

Repliek op Van de Grifts 'Onderzoek naar het functioneren van de onderwijsbegeleiding op basisscholen'

G. VAN DEN BERG, E. HARKAMP, J. PRINS, P. WOLFGRAM
RION – Instituut voor
Onderwijsonderzoek Rijksuniversiteit
Groningen

Samenvatting

Van de Grifts kritiek is gebaseerd op twee hoofdpunten: 1. er zou in het onderzoek een onzorgvuldig onderscheid zijn gemaakt tussen wel en niet begeleide scholen, en 2. onderzoeksinstrumenten zouden non-valide zijn.

Wat betreft punt 1 zij opgemerkt dat het onderzoek zich heeft geconcentreerd op die scholen die begeleiding kregen van een OBD. Toch hebben wij variantie-analyses uitgevoerd om na te gaan of de verschillende categorieën van scholen die wel of niet worden begeleid op relevante variabelen van elkaar verschillen. Deze analyses laten zien dat de verschillen tussen deze groepen scholen te verwaarlozen zijn.

Wat betreft punt 2 zij opgemerkt dat de instrumenten in redelijke mate voldoen aan verschillende vormen van validiteit, te weten: begripsvaliditeit, discriminant-validiteit en inhoudsvaliditeit. Van de Grift probeert de non-validiteit van de drie belangrijkste onderzoeksinstrumenten aan te tonen. Hij gebruikt hiervoor drie argumenten, die niet steekhoudend blijken te zijn:

- a. *De correlatie tussen twee op elkaar lijkende variabelen als een maat van soortgenoot validiteit is naar onze mening niet toegestaan als de paren van variabelen niet hetzelfde object van bevraging hebben.*
- b. *Correlaties tussen twee versies van dezelfde itemlijsten zijn geen correcte maat voor validiteit, omdat het object van be-*

vraging voor beide respondentgroepen verschillend is:

- c. *Het gebruik van de frequentieverdeling zoals Van de Grift doet, is als maat voor validiteit niet aanvaardbaar.*

Het geheel van argumenten overziend kan de conclusie alleen maar zijn dat de kritiek van Van de Grift wetenschappelijk niet adequaat is.

Alvorens in te gaan op de kritiek van Van de Grift op het RION-onderzoek naar het functioneren van de onderwijsbegeleiding op basisscholen (zie dit nummer p. 413-419) spreken wij onze erkentelijkheid uit over het feit dat Van de Grift zich niet beperkt tot de samenvatting maar kennis neemt van het hele rapport en dus ook van de voorbehouden die de onderzoekers bij de onderzoeksresultaten maken. De methodologische kritiek van Van de Grift op het onderzoek is voor ons echter aanleiding te reageren op de volgende twee aspecten die in zijn kritiek centraal staan:

1. het doen van causale uitspraken op grond van resultaten uit een niet-experimenteel design en
2. de validiteit van de onderzoeksinstrumenten.

De gewraakte causale uitspraken

We zijn het volledig met Van de Grift eens dat men voorzichtig dient te zijn met het doen van causale uitspraken op grond van resultaten uit correlatieonderzoek. Zeer hardnekkig is het misverstand dat wij de niet gevonden verschillen tussen scholen die wel en niet worden begeleid interpreteren alsof begeleiding geen effect heeft gehad. Ook Van de Grift doet dit, daarbij verwijzend naar par. 5.6. van het eindrapport. Het ware mischien verstandiger geweest dit onderzoeksresultaat een minder prominente plaats te geven in het rapport. We hebben in ons onderzoek inderdaad geen verschillen kunnen

constateren tussen scholen die wel en die geen ondersteuning hebben gehad van een OBD. Met behulp van het onderhavige onderzoek kan niet worden uitgemaakt of scholen op eigen kracht evenveel bereiken als scholen die begeleiding ontvangen. Op pag. 82 van het eindrapport stellen wij: 'Het feit dat wij geen verschillen hebben kunnen vaststellen tussen scholen die wel en scholen die geen begeleiding ontvingen hoeft niet noodzakelijk te betekenen dat begeleiding geen effect heeft'. Uiteraard zullen er scholen zijn die op eigen kracht in staat zijn tot het realiseren van belangrijke effecten en uiteraard zullen er ook scholen zijn die daar ondanks intensieve begeleiding niet toe in staat zijn.

Het doel van het onderzoek was om na te gaan of de soort en hoeveelheid begeleiding invloed hebben op een aantal effectvariabelen. Om dit te kunnen nagaan zijn de 'niet-begeleide' scholen buiten de verdere analyse gehouden en hebben we ons beperkt tot de gegevens van scholen die wel begeleiding hebben ontvangen. Het zijn juist deze analyses die een prominente plaats hebben in het eindrapport. Van de Grifts conclusie dat er geen keurige scheiding is gemaakt tussen wel en niet begeleide scholen is dus onterecht.

Verder willen we opmerken dat in het onderzoek geen onderscheid is gemaakt tussen wel en niet begeleide scholen an sich, maar tussen wel en niet begeleide scholen op steeds één van de drie onderscheiden begeleidingsdomeinen. Dit betekent dat een school op het ene begeleidingsdomein wel en op andere begeleidingsdomeinen geen begeleiding kan hebben ontvangen. Hoewel we de 'niet-begeleide' scholen steeds buiten de analyses hebben gehouden, willen we hier toch proberen antwoord te geven op de door Van de Grift gestelde vraag of er verschillen zijn tussen de groepen wel en niet begeleide scholen. Daartoe hebben we de scholen ingedeeld in a) begeleiding ontvangen van een OBD, b) begeleiding ontvangen van een andere instantie, c) geen begeleiding omdat er geen behoefte aan was en d) geen begeleiding ontvangen, omdat zij geen veranderingen wensen. Uit de analyses blijkt dat er nauwelijks verschillen tussen deze vier groepen bestaan. Met behulp van variantie-analyse is nagegaan of deze scholen van elkaar verschillen op de zeven door ons gehanteerde effectvari-

abelen en dat steeds per begeleidingsdomein. Tussen de vier categorieën van scholen zijn slechts drie significante verschillen gevonden. Zo blijkt dat:

1. scholen ($n=8$) die wel leerlinggerichte hulp krijgen, maar niet van de OBD, een significant hoger doubleerpercentage hebben dan de andere drie groepen scholen,
2. op scholen ($n=14$) waar leerkrachten veranderingen op het gebied van leerinhouden en werkwijzen niet nodig vinden, significant traditioneler wordt lesgegeven dan op scholen die worden begeleid door een andere instantie dan de OBD. Deze groep scholen verschilt echter niet significant van de andere groepen van scholen,
3. op scholen ($n=88$) waar de leerkrachten zeggen geen hulp van de begeleider nodig te hebben bij het doorvoeren van veranderingen van lesinhouden en werkwijzen de leerlingen significant vaker blijken door te stromen naar hoger onderwijs dan op scholen die wel worden begeleid maar niet door een OBD. Er is echter geen significant verschil met de andere groepen van scholen.

Verder hebben wij geen verschillen gevonden tussen de vier categorieën scholen. Het vermoeden van Van de Grift dat de begeleide scholen op relevante kenmerken verschillen van de niet begeleide scholen wordt hiermede ontkracht.

Validiteit van de instrumenten

De argumenten van Van de Grift waarmee hij het door ons gebruikte instrumentarium naar de prullebak verwijst zijn teleurstellend te noemen. We zullen die hieronder toelichten.

De door ons ontwikkelde schalen voldoen aan drie kenmerken van validiteit, te weten begripsvaliditeit, discriminantvaliditeit en inhoudsvaliditeit. De interne betrouwbaarheid van de schalen is een indicator voor *begripsvaliditeit*. Deze is gemeten door analyse van de p-waarden, item-restcorrelaties en interne consistentie. Uiteraard dekt begripsvaliditeit niet alle vormen van validiteitseisen die aan onderzoeksinstrumenten moeten worden gesteld. Voor begripsvaliditeit zijn de gekozen maten goede indicatoren.

De *discriminant-validiteit* van de drie in-

strumenten is vastgesteld met behulp van principale componentenanalyse (met vari-max-rotatie). Deze is uitgevoerd op alle items uit de drie schalen. De analyse laat zien dat er drie factoren zijn te onderscheiden, die overeenkomen met de door ons onderscheiden schalen.

Ook zijn er duidelijke indicaties voor de *inhoudvaliditeit* van de schalen. Verreweg de meeste begeleiders blijken het merendeel van hun tijd te besteden aan de in de schalen aangegeven activiteiten. De activiteiten die in de drie schalen zijn ondergebracht dekken gemiddeld steeds zo'n 70% van de totale tijd die begeleiders aan de drie domeinen besteden. Voor de ene begeleider is de schaal 100% en voor de andere begeleider 40% dekkend.

Wat betreft de schaal 'leerlinggerichte begeleiding' zegt 87% van de begeleiders dat de activiteiten in de schaal voor meer dan 50% dekkend zijn voor de door hen ondernomen activiteiten op dit domein; 50% van de begeleiders zegt dat meer dan 75% van deze activiteiten dekkend zijn voor hetgeen ze doen. Wat betreft de schaal 'begeleiding van lesinhouden' zegt 72% van de begeleiders dat de activiteiten in deze schaal voor meer dan 50% dekkend zijn voor hun activiteiten op dit domein; 50% van de begeleiders zegt dat 75% van deze activiteiten dekkend zijn voor hun activiteiten op dit domein. Wat betreft de schaal 'begeleiding van de schoolorganisatie' zegt 75% van de begeleiders dat deze activiteiten meer dan 50% dekkend zijn voor hun activiteiten op dit domein; en 50% van de begeleiders zegt dat ze voor meer dan 80% dekkend zijn.

Van de Grift probeert aan te tonen dat de door ons gebruikte schalen niet concurrent valide zijn. Hij probeert dit op drie manieren aan te tonen. Allereerst probeert hij te argumenteren dat een correlatiemaat tussen de activiteitenschaal en de hoeveelheid tijd die begeleiders aan de activiteiten in de schaal besteden een goede maat is voor validiteit. De correlatie tussen de tijdschatter en de activiteitenschaal blijkt dan .50 te zijn. De correlaties tussen de tijdschatters en de andere activiteitenschalen zijn .57 en .32. Bij deze correlatieberekening willen we de volgende kanttekening plaatsen. Het aantal activiteiten dat een begeleider zegt uit te voeren gecorreleerd aan het percentage van de totale

tijd die hij aan een bepaald begeleidingsdomein besteedt is naar onze mening betekenisloos. Een begeleider kan zeggen weinig activiteiten te ondernemen die echter toch 100% van zijn tijd in beslag nemen. Anderzijds kan een begeleider veel activiteiten ondernemen die eveneens 100% van zijn tijd in beslag nemen. Er is dan ook geen reden een lineaire samenhang te verwachten.

Vervolgens meent Van de Grift de non-validiteit van de schalen te kunnen aantonen met behulp van een correlatie tussen twee versies van de schalen. Bij de interpretatie van deze correlatie gaat Van de Grift voorbij aan het feit dat de twee versies betrekking hebben op verschillende objecten. De leerkrachten is gevraagd de frequentie van de activiteiten van de begeleider in hun *klas* aan te geven. De schoolbegeleider is gevraagd dit te doen voor de gehele *school*. Het object van bevraging is dus geheel verschillend. Men dient dan ook buitengewoon voorzichtig te zijn op basis van een correlatie tussen deze twee metingen uitspraken te doen over de validiteit van deze schaal. Ook als twee metingen hetzelfde object betreffen dan is het nog maar de vraag of je een hoge correlatie mag verwachten. We zijn van mening dat een oordeel over de validiteit van deze schalen gebaseerd op de correlaties tussen beide versies niet geoorloofd is, omdat deze correlaties volstrekt geen indicatieve waarde hebben. Er zijn allerlei factoren die verschillen in percepties van verschillende personen in dezelfde situatie kunnen beïnvloeden, zoals selectieve attentie, emotionele betrokkenheid, perspectiefvertekening.

Van de Grift doet nog een derde poging om aan te tonen dat de schalen niet valide zijn. Hij meent op basis van frequentieverdeling van het gemiddelde item van de schaal iets te kunnen zeggen over de validiteit van deze schaal. Daartoe telt hij de door ons rijgewijs berekende percentages kolomsgewijs op. Op grond van deze berekening kan absoluut niets geconcludeerd worden over de validiteit van het instrument. Met deze berekening wil van de Grift aantonen dat wij nauwelijks voorkomende activiteiten hebben gemeten en dat er een grote discrepantie bestaat tussen de percepties van begeleiders en leerkrachten c.q. scholen. Maar so wie so is de validiteit van de schalen niet af te meten aan de mate waarin begeleiders deze activi-

teiten doen. Validiteit staat los van de verdeling van scores op een item-schaal.

Conclusie

De kritiek van Van de Grift is teleurstellend. Hij heeft geprobeerd aan te tonen dat de drie begeleidingsschalen (de ruggespraak van het onderzoek) non-valide zijn en dat derhalve ook de uitspraken uit het onderzoek ongeldig zijn. Dit is een zeer ernstige kritiek; een kritiek die niet houdbaar is, als we het geheel van argumenten en tegenargumenten overzien.

Summary

Van den Berg, G., E. Harskamp, J. Prins, P. Wolfgram. 'A reply to Van de Grift's criticism: Research on the functioning of the local educational support-systems.' *Pedagogische Studiën*, 1986, 63, 420-423.

Van de Grift's criticism is based on two main points: 1 the in Van de Grift's opinion incorrect difference between forms of (not) being counseled and 2 the validity of the most important research-instruments.

1. Our research is concentrated on those schools which did have assistance of the OBD's. An analysis of variance between different forms of (not) being counseled shows that the differences in the dependent variables are negligible.
2. Van de Grift tries to point out the non-validity of the three main-research-instruments. He uses three arguments which are, all things, considered not sound:
 - a. the correlation between two 'concurrent' variables as a measure of concurrent validity is in our opinion not allowed, because the pairs of variables didn't have the same object of questioning,
 - b. the correlation between two versions of the same item-list is not a correct measure of validity, because the object of questioning for the two respondentgroups was different,
 - c. the frequency-distribution of a mean-item of a scale is an absurd measure. When analysed in a usual way the instruments are content-valid.

The conclusion is that the criticism of Van de Grift is scientifically inadequate.

Curricula vitae

De auteurs zijn allen werkzaam bij het RION - Instituut voor Onderwijsonderzoek.

Adres: Postbus 1286, 9701 BG Groningen

Manuscript aanvaard 6-8-'86