



Universiteit
Leiden
The Netherlands

Mooie modus!

Haak, C.P.M. van der

Citation

Haak, C. P. M. van der. (1971). Mooie modus!. *Acta Politica*, 6: 1971(3), 304-309. Retrieved from <https://hdl.handle.net/1887/3451369>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3451369>

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

Mooie modus!

Reactie op het artikel van C. Boef 'Politiek op de televisie' in Acta Politica van januari 1971. p. 92—107

door C. P. M. van der Haak

Boef's analyse van de kijkdichtheid- en waarderingscijfers voor een aantal politieke televisie-uitzendingen is uitgelopen op een methodologische verhandeling over het meten van de waardering voor tv-programma's in het algemeen.

Bij de omroep hanteert men als maatstaf het *waarderingcijfer*, zijnde het rekenkundig gemiddelde (r.g.) van de beoordelingscijfers op een schaal van 1 tot 10 ('schoolrapportcijfers'), die door een steekproef uit de kijkers aan een programma worden toegekend.

Boef stelt voor, de 10-puntsschaal bij de verwerking te transformeren tot een 5-puntsschaal en het r.g. als maat van centrale tendentie te vervangen door de modus.

Vanuit het gezichtspunt van de meest betrokkenen, nl. de onderzoekers werkzaam bij de omroep, zijn hierbij de volgende kanttekeningen te maken.

1 Boef doet zijn voorstel zonder zichzelf (of iemand anders) de vraag te stellen, waartoe het waarderingscijfer dient en of de modus daartoe even bruikbaar is als het r.g. Het waarderingscijfer dient, evenals het hele kijk- en luisteronderzoek, om degenen die het programmabeleid bepalen, een indruk te geven van de reacties op de programma's.

De modus nu is zó grof, dat alle nuances tussen afzonderlijke programma's van één soort wegvallen. Daaruit konkludeert Boef dan weer dat de waardering kennelijk niet zozeer door de kwaliteit van het afzonderlijke programma bepaald wordt als wel door de programmasoort. Vervolgens konkludeert hij dat je, als dit zo is, wel kunt ophouden met de waardering continu (d.w.z. van alle programma's) te meten, omdat je die waardering inmiddels wel kent en de modi alleen zullen veranderen als de algemene waardering voor een programmakategorie verandert. Met deze gedachtengang schiet het programmabeleid niet veel op.

2 Ook theoretisch valt er op de probleemstelling, zoals Boef die terloops formuleert, wel wat af te dingen. Hij schrijft: 'de informatie die wij wensen: welke beoordeling (exact) haalt het grootste aantal stemmen en hoeveel zijn dat er?', wordt ons wel geleverd door de modus en niet door het gemiddelde' (pag. 97). Een gaaf voorbeeld van aanpassing van de probleemstelling aan de kennelijk bij voorbaat gekozen methode. Het gaat hier echter niet om de verkiezing van bijv. de 7 tot Miss Waardering, maar om de registratie van de waardering van het gehele publiek met behulp van een schaaltechniek. En bij de

verwerking van schaal-gegevens is het r.g. minstens zo gebruikelijk en bruikbaar als de modus.

3 De modus, berekend over de (van de 10-puntsschaal afgeleide) 5-puntsschaal, leidt tot verkeerde conclusies. Alle uitzendingen van politieke partijen krijgen in de nieuwe 5-puntsschaal als modus een 3. Echter, het percentage dat de modus vormt van het totaal aantal gegeven beoordelingen, is bij de grote partijen significant lager dan bij de kleine partijen.

Hieruit konkludeert Boef: 'Bij de grote partijen is dus het aantal niet-modale beoordelaars relatief groot, dat wil dus zeggen dat het aantal relatief felle voor- of tegenstanders van het programma betrekkelijk talrijk is' (p. 103).

Dit statistisch significante gegeven is een artefact van de gevolgde werkwijzen. Aan de uitzendingen van de kleine partijen worden door de respondenten hoofdzakelijk zessen en vijven toegekend, die bij de transformatie tot 5-puntsschaal in hetzelfde hokje terechtkomen, nl. 3. De uitzendingen van de grote partijen krijgen iets hogere beoordelingscijfers, nl. vooral zessen en zevens, die in de 5-puntsschaal in twee verschillende categorieën komen (6 wordt 3; 7 wordt 4) en daarmee de grotere spreiding veroorzaken. Boef zelf constateert trouwens, voordat hij de 5-puntsschaal ten tonele voert, dat alle 28 uitzendingen van de politieke partijen als modus een 6 hebben, en rept daarbij niet van een verschil in percentage tussen grote en kleine partijen. Bij hanteren van het r.g. blijkt dat de grote partijen, naast een iets hogere waardering, een gelijke of iets kleinere standaard-deviatie hebben dan de kleine partijen.

4 Boef's hypothese dat (ik kies de scherpste formulering) 'niet de kwaliteit van de individuele uitzending wordt beoordeeld, maar het soort programma' (p. 105) kan met behulp van het r.g. als waarderingsmaatstaf aanzienlijk genuanceerd worden, zelfs met betrekking tot de twee uitzonderlijk homogene programmakategorieën waarop Boef zijn hypothese baseert. Brandpunt krijgt normaliter een waardering van 72 (7,2), maar de Brandpuntreportage 'Bommen op Biafra' werd gewaardeerd met 75 (7,5) en de Brandpuntreportage over de toneelgroep Studio met 67 (6,7).

Leden/abonnees van de KRO gaven aan uitzendingen van de KVP systematisch een iets hoger cijfer dan aan uitzendingen van andere partijen. Hetzelfde geldt voor NCRV-leden t.a.v. CHU en ARP en voor VARA-leden t.a.v. P.v.d.A. Hieruit blijkt dat de kijkers wel degelijk verschil zien tussen de afzonderlijke programma's van één soort. Of het de 'kwaliteit' van het programma is, die de waardering bepaalt, is natuurlijk niet te zeggen. Bij de politieke partijen zal het eerder de meer of minder konsonante informatie zijn.

Over het verband tussen de waardering voor een programmakategorie en die voor een afzonderlijk programma hopen we binnenkort meer te weten te komen,

wanneer we (m.i.v. het 2e kwartaal 1971) de waarderingscijfers per programma geanalyseerd krijgen naar het al dan niet liefhebber zijn van de betreffende programmakategorie.

5 Ook al levert het r.g. bruikbaarere informatie dan de modus, het moet natuurlijk wel vaststaan, dat deze informatie valide is, en niet het produkt is van een ongeoorloofde verfijning van de basisgegevens. De kritiek op dit punt is ook het uitgangspunt geweest voor Boef's poging tot een alternatief (p. 96/97).

Hierover valt het volgende te zeggen.

De statistische bezwaren die Boef noemt (scheve verdeling en ongelijke intervallen) delen wij uiteraard, evenals zijn opvatting dat deze niet doorslaggevend zijn.

Wanneer men van die statistische bezwaren afziet, is er geen enkele reden waarom men over de door de individuele respondenten gegeven beoordelingscijfers geen gemiddelde met één decimaal zou mogen berekenen.

'De zeer globale aard van de gegevens vraagt een grove benadering van de centrale tendens' schrijft Boef. De scores op een 10-puntsschaal zijn niet globaler dan de scores op welk soort schaal dan ook. Waarschijnlijk (vgl. p. 93) doelt Boef hier op het feit dat het beoordelingscijfer dat een respondent aan een programma toekent, een uitdrukking is van zijn *globale (on)tevredenheid* over het programma, zonder specificatie van de factoren welke tot die (on)tevredenheid bijdragen, zoals het onderwerp van de uitzending, de presentatie, de informatieve waarde dan wel de amusementswaarde, de uitslag van een via de tv uitgezonden voetbalwedstrijd (reportages van door de nederlandse clubs gewonnen wedstrijden krijgen een hogere waardering dan verloren wedstrijden). Maar dat verhindert toch niet een verfijnde meting van die tevredenheid?

Daarentegen leidt hantering van de modus, zelfs bij een 10-puntsschaal, tot een ongerechtvaardigde vergroving van de basisgegevens. Wanneer we de waarderingscijfers voor de programma's van een willekeurige week (21-27 febr. 1971) 'modifieren', blijkt dat ca. 70% van alle programma's een 7 als modus krijgt, ruim 20% een 8 en bijna 10% een 6.

Van de programma's met modus 6 varieert het r.g. van 5,3 tot 7,0, van de programma's met modus 7 loopt het r.g. uiteen van 6,3 tot 7,5; van de programma's met modus 8 loopt de range van 7,0 tot 8,1.

In alle gevallen is de variatiebreedte ontoelaatbaar groot, gezien de beperkte spreiding die de waarderingscijfers doorgaans vertonen (waarover direkt meer). Het tweede argument van Boef (p. 97) is hierboven weerlegd onder 2.

Zijn derde argument richt zich tegen de gewoonte, het waarderingscijfer zonder komma te presenteren, dus 71 i.p.v. 7,1.

Dit wordt gedaan om verwarring met de kijkdichtheidcijfers, die vaak in de-

zelfde orde van grootte liggen, te voorkomen.

Ook afgezien daarvan is deze praktijk goed te verdedigen. Zeker 90% van de programma's krijgt een waarderingscijfer dat ligt tussen 6,0 en 8,0, zodat de verschillen hoofdzakelijk in de decimaal tot uiting komen. In dat geval ligt het voor de hand de aandacht dan ook te vestigen op die decimaal door dit stukje van de schaal a.h.w. onder een vergrootglas te leggen.

Iedereen in de omroep, waarvoor deze gegevens in de eerste plaats bedoeld zijn, weet, dat het oorspronkelijk geen 100-puntsschaal was.

Het laatste argument van Boef luidt: 'Voor onze doelstelling geeft het gemiddelde zonder nadere gegevens minder informatie dan de modus, en bovendien is de laatste eenvoudiger en sneller te bepalen'. Op de doelstelling is hierboven onder 1 en 2 ingegaan, op het feit dat i.t.t. wat Boef beweert, het r.g. meer informatie geeft dan de modus, onder 4. en 5. Daar kan nog aan toegevoegd worden, dat in de omroep het waarderingscijfer nooit 'zonder nadere gegevens' gehanteerd wordt, maar altijd in combinatie met het kijkdichtheidcijfer en dikwijls met standaard-deviatie en frekwentieverdeling, die van elk waarderingscijfer bekend zijn.

Met excuses aan de lezers van *Acta Politica* voor deze technische discussie. Maar het kan geen kwaad, het instrumentarium eens te onderzoeken, waarmee zowel het trio Manschot, Niemantsverdriet en Noldus¹), als Boef en Hoogendijk² tot hun konklusies inzake de belangstelling voor politiek op de televisie kwamen.

Een mooie modus is niet lelijk!

Antwoord van C. Boef

Op verzoek van de redactie zal ik in het hierna volgende reageren op de kanttekeningen die de heer Van der Haak van de Afdeling Studie en Onderzoek van de NOS gemaakt heeft bij mijn artikel 'Politiek op de Televisie' in het januari-nummer van dit blad. Hij meent een aantal argumenten te hebben gevonden, die tegen het gebruik van de door mij voorgestelde parameter pleiten.

Voordat ik op de puntsgewijze geformuleerde kritiek inga nog dit: Van der Haak meent, dat ik voorstel, de tienpuntsschaal te transformeren in een vijfpunts. Dat is niet juist. Ik heb mij afgevraagd waarom er geen vijfpuntsschaal gebruikt is. De transformatie heb ik alleen gebruikt om aan te geven wat er zou kunnen gebeuren wanneer een schaal met een vijfpuntsnotering zou zijn gebruikt. Uit het

¹ B. J. A. M. Manschot, L. J. Niemantsverdriet en T. W. J. J. M. Noldus, 'TV-uitzendingen van politieke partijen', in: *Massamedia en politiek*, Utrecht, Ambo, 1968.

² F. A. Hoogendijk, *Partijpropaganda in Nederland*, Amsterdam 1971.

voorbehoud dat ik maak ten aanzien van de transformatie, blijkt duidelijk dat ik vervanging wil (vgl. p. 102).

Ad. 1 — Het verwijt dat ik mijzelf of een ander niet gevraagd heb waartoe het waarderingscijfer dient, wordt gelogenstraft door mijn inleiding, waarin ik uitgebreid inga op het doel van het continu programma-onderzoek, en waarbij ik mij heb laten leiden door hetgeen Hendriksen, het huidige hoofd van de afdeling onderzoek van de NOS, daarover in de Sociologische Gids publiceerde.

Dat de modus als parameter vrij grof is, is juist. Maar als het materiaal minder goed geschikt is voor een fijner instrument, dan is dát er de oorzaak van dat er slechts globale gegevens uitkomen.

Vervolgens poneert Van der Haak dat het programmabeleid weinig opschiet met mijn gedachtengang. Dat zal best zo zijn, maar dan doet men er beter aan, de materiaalverzameling, en dus het onderzoek, aan de wensen van het beleid aan te passen in plaats van de uitkomsten.

Ad. 2 — Dat op grond van het materiaal bepaalde informatie verwacht kan worden en dat die met behulp van de modus zichtbaar gemaakt kan worden was in dit geval voor mij een eenvoudige redenering. Dat Van der Haak dit een gaaf voorbeeld van aanpassing van de probleemstelling aan de kennelijk van tevoren gekozen methode noemt is zijn zaak.

Bij verwerking van schaalgegevens is het rekenkundig gemiddelde minstens zo bruikbaar en bruikbaar als de modus, stelt Van der Haak. Over de bruikbaarheid kan men van mening verschillen; bruikbaarheid is een argument dat beter niet op wetenschappelijke problemen kan worden toegepast.

Ad. 3 — Bij mijn conclusie, dat het percentage dat de modus vormt van het totaal aantal gegeven beoordelingen, bij grote partijen significant lager is dan bij kleine partijen, heb ik er inderdaad geen rekening mee gehouden, dat dit wel eens het resultaat zou kunnen zijn van de transformatie. Het zou dus beter geweest zijn, dat de nog al detaillistische conclusie achterwege was gebleven. Gelukkig tast dit mijn algemene argumentatie voor het gebruik van de modus niet aan, omdat die ook gebruikt kan worden bij een tienpuntsschaal. De transformatie diende alleen maar als voorbeeld, zoals ik al eerder betoogd heb. Op p. 102 heb ik er uitgebreid op gewezen dat de uitkomsten van een getransformeerde schaal wel eens heel anders zouden kunnen zijn dan de uitkomsten van een echte vijf-puntsschaal in hetzelfde geval toegepast.

Ad. 4 — De bedoeling van deze opmerking is niet helemaal duidelijk. Mijn uitspraak: 'Niet de kwaliteit van de individuele uitzending wordt beoordeeld maar het soort programma', is wellicht wat te ongenueanceerd. Onder hetzelfde hoofd

enige regels eerder stel ik echter, dat de waardering *in de eerste plaats* een waardering voor een bepaalde categorie programma's is. Daaruit blijkt toch duidelijk dat ik niet denk dat er geen uitzonderingen zijn. Verder heb ik gezegd dat de relatief sterk affectieve reacties buiten de routinematige modale beoordelingen vallen (p. 106). Dat is kennelijk van toepassing op de KRO-leden die de KVP beoordelen enz. De Brandpunt-voorbeelden tasten evenmin mijn algemene hypothese aan.

Het begrip kwaliteit is door Van der Haak in waarderende zin opgevat, terwijl ik het heb gebruikt zoals in de tegenstelling kwalitatief-kwantitatief.

Ad. 5 — De globale aard van de gegevens vraagt om een grove benadering van de centrale tendens. Dat ik met globaal bedoeld heb dat het gegeven niet unidimensioneel is blijkt duidelijk uit hetgeen ik eraan vooraf liet gaan (vgl. bijv. p. 93).

De ingebouwde twijfel van Van der Haak („waarschijnlijk bedoelt Boef . . .”) is dus enigszins overbodig.

Maar, zegt hij, dat verhindert toch niet de verfijnde meting van de tevredenheid? Natuurlijk niet, maar meting verricht je in eerste instantie met behulp van je operationaliseringsvragenlijst. Als je je items toespitst op een fijnere meting, kan je ook je schaaltechniek daarbij aanpassen. Verfijnde meting zou in dit geval kunnen plaatsvinden door de multidimensionale globale waardering op te splitsen in zoveel mogelijk unidimensionale gegevens. De centrale tendens weergegeven door een parameter die minder geschikt is dan een andere is in ieder geval geen goed voorbeeld van verfijnde meting.

Vervolgens levert Van der Haak een gave ondersteuning van mijn hypothese, dat de mensen in het algemeen naar een programma kijken als ze van tevoren reeds een positieve waardering hebben voor de categorie waarin dat programma thuishoort. 7 is een typisch waarderingscijfer, dat deze algemene waardering weergeeft. Het is niet uitgesproken goed, maar ook niet op het nippertje. Van iemand die kijkt is de kans dus groot dat hij een programma op een 7 waardeert. Dat gold in de week van 21-27 februari 1971 voor 70% van alle programma's! Tenslotte nog een opmerking. Dat men bij de waarderingscijfers een tienpuntsschaal tot een honderdpunts transformeert om verwarring met andere cijfers te voorkomen, is het meest curieuze argument dat ik voor een dergelijke handeling zou kunnen verzinnen.