Universiteit Leiden
The Netherlands

Tien kleine vergissingen en één grote: Reactie op Tien kleine negertjes, AP XI, p, 106

Heunks, F.J.

Citation
Heunks, F. J. (1976). Tien kleine vergissingen en één grote: Reactie op Tien kleine negertjes, AP XI, p, 106. *Acta Politica*, *11: 1976*(3), 406-408. Retrieved from https://hdl.handle.net/1887/3451639

Version:	Publisher's Version
License:	Leiden University Non-exclusive license
Downloaded from:	https://hdl.handle.net/1887/3451639

Note: To cite this publication please use the final published version (if applicable).

Literatuur

Tien kleine vergissingen en een grote reactie op M. E. Boon en B. Niemöller 'Tien kleine negertjes' (AP, blz. 106, 107)

door F. J. Heunks

Belangstelling voor mijn aliënatie-studie stel ik steeds op prijs. De reaktie van Boon en Niemöller vergt zelfs een repliek, die ik overigens graag meteen had willen geven, wanneer ik daartoe in de gelegenheid was gesteld.
Goed, dan eerst de kleine vergissingen:
1 – Ik trek in mijn proefschrift niet de kwaliteit van mijn schalen in het algemeen in twijfel, maar geef enige oorzaken van de lagere schaalbaarheid in het hoofdonderzoek, in vergelijking met de vooronderzoekingen (een daling van gemiddeld .45 naar .34).
2 – Deze genoemde oorzaken van geringere schaalbaarheid t.a.v. een bepaalde dataset kunnen natuurlijk nooit gesteld zijn als bewijs van betrouwbaarheid. Integendeel, de relatief hoge non-response op de items en daardoor hun geringere betrouwbaarheid in het hoofdonderzoek benadeelt juist de schaalbaarheid.
3 – Het is niet zo dat sommige van de schalen het in het vooronderzoek wat beter deden, maar bijna alle schalen deden het in de drie vooronderzoekingen aanzienlijk beter.
4 – De enige van de schalen (ENK) die duidelijk niet het nivo van .25 van Green's I haalde, is mede om die reden, verder niet gebruikt.
5 – De twijfelachtige kwaliteit van een schaal met een I beneden .50 is niet veel anders dan die met een H van Loevinger beneden .50. De leermeester van mijn kritici, Prof. R. J. Mokken wijst op de gelijkheid van beide koëfficiënten (p. 58 voetnoot), hoewel de berekening van beide voor de gehele schaal iets verschilt (p. 59). Tegelijk roept deze echter (evenals Green) op om de grens van .50 betrekkelijk op te vatten, en kiest arbitrair voor .30 als laagste grens, zoals Green arbitrair voor .25 kiest. Van .30, resp. .25 tot .50 gaat het dan om een "weak" of "medium schale", resp. "quasi-scale". Beide auteurs staan min of meer sceptisch tegenover een evt. minimumgrens van .50 (Mokken, p. 185, Green, p. 87).
6 – Naast een toetsende is door mij in het hoofdonderzoek ook een ekploratieve werkwijze gevolgd, ter kontrole van alternatieve referentiekaders van de respondenten in vergelijking met de aan de literatuur ontleende kaders. Dergelijke alternatieve strukturen traden daarbij minder sterk naar voren dan onze theoretische struktuur (die dan ook mede op honderden onderzoekingen van anderen gebaseerd is). Overigens heeft een eksploratieve werkwijze bij schaalkonstruktie als gevaar dat door het selekteren van een kleine subset van items uit een groter aantal items, de schaalbaarheid bij toeval kan ontstaan (niet te verwarren met de door de I en H bezworen effekten van de randverdelingen der inter-item-tabellen).
7 – Er is door mij wel gewerkt met Green-schaalbaarheid in het kader van het scalagram- of Guttman-model, maar niet met Guttman-schaalbaarheid. De kri-

teria van Guttman inzake schaalbaarheid zijn m.u.v. de reproduceerbaarheidskoëfficiënt (en dan nog in een ondergeschikte betekenis) juist overbodig gemaakt door de I van Green, e.d. koëfficiënten.
8 – Voor het ondervangen van de drie zgn. bezwaren van mijn werkwijze is noch de methode-Mokken, noch diens schaalmodel, noch de H van Loevinger nodig. Voorzover men eksploratief te werk wil gaan kan dit bv. ook met bepaalde kluster-analyses (zoals in mijn proefschrift gebeurd is, of met faktor-analyse met scheve assen (zoals in geval van Bergmans c.s.). In dit laatste geval werkt men uiteraard met een ander (op een korrelatie-koëfficiënt gebaseerd) dimensiebegrip. Het gaat dan ook niet aan om de resultaten daarvan te beoordelen met een kriterium (een niet nader omschreven „aanvaardbaar minimum") in termen van H-schaalbaarheid.
He zgn. tweede bezwaar is reeds als onjuist terzijde gelegd. Het derde bezwaar, het niet nagegaan hebben van de robuustheid van mijn schalen, heeft een grond van waarheid. Wij hadden inderdaad de schaalbaarheid van elke schaal na kunnen gaan bij verschillende subsets van respondenten. Dit is niet gebeurd, omdat al een heel programma van meer algemeen gangbare kontroles op de schalen was afgewerkt, met inbegrip van betrouwbaarheid- en geldigheid-metingen.
9 – Wat betreft de ENE-schaal zijn mogelijk niet precies de juiste items genomen door mijn kritici. Dit gebeurt nl. ook niet t.a.v. de EA-schaal, waarin VAR 42 niet door ons is opgenomen (p. 383, item 98). T.a.v. de ENE-schaal heeft men zich bovendien kunnen vergissen, omdat in ons overzicht van items een ENE-item (nr. 80), slechts via een voetnoot als zodanig herkenbaar was.
10 – De konklusies t.a.v. de robuustheid van onze schalen komen nogal boud over. Wij zouden willen wijzen op het verschijnsel spreiding. Ter vergelijking, wanneer in ons land de meerderheid uit mannen bestaat, is het toch niet schokkend wanneer blijkt dat dit niet in alle provincies, leeftijdsgroepen, enz. het geval is. Zo vindt men ook vrij gemakkelijk subgroepen waar de schaalbaarheid van een set items te laag uitvalt. Het wordt pas interessant wanneer, net zoals bij elaboratie van tabellen, een duidelijke specifikatie optreedt of spuriousness kan worden vastgesteld, m.a.w. in dit geval de schaalbaarheid geheel zou ontbreken bij sommige, resp. alle subgroepen.

De grote vergissing die Boon en Niemöller maken is hun simplisme. Wie met het Guttman-model werkt moet inderdaad verantwoording afleggen van de schaalbaarheid in termen van I- of H-koëfficiënt. Daarnaast zijn er echter vele andere belangrijke zaken.
Nemen we als voorbeeld de enige schaal die we in feite toch bij ons hoofdonderzoek gebruikt hebben, hoewel deze net niet de minimumgrens van Green behaalde (.24 i.p.v. .25). De desbetreffende dimensie (MMS ofwel ontevredenheid over de maatschappelijke positie) werd ontleend aan de literatuur; de desbetreffende schaal behaalde in twee vooronderzoekingen een I van .47, resp. .50; ook de betrouwbaarheid bleek in het kader van een vooronderzoek een redelijk nivo te behalen (split-half r van .67). In het hoofdonderzoek konden we, door omstandigheden, slechts één item van de schaal opnemen, en moesten het ter aanvulling doen met drie andere nog niet eerder beproefde items. Schaalbaarheid en betrouwbaarheid zakten mede hierdoor. De validiteit van de schaal kon echter zeer duidelijk worden aangetoond. Ook via kluster-analyse en non-metrisch faktor-analyse kwam de eigenheid van de MMS-dimensie duidelijk boven drijven (op één item na).

De konklusie moet dan zijn dat men ondanks aanwijsbare ongunstige omstandigheden bij de dataset van het hoofdonderzoek, de schaal toch betrekkelijk veilig kan gebruiken bij de analyse van die dataset. De samenhangen die in het onderzoeksmodel worden verwacht, zullen echter waarschijnlijk wat lager uitvallen. Tegelijk dient men de lezer er op te wijzen dat hij bij verdere toepassingen ook de nodige aandacht geeft aan de items die in het vooronderzoek in de schaal worden opgenomen. Dit alles kan men lezen in mijn proefschrift.
Schaalontwikkeling is een langdurige en veelzijdige zaak, en bepaald ingewikkelder dan het automatisme van een bepaalde rekenprocedure.

Literatuur

Green, B. F., A method of scalogram analysis using summary statistics', in: *Psychometrica*, Vol. 21, 1956, pp. 79-88
Heunks, F. J., *Aliëntatie en stemgedrag*, Tilburg, 1973.
Mokken, R. J., *A theory and procedure of scale analysis*, Den Haag, 1970

Tien?
Reaktie op F. J. Heunks 'Tien kleine vergissingen en één grote'
door M. E. Boon en B. Niemöller

Laten we Heunks' verslag van de ontwikkeling van zijn aliënatieschalen eens op de voet volgen om na te gaan of de 'leerlingen van Mokken' wel zulke simpele zielen zijn als Heunks suggereert. Wellicht wordt dan tevens duidelijker dan in ons eerste beknopte stuk waarom we vinden dat de tien schalen van Heunks nog lang niet allen voor algemeen gebruik geschikt zijn.

1
Heunks noemt in het begin van zijn verantwoording twee mogelijke werkwijzen om tot schaalkonstruktie te komen, namelijk de eksploratieve – en de toetsende – werkwijze. Zijn besluit om voor de toetsende werkwijze te opteren rechtvaardigt hij als volgt: 'Wij meenden theoretisch sterk genoeg te staan om deze laatste weg aanstonds te bewandelen. Begrippen als machteloosheid zijn theoretisch duidelijk plaatsbaar, men zie daarvoor de laatste paragrafen van hoofdstuk 1 en 2. Voor de volledigheid is in een latere fase de alternatieve werkwijze eveneens nog gevolgd.' (Heunks, p. 263)
Of de genoemde *begrippen* inderdaad zo duidelijk plaatsbaar zijn lijkt ons op dat punt van zijn betoog minder relevant. Het gaat erom of de door Heunks gekozen items goede *indikatoren* zijn voor die begrippen. Dat later 'voor de volledigheid' de alternatieve werkwijze nog wordt gevolgd lijkt ons een minder juiste volgorde.

2
Het toetsen van zijn a priori schalen doet Heunks m.b.v. Green's Index of Consistency' waarover hij o.a. opmerkt: 'Deze "Index of Consistency" is behalve op

de gehele groep van schaalitems ook toepasbaar op itemparen daaruit, hetgeen in staat stelt tot het selecteren van die items welke tesamen de beste schaalbaarheid vertonen.' (Heunks, p. 264)
Dit lijkt ons om een aantal redenen een merkwaardige uitspraak:
- Uit de indices voor item-*paren* valt niets met zekerheid te zeggen omtrent de index voor een eventuele kombinatie van meer dan twee items. Dit vormt een van de problemen die men bij eksploratieve schaalanalyse tegenkomt.
- Heunks wijst later (Heunks, p. 266 voetnoot 2 en Heunks' reaktie punt 6) dit soort zoek procedures af.
- Bestudering van de matrices met Green's I voor alle itemparen (Heunks, pp. 388/9) zou het zelfvertrouwen van Heunks ten aanzien van zijn theoretische a priori-schalen moeten ondermijnen.
  (zie verder punt 6)

De vraag rijst nu hoe serieus Heunks de mogelijkheden van de door hem gekozen index heeft benut.

3
Het onder punt 2 te berde gebrachte heeft overigens niets te maken met onze waardering voor Green's index I. Integendeel, voor Green's I geldt, evenals voor Jackson's PPR en Loevinger's H, dat zij '. . . are all objective, and result in measures which are not affected by the distribution of item difficulties.' (White and Saltz, p. 189) Aangezien wij in het navolgende gebruik zullen maken van Loevinger's H lijkt het volgende citaat nog op zijn plaats.
'All have the same underlying rationale, but differ slightly in the way in which "errors" are counted. Loevinger's Ht is the most conservative of the three since all possible errors are counted; Jackson's PPRt is the least conservative, and Green's I will usualy fall between the two.' (White and Saltz, p. 189)[1]

4
Alvorens tot de presentatie van de met behulp van Green's I gemeten kwaliteit van zijn a priori schalen over te gaan, geeft Heunks nog een toelichting m.b.t. de aan I te stellen 'ondergrens'.
'Een I van omstreeks .50 wordt door Green bevredigend geacht, terwijl een lagere I tot .25 toe twijfelachtig wordt geacht.
Nog lagere I-waarden duiden volgens hem op het ontbreken van schaalbaarheid in een Guttman-zin.' (Heunks, p. 264)
Deze aanduidingen moeten worden opgevat als empiriese bevindingen; een theoretische basis ontbreekt. Hetzelfde geldt voor de vuistregel die Mokken geeft voor Loevinger's H:
a – $0.50 \leq H$: a strong scale;
b – $0.40 \leq H : < 0.50$: a medium scale;
c – $0.30 \leq H : < 0.40$: a weak scale.
(Mokken, 1971, p. 185)
Wel kunnen we hieruit konkluderen dat men met item-kombinaties die een index hebben in de buurt van de .30 of zelfs .25 wel bijzonder voorzichtig dient te zijn als het om toekenning van het predikaat 'schaal' gaat.
Gelet op het onder punt 3 gestelde moeten we voorts konkluderen dat een ondergrens voor Green's I als regel *hoger* gesteld moet worden dan die voor Loevinger's H.