

ONDERWIJS RESEARCH

INLEIDING

De studie van „experimental designs”, i.e. van modellen van experimenteel onderzoek is een legitiem onderdeel van een empirische onderwijskunde. Uitvoering van experimenten en interpretatie van experimentele resultaten vereist naast kennis van experimentele modellen een zeker fundamenteel inzicht in de beginselen van de statistiek. Het gevaar van oppervlakkige interpretatie van experimentele resultaten blijkt duidelijk in twee vrij recente publicaties. \*

Een bevredigend antwoord op de in een experiment gestelde vragen is alleen mogelijk wanneer de onderzoeker een experimenteel model kiest (of ontwerpt) waardoor het hem mogelijk wordt de in het experiment aanwezige variantie te controleren (*Variantie* is hier gebruikt in statistische zin. In algemene termen kan gezegd worden dat de in een experiment werkzame factoren (variabelen) hun invloed kenbaar maken door beïnvloeding van de grootte van de variantie in de observaties (e.g. test scores, etc.).

De onderzoeker zal in het algemeen trachten:

- a. de variantie als gevolg van de tot zijn hypothesen relevante variabelen zo groot mogelijk te maken;

\* *Rubbens* (Geprogrammeerd Onderricht en Onderwijsresearch, Eindhoven, 1965) spreekt zijn verwondering uit over het feit dat *Silberman* wantrouwend is t.o.v. de geldigheid van bepaalde researchstudies waarin nergens significante verschillen blijken (p. 87.) Echter, het al of niet bereiken van significantie in de toetsing van hypothesen garandeert noch validiteit noch invaliditeit van een studie. In tegenstelling tot *Rubbens'* verzwegen vooropstelling is het juist zelfs zo dat niet correct opgezette, en daardoor geïnvallideerde, experimenten door verwarring van het effect van diverse variabelen kunnen falen significante resultaten te leveren. *Van Calcar* (Leren Lezen; Groningen, 1967) heeft verwante problemen met statistische inferentie. M.b.t. tabellen 9.9, 9.10 (blz. 115 e.v.) en de erbij behorende uitleg kan gesteld worden, dat, wanneer een verschil tussen groepen geen statistische significantie bereikt de onderzoeker maar één ding kan concluderen: de verschillen zijn het gevolg van kansfuncties. Rubbiken als „beter maar niet significant” zijn in het gemaakte van de door hem gehanteerde statistische inferentie betekenisloos. Overigens valt niet uit te maken of hij eenzijdige toetsen gebruikt met een 5 % significantie peil, 2-zijdige met een 5 % significantie peil of een 2-zijdige met in feite een 10 % significantie peil. (Afgezien van deze en andere soortgelijke kritiek mogen we blij zijn met onderzoekers als Dr. Van Calcar die een duidelijk empirische inslag vertonen.)

- b. de variantie tengevolge van allerlei toevallige factoren, invloeden, te controleren;
- c. de variantie als gevolg van toevallige individuele verschillen tussen proefpersonen en van meetfouten (vaak samen genomen als 'error variance') zo klein mogelijk te maken.

In de onder b) genoemde vereiste voor een succesvol experiment schuilen met name tal van moeilijkheden, tal van bronnen van (interne) invalidatie van het experiment. In principe kan de onderzoeker op 3 manieren aan deze vereiste voldoen:

- a. mogelijk verwarrende variabelen te elimineren;
- b. zorgen voor een a-selectieve (kans)verdeling van de invloed van zulke variabelen („randomizing”) en/of
- c. het onderzoek zo opzetten dat de invloed van deze opzichzelf niet belangrijke variabelen gemeten of geschat kan worden.

Modellen van experimenteel onderzoek verschillen in de mate waarin deze diverse controle technieken aan de onderzoeker ter beschikking staan. Het kiezen of ontwerpen van een model van onderzoek dat gezien het experimentele probleem, de experimentele omstandigheden, de doelstellingen van de experimentator en de beschikbare statistische hulpmiddelen optimaal geschikt is, vormt een eerste voorwaarde voor het verkrijgen van betekenisvolle researchresultaten.

Bij deze keuze van onderzoeksmodel (een term die dus dient als vertaling van „research design”) dient de onderzoeker speciaal aandacht te schenken aan potentiële bronnen van invaliditeit van zijn experiment. In een inmiddels klassiek geworden artikel maken *Donald T. Cambell* en *Julian Stanley* onderscheid tussen interne (in)validiteit en externe (in)validiteit. (Dit artikel is herdrukt in boekvorm: *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally, 1966, 84 blz.). *Interne validiteit* heeft te maken met de vraag in hoeverre de uitkomst van het experiment inderdaad kan worden toegeschreven aan de experimentele behandeling (e.g. zelfstudie vs. geleide studie). „Is het mogelijk dat de resultaten veroorzaakt werden door andere, onbekende en/of oncontroleerbare omstandigheden?” *Externe validiteit* houdt verband met de vraag in hoeverre de onderzoeksresultaten generaliseerbaar zijn.

*Cambell* en *Stanley's* beschrijving van mogelijke bronnen van invaliditeit en van onderzoeksmodellen die gebruikt kunnen worden deze bronnen van ongeldige resultaten te elimineren vormt de meest belang-

rijke bron van informatie op voor de „experimentele leek” leesbaar peil. Kort geleden verscheen een artikel van *Glenn H. Bracht* en *Gene V. Glass* („The external validity of Experiments”, *American Educational Research Journal*, V(4), November 1968) waarin de potentiële bronnen van externe validiteit die bij *Cambell* en *Stanley* niet helemaal uit de verf waren gekomen gedetailleerd zijn uitgewerkt.

Ik maak gebruik van het voorrecht van de kroniekschrijver van belang zijnde gegevens simpelweg door te geven, zonder (uitvoerige) interpretatie, om *Cambell* en *Stanley's* overzicht van bronnen van interne validatie en *Bracht* en *Glass'* overzicht van bronnen van externe invalidatie in het kort weer te geven. (Meer opmerkingen m.b.t. interne validatie en actuele voorbeelden zijn te vinden in: Dr. Gene Glass, Dr. Melvin P. Robbins, m.m.v. Jaap Tuinman: „Delacato- Neurologische Organisatie en Leesonderwijs. Een Kritiek.” Te verschijnen in *Pedagogisch Forum*, begin 1970). Gezien het voornemen van de redactie van P.S. om in de toekomst meer aandacht te gaan schenken aan researchverslagen kunnen we binnenkort allerlei terminologische verschillen tussen onderzoekers verwachten. Mijn vertalingen van researchjargon zijn echter zeer voorlopig en niet bedoeld als een definitief voorstel voor terminologische standaardisering.

### *Interne invaliditeit*

De volgende factoren kunnen de interne validiteit van een experiment nadelig beïnvloeden. *Wanneer de juiste controle afwezig is, kan hun invloed nl. gemakkelijk verward worden met het effect van de experimentele variabele(n).*

1. *Tussentijdse externe voorvallen*: („history”): gebeurtenissen tussen de eerste en volgende waarnemingen (bijv. tests) die onbedoeld invloed uitoefenen op de experimentele eenheid (e.g. klas, individu).
2. *Natuurlijke verandering in prestatie als gevolg van tijdsverloop*: („Maturation”): biologische, fysiologische en psychologische processen die een functie van tijd zijn en niet het gevolg van specifieke gebeurtenissen buiten de proefpersonen. ‘Rijping’ valt binnen het geheel van deze factoren. Andere voorbeelden: ouder worden, vermoeid raken, sterker worden.
3. *Herhaald testen*: het effect van het nemen van een test op de uitslag van een volgende test.
4. *Inconsequent gebruik van meettechnieken*: onbedoelde variaties in het gebruik van meettechnieken (in de ruime zin die zowel tests, beoordelaars, als andere psychologische meettechnieken omvat)

kan leiden tot verandering in scores die worden aangezien voor effect van de experimentele variabele.

5. *Statistische regressie*: In onderwijsresearch beschikken we niet over volkomen betrouwbare meetinstrumenten. Weinig schooltoetsen, enz. hebben een betrouwbaarheidscoëfficiënt boven 95 en veel vaker moet de onderzoeker gebruik maken van veel minder betrouwbare instrumenten. Dit heeft o.a. tot gevolg het verschijnsel van regressie naar het gemiddelde. D.w.z.: proefpersonen die bij een eerste meting uitzonderlijk hoog of laag scoorden worden verwacht — in dit geval als gevolg van statistische overwegingen — bij een tweede meting minder uitzonderlijk (hoog of laag) te scoren. Zulke veranderingen in *waargenomen* scores vertegenwoordigen geen verandering in *werkelijke* prestatie.
6. *Selectie van proefpersonen*: het is duidelijk dat verschillen in prestaties op de afhankelijke variabele het gevolg kunnen zijn van vóór-experimentele verschillen tussen (bijv.) controle en experimentele groep. Dergelijke verschillen ontstaan door een niet a-selectieve (of „onbevooroordeelde”) keuze van proefpersonen.
7. *Experimentele uitval*: („experimental mortality”): Vaak vallen gedurende een experiment bepaalde proefpersonen uit. Als die uitval niet wordt bepaald door toevalsfactoren (zodat het netto effect theoretisch gelijk is voor experimentele en controle groepen) maar het gevolg van systematische factoren (zoals het feit dat de experimentele groep een meer inspannende taak moet volbrengen) wordt het oorspronkelijke evenwicht tussen controle en experimentele groep(en) verstoord.

In aanvulling van deze 7 potentiële bronnen van interne invaliditeit bestaat dan nog het gevaar van een interactie tussen selectie en sommige of de andere factoren die eveneens tot gevolg heeft dat de experimentele resultaten moeilijk interpreteerbaar worden.

#### *Externe invaliditeit*

Een experiment dat geen interne validiteit bezit kan ons niets leren. De resultaten zijn oninterpreteerbaar. Het begrip externe validiteit heeft een wat meer relatief karakter. Een experiment dat minimum externe validiteit bezit kan zeer belangrijke informatie verstrekken. Generalisatie van die informatie tot andere proefpersonen en omstandigheden is in dat geval echter maar erg beperkt. De onderzoeker moet zijn conclusies beperken tot de situatie waarin hij het experiment uitvoerde.

Bracht en Glass (1968) onderscheiden twee hoofdgroepen van factoren die de externe validiteit kunnen beïnvloeden: (1) die factoren die verband houden met generalisatie van de experimentele uitkomsten naar populaties van personen; (2) die verband houden met de totale opzet van het experiment: de omstandigheden waarin geëxperimenteerd werd, de experimentele behandeling, de variabelen, enz. De eerste groep factoren hangt samen met *populatie validiteit*, de tweede met *ecologische validiteit*: „Naar welke situaties kunnen mijn resultaten ge-generaliseerd worden?”

Bronnen van moeilijkheden voor:

A. *Populatie validiteit*:

1. *Beschikbare vs. Verlangde populatie*: in veel gevallen kan niet ongestraft gegeneraliseerd worden van de populatie die de experimentator ter beschikking staat (bijv. „10-jarigen in Zwolle”) naar de populatie waar hij uiteindelijk in geïnteresseerd is (e.g. „10-jarigen in Nederland”).
2. *Wisselwerking (interactie) tussen persoonsvariabelen en effect van de experimentele behandeling*. Dit verschijnsel kan worden geïllustreerd als volgt. Stel dat behandeling A inhoudt: een paragraaf lezen en dan uitleg krijgen en behandeling B: eerst uitleg, dan lezen. Een test voor begrip van de leerstof wordt gegeven en uit de statistische analyse van de resultaten blijkt dat A beter werkt voor hoog-intelligente en B voor laag-intelligente proefpersonen. In dit geval spreekt men van interactie tussen persoonsvariabelen (als intelligentie) en de behandeling. Het niet herkennen van deze wisselwerking leidt gemakkelijk tot generalisaties van conclusies betreffende het effect van de behandeling tot verkeerde populaties.

B. *Ecologische validiteit*:

1. *Gedetailleerde beschrijving van de onafhankelijke variabele*. Herhaling en generalisatie van een experiment hangt af van uitvoerige en betrouwbare beschrijving van de details van de behandeling en de variabelen.
2. *Gekruist effect van meervoudige experimentele behandeling*: Wanneer eenzelfde groep personen meer dan één behandeling ondergaat wordt het vaak onmogelijk de oorzaak van het experimentele resultaat aan te wijzen (hier raken we de interne validiteit) of te generaliseren naar omstandigheden waar maar één behandeling plaats vindt.
3. *Hawthorne-Effect*: het gedrag van de proefpersoon wordt beïnvloed door zijn besef aan een experiment deel te nemen. Dit be-

leemert generalisatie maar niet-experimentele situaties.

4. *Het effect van nieuwigheid en verstoring van normale gang van zaken.* Het effect van een bepaalde experimentele behandeling in een omgeving waarin deze behandeling iets „nieuws” vertegenwoordigt kan niet zomaar gegeneraliseerd worden naar omstandigheden waarin dit niet het geval is.
5. *Het effect van de onderzoeker:* generalisaties naar andere situaties kunnen mislukken omdat in het oorspronkelijke experiment de proefpersonen (onbewust) beïnvloed waren door de (aanwezigheid van) de onderzoeker.
6. *Effect van de na-test.* Het effect van de behandeling kan latent zijn en pas tot uiting komen wanneer de proefpersonen een na-test krijgen. Generalisaties van het behandelingseffect naar situaties waarin zo'n na-test ontbreekt is in dat geval dus niet zonder meer verantwoord.
7. *Rol van de afhankelijke variabele:* generalisatie van de resultaten hangt nauw samen met de keuze van afhankelijke variabele(n) en van de instrumenten waarmee dat effect gemeten wordt.

Ik heb enkele van de door de *Bracht* en *Glass* opgesomde bronnen van externe invaliditeit weggelaten omdat ze n.m.b.m. thuishoren in de bespreking van interne invalidatie en daar reeds aan de orde zijn gesteld.

#### *Samenvattend*

„Kan het in het experiment geobserveerde effect misschien verklaard worden door andere, niet relevante, factoren?” „In hoeverre zijn mijn resultaten generaliseerbaar?” Dat zijn de twee vragen waarop de onderzoeker antwoord moet geven; het grootste deel van de literatuur betreffende onderzoeksmodellen en statistische hulpmiddelen voor de onderwijsresearch behandelt vraagstukken die samenhangen met het geven van een zo correct mogelijk antwoord erop. Hierboven werd een beknopt overzicht van enkele facetten van die twee vragen gegeven.

#### *Literatuur m.b.t. Onderzoeksmodellen en relevante statistische technieken*

In onderwijsresearch kringen in de Amerikaanse universiteiten bestaat de opvatting dat zelfs nu nog, na een ruim 50 jarige experimentele traditie, naïviteit in keuze van onderzoeksmodel en statistische analyse nog maar al te vaak overheerst. Met name door deskundigen in een be-

paald didactisch specialisme (zoals leesonderwijs) zelfstandig — d.w.z. zonder behulp van een speciaal getrainde researchspecialist — uitgevoerde onderzoeken vertonen vaak ernstige methodische tekortkomingen.

Als gevolg van voortschrijdende teleurstelling in de kwaliteit van de onderwijsresearch, zijn de pedagogische faculteiten in de V.S. de laatste jaren zeer bewust gaan concentreren op onderzoeksmethodologische aspecten. Het resultaat van deze instelling is een groot aantal publikaties bedoeld als hulp bij de introductie van de pedagogiek student in deze materie.

Wat nu volgt is een persoonlijk gekleurde selectie uit de overvloed van literatuur. Enkele, niet direct voor de pedagogiekstudent geschreven, maar voor hem toch belangrijke werken zijn eveneens opgenomen. Boeken met voornamelijk psychometrische inhoud zijn achterwege gelaten. Zij vormen een klasse apart, hoewel er natuurlijk nauwe relaties tussen testen en experimenteren bestaan. Volledigheid is niet nagestreefd; wel is geprobeerd de allerbelangrijkste werken in te sluiten.

#### *Statistische methoden*

1. BRADLEY, JACK I. en JAMES N. MCLELLAND. *Basic Statistical Concepts*. Chicago: Scott, Foresman en Co., 1963, 168 pp.
2. COOLEY, W. en P. LOHNES. *Multivariate Procedures for the Behavioral Sciences*. New York: Wiley, 1962.
3. DIAMOND, S. *Information and Error*. New York: Basic Books, 1959.
4. DOWNIE, N. M. en R. W. HEATH. *Basic Statistical Methods*. 2e druk, London: Harper and Row, 1965. \*
5. DUBOIS, PHILIP H.. *An Introduction to Psychological Statistics*. London: Harper and Row, 1965, 530 blz.
6. EDWARDS, A. *Statistical Analysis*, Herziene uitgave, London: Holt, Rinehart and Winston, 1958, 234 blz.
7. FRUCHTER, B. *Introduction to Factor Analysis*. Princeton, N. J.: Van Nostrand, 1954, 280 blz.
8. GOROW, FRANK F. *Statistical Measures: A Programmed Text*. San Francisco: Chandler Publishing Company, 1962.
9. GUENTHER, WILLIAM C. *Concepts of Statistical Inference*. London: McGraw-Hill Book Co., 1965, 353 blz.

\* Van uitgaven door uitgevers gevestigd zowel in de V.S. als in Londen wordt Londen als plaats van uitgave vermeld, hoewel de auteurs doorgaans Amerikaanse professoren zijn.

10. GUILFORD, J. P. *Fundamental Statistics in Psychology and Education*. 3e druk, London: McGraw-Hill, 1956, 561 blz.
11. HAYS, W. *Statistics for Psychologists*. London: Holt, Rinehart and Winston, 1963.
12. KRAFT, CHARLES H. en CONSTANCE VAN EEDEN. *A Non-Parametric Introduction to Statistics*. London: Collier-McMillan, Ltd, 1968, 342 blz.
13. MORRISON, DONALD F. *Multivariate Statistical Analysis*. London: McGraw-Hill, 1967, 338 blz.
14. MCNEMAR, Q. *Psychological Statistics*. 2e druk. New York: Wiley, 1955.
15. POPHAM, JAMES W. *Educational Statistics—Use and Interpretation*. London: Harper and Row, 1967.
16. SIEGEL, S. *Non-parametric Statistics for the Behavioral Sciences*. London: McGraw-Hill, 1956.
17. TATE, MERLE W. *Statistics in Education and Psychology—A First Course*. London: Collier MacMillan Ltd., 1965, 355 blz.
18. WEINBERG, GEORGE H. en SCHUMAKER, JOHN A. *Statistics—an intuitive approach*. Belmont, California: Wadsworth Publishing Company, Inc.: 1962, 338 blz.

#### Toelichting:

In bovenstaande lijst zijn een aantal werken voor beginners opgenomen. De reden daarvoor, terwijl we in ons land het werk van *De Jonge en Wielenga* (Groningen, 1953) hebben, is voornamelijk dat de onderwijsresearcher die op de hoogte wil raken van huidige ontwikkelingen op dit terrein ermee gediend is de Anglo-Saksische terminologie te beheersen. Verder is mij alleen de 1953 uitgave *De Jonge-Wielenga* bekend en die is nu enigszins verouderd.

No's 1 en 8 zijn geprogrammeerde teksten; nuttig alleen als vingeroefening, enkele basis begrippen nodig voor het lezen van researchverslagen worden behandeld.

No's 4, 6, 17 en 18 zijn alle vier goede introducties in de eenvoudigste grondbeginselen van statistische berekening en interpretatie van statistische grootheden. *Edwards* is sterk in het in een paar zinnen raken van de kern van de onderliggende gedachtengang; *Downie* en *Tate* hebben het voordeel oefeningen met antwoorden te geven, terwijl *Weinberg* grotere nadruk dan de andere drie legt op het begrijpen van de basis van statistische inferentie.

No's 9, 10 en 15 zijn eveneens tekstboeken van een inleidend karakter, maar een tikkeltje moeilijker geschreven dan de voorgaande



boeken. *Guilford* begint eenvoudig maar behandelt de gebruikelijke onderwerpen zeer grondig zodat hij vrij lastig te lezen is. *Guenther's* boek is vooral waardevol vanwege de ongebruikelijk uitgebreide verzameling statistische tafels achterin; verder veel voorbeelden, opgaven. Een voordeel van *Popham's* werk is het feit dat van elk onderwerp (zoals de t-test) éérst de logische achtergrond gegeven wordt en daarna in een apart hoofdstuk de wijze van berekening van de betreffende statistische grootheid wordt uiteengezet.

No 3 kan dienen als aanvulling bij elk van de hierboven besproken zeven inleidende teksten. Zeer aanbevolen vanwege originele kijk op diverse statische problemen.

No's 5, 11 en 14 geven een behandeling van fundamentele statistische methoden op een wat hoger peil. *McNemar's* werk is ongetwijfeld het meest bekend; nuttig als naslagwerk voor de researcher. *DuBois'* boek reflecteert de nieuwe ontwikkelingen in statistische methoden in onderwijsresearch door een hoofdstuk over matrix algebra en over factor analyse. Problemen (met antwoorden in losse handleiding) aanwezig. *Hays'* boek is eveneens een uitstekende introductie tot de wat meer complexe statistische methoden. Al deze boeken zijn geschikt voor de wat gevorderde student. In deze groep zouden ook *Edwards* (1968), *Lindquist* (1953) en *Winer* (1962) genoemd kunnen worden (zie research-bibliografie).

No 16 (*Siegel*) is voor lange tijd de autoriteit op het gebied van non-parametrische statistiek geweest in pedagogiekringen. Zijn uitstekend boek is onlangs aangevuld door het enigszins theoretischer opgezette werk van *Kraft* en *Van Eeden*, vooral waardevol ook door naar mijn weten niet eerder gepubliceerde tafels van diverse non-parametrische distributies. Wie veel met kleine steekproeven werkt of om andere redenen behoefte heeft aan non-parametrische toetsen zal beslist beide werken willen bezitten.

No's 2 en 13 zijn van interesse voor de onderzoeker die behoefte heeft aan methoden om de resultaten van experimenten waarin waarnemingen op meer dan één afhankelijke variabele verricht worden te analyseren. *Cooley's* boek is minder uitgebreid, maar is waardevol vanwege beschrijvingen van computerprogramma's voor het uitvoeren van diverse toetsen.

*Fruchter's* boek tenslotte (no 7) is de beste en eenvoudigste inleiding tot factor analyse, hét stokpaardje van veel onderwijsresearchers tegenwoordig, die mij bekend is.

*Researchmethoden en -modellen.* \*

1. COCHRAN, W. G. en GERTRUDE COX. *Experimental Designs*. 2e druk, London: Wiley, 1957, 611 blz.
2. COOK, DAVID R. *A Guide to Educational Research*. Boston: Allyn and Bacon, 1965, 231 blz.
3. COX, D. R. *Planning of Experiments*. London: Wiley, 1958.
4. EDWARDS, A. *Experimental Design in Psychological Research*. 3e druk, London: Holt, Rinehart en Winston, 1968, 455 blz.
5. FISHER, R. A. *The Design of Experiments*. 6e druk, New York: Hafner.
6. GALFO, ARMAND J. en EARL MILLER. *Interpreting Education Research*. Dubuque, Iowa: Wm. C. Brown Co. Publishers, 1966, 367 blz.
7. KIRK, ROGER E. *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, Calif.: Brooks-Cole Publishing Co., 1968, 577 blz.
8. KERLINGER, FRED N. *Foundations of Behavioral Research*. London: Holt, Rinehart and Winston, 1964, 739 blz.
9. LATHROP, RICHARD G. *Introduction to Psychological Research. Logic, Design, Analysis*. London: Harper and Row, 1969, 297 blz.
10. LINDQUIST, E. *Design and Analysis of Experiments*. Boston: Houghton Mifflin, 1953.
11. TOLBERT, E. L. *Research for Teachers and Counselors*. 2e druk, Minneapolis: Burgess Publishing Co., 1968, 100 blz.
12. TRAVERS, ROBERT M. W. *An Introduction to Educational Research*. 2e druk, London: Collier-McMillan Ltd., 1964.
13. WEBB, EUGENE J., et al. *Unobtrusive Measures — Nonreactive Research in the Social Sciences*. Chicago: Rand McNally, 1966.
14. UNDERWOOD, B. *Psychological Research*. New York: Appleton, 1957, 300 blz.
15. WINER, BEN J. *Statistical Principles in Experimental Design*. London: McGraw-Hill Book Co., 1962, 672 blz.

De in deze lijst opgenomen werken vallen uiteen in een aantal duidelijk gescheiden categorieën.

No's 2, 11, 12 en 14 zijn allemaal introducties in de algemene problematiek en methoden van onderzoek die geen statistische kennis voor onderstellen. De potentiële research-producent heeft weinig aan

\* Zie vooral ook: Cambell en Stanley.

no's 2 en 11: te summier voor hèm, maar een bruikbaar begin voor de research-consument. No. 12 geeft een grondiger introductie in principiële vraagstukken in en methoden van onderwijsresearch; 14 behandelt zeer bekwaam een aantal fundamentele aspecten van het experiment en gaat in op het wezen van de logica van het experimenteren.

*Webb* et al (no. 13) suggereren talrijke originele oplossingen voor de moeilijkheid van beïnvloeding van het waargenomene door de waarnemer. Zeer interessant.

No's 6, 8 en 9 zijn eveneens introducties in onderwijsresearch maar met meer nadruk op statistische aspecten. *Galfo* en *Lathrop* zijn misschien wat gemakkelijker te lezen, maar *Kerlinger* lijkt mij verreweg de beste koop: systematisch, uitgebreid en grote nadruk op onderzoek-logica. Enige kennis van de statistiek vergemakkelijkt het bestuderen van dit uitstekende boek aanzienlijk.

No's 4, 7, 10 en 15 behandelen statistische analyses van de diverse onderzoeksmodellen zoals factoriële experimenten, Latijnse vierkanten, enz. Ze horen alle vier thuis op de werktafel van de onderzoeker, maar *Winer* geeft misschien het meeste directe rendement: zijn boek is een soort kookboek voor experimentele modellen (zeer uitvoerig!). *Edwards* leest het gemakkelijkst en vereist wellicht de minste staisische voorbereiding. *Lindquist* is voornamelijk aanbevolen vanwege zijn benadrukken van de achtergrond van de diverse designs; moeilijke lectuur. *Kirk* lijkt me o.a. goed als achtergrond literatuur voor de gebruiker van *Winer*.

No's 1, 3 en 5 zijn niet direct geschreven voor onderzoek van onderwijs- of psychologische problematiek. Voor het verkrijgen van een beter inzicht in de functie van een goed experimenteel model zijn deze werken onmisbaar. Enige kennis van de statistiek is wel vereist. No. 5 is natuurlijk een van *Sir Ronald Fisher's* klassieke bijdragen tot de statistiek. Zijn logica is nog altijd actueel.

JAAP TUINMAN

University of Georgia, V.S.