laten liggen.

Het mag dan ook geen verbazing wekken dat het bijstellen van ons programma op secundaire punten slechts heeft geleid tot de te verwachten (artificiële) verhoging van de correlaties tussen toetsen en rapportcijfers.

Men kan zich afvragen wat het nut is van dergelijk 'replicatie-onderzoek', waarin enerzijds op een aantal punten wordt afgeweken van de oorspronkelijke opzet, maar anderzijds de kansen onbenut worden gelaten om het effect te onderzoeken van essentiële verbeterin-

gen.

Tot slot willen we nog ingaan op de opmerking van VdB & B dat wij de resultaten van onze toetsen op de verschillcores hebben genegeerd of althans 'slechts genoemd en niet als zodanig in de conclusies betrokken' (par. 3); in par. 4 stellen de auteurs zelfs: 'Zouden zij (d.i. Assink en Verloop) de toetsing door middel van verschillcores als belangrijkste resultaat aangemerkt hebben, dan zou voor de beide onderzoeken gelden dat het experimentele onderwijsprogramma niet aan de verwachtingen heeft voldaan'. Dat dit verwijt volkomen onterecht is kan het best worden aangetoond met een citaat uit de laatste paragraaf van ons artikel (p. 141): 'Wij menen dat de meest "kritische" toets, de toets op de verschillcores het best als de definitieve toetsing van de door ons bestudeerde effecten kan worden beschouwd. Aangezien deze toetsing resultaten opleverde die, hoewel zij duidelijk in de richting van onze vermoedens wezen, niet significant waren, mogen wij niet stellen dat de verschillen tussen de condities A, B en C statistisch aangetoond zijn'.