

# Hoe betrouwbaar is de redactie van De Psycholoog?

H.G. Schmidt<sup>1</sup>

*Voor de bewoners van de academische wereld wordt publiceren hoe langer hoe belangrijker. Dat legt een extra verantwoordelijkheid op de schouders van redactie van wetenschappelijke tijdschriften. Redacties doen er dan ook goed aan af en toe eens naar hun eigen functioneren te kijken. Henk Schmidt, redactielid, becijferde de overeenstemming tussen de oordelen van de redactieleden van De Psycholoog.*

De redactie van een wetenschappelijk tijdschrift is als de portier van een exclusieve discotheek. Hij selecteert uit het aanbod diegenen die worden toegelaten en wijst anderen af. De vraag welke criteria bij dat selectieproces gehanteerd worden is van niet onaanzienlijk maatschappelijk en – in het geval van de besluitvorming over potentiële bijdragen aan een tijdschrift – wetenschappelijk belang. Immers, omdat communicatie tussen wetenschapsbeoefenaren veelal via de kolommen van hun tijdschriften verloopt, speelt een redactie een sleutelrol bij het bepalen van wat aanvaardbare producten van wetenschap zijn en wat niet. Daarnaast heeft de mogelijkheid te publiceren maatschappelijke consequenties. Bij bevorderingsbeslissingen speelt het aantal en de aard van de publicaties waarop iemand zich kan beroepen een belangrijke rol.

Het zal dan ook niemand verbazen dat de beoordeling van ingezonden artikelen

door redacties van psychologische tijdschriften, of door daarbij ingeschakelde externe beoordelaars, met name in de Verenigde Staten frequent object van onderzoek is geweest. Zie voor een overzicht Marsh & Ball (1981).

In Nederland is die praktijk minder gemeengoed. Ik herinner me enkel een onderzoek van Buiters en Hoogstraten (1979), dat zich richtte op de wijzen waarop redacties van acht vooraanstaande tijdschriften hier te lande hun beoordelingsproces hadden georganiseerd en welke normen zij bij die beoordeling hanteerden. Over de *kwaliteit* van dat beoordelingsproces, uitgedrukt bijvoorbeeld in een of andere index voor de interbeoordelaarsovereenstemming, is echter nog geen informatie beschikbaar. Dit artikel is bedoeld als een eerste poging om – in ieder geval voor dit tijdschrift – in die leemte te voorzien.

Eerst echter enkele zinnen over de manier waarop de redactie van *De Psycholoog* een ingezonden bijdrage gewoonlijk behandelt. De redactie maakt in tegenstelling tot de meeste andere bla-

den in het veld nauwelijks gebruik van externe beoordelaars. De leden ontvangen de potentiële artikelen thuis, vormen zich een oordeel over de kwaliteiten van het stuk en maken zonodig aantekeningen in de tekst die later als mogelijke suggesties voor verbeteringen gebruikt kunnen worden.

Maandelijks komt de redactie bijeen voor een vergadering waarin de bijdragen en het oordeel van de leden daarover gedetailleerd besproken worden. Meestal wordt geprobeerd tot een unanieme beslissing te komen. Dat betekent dat de bespreking van een enkel artikel soms aanzienlijke hoeveelheden tijd vergt. Als de redactie zich onvoldoende competent acht, wordt een oordeel van derden ingewonnen.

Welke criteria legt de redactie aan bij het beantwoorden van de vraag of een bijdrage voldoende kwaliteit heeft om geplaatst te worden? Voor het onderzoek dat Buiters en Hoogstraten (1979) deden naar het functioneren van redacties heeft de toenmalige redactiesecretaris een inhoudsanalyse uitgevoerd op de notulen van de vergaderingen die

1. Met dank aan B. Kerkhofs.

gedurende één jaar gehouden werden en op de brieven die hij in die periode aan auteurs over hun stukken schreef. Als belangrijkste criteria kwamen daarbij naar voren: relevantie, theoretische en methodologische adequaatheid, en vooral: helderheid van opzet en stijl: heeft de tekst een duidelijke structuur? wordt de probleemstelling op een goede manier geïntroduceerd? is het leesbaar geschreven? enzovoorts.

Naast deze overwegingen is er nog een andere, meer prozaische vorm die bepaalt of een stuk geplaatst wordt. Dat is de omvang van het aanbod. Bij een groter aanbod wordt de drempel die de toegankelijkheid van de kolommen reguleert automatisch hoger. Dat gebeurt om te voorkomen dat er een reservoir van artikelen ontstaat, die de publicatiesnelheid van nieuwe bijdragen zou vertragen. De tijd die verstrijkt tussen het aanbieden van een stuk en de publicatie ervan is nu acht weken, aanzienlijk korter dan die van andere tijdschriften, en essentieel voor een blad dat actueel wil zijn. Dat betekent echter dat wanneer in een bepaalde maand het aanbod groot is, de normen automatisch wat opgeschroefd worden. Het afwijzingspercentage van het blad is overigens door de jaren heen min of meer constant gebleven: ongeveer 40%.

Het doel van het hier gerapporteerde onderzoek was niet zozeer na te gaan welke normen gehanteerd worden bij een oordeel over een bepaalde bijdrage, maar wat de *kwaliteit* van dat oordeel is. In welke mate stemmen redactieleden overeen in hun visie op een tekst? Zijn er notoir strenge beoordelaars en ongeneeslijk goedgehartigen onder hen? Met andere woorden: hoe betrouwbaar oordelen zij? Uit ander onderzoek naar de kwaliteit van redacties is gebleken dat de interbeoordelaarovereenstemming ook voor globale oordelen als de dichotomie „accepteren–afwijzen” meestal laag is. Marsh en Ball (1981) rapporteren in hun onderzoek waarden die variëren tussen 0,08 en 0,54 voor enkelvoudige oordelen.

## Methode

Vijf redactieleden beoordeelden gedurende het jaar 1984 alle bijdragen voor *De Psycholoog*. Zij gaven voorafgaande aan de discussie over een stuk mondeling hun oordeel, uitgedrukt in één van drie categorieën: „afwijzen”, „slechts accepteren na aanzienlijke wijzigingen” en „accepteren (eventueel met

kleine wijzigingen)”. Dit oordeel werd genoteerd. Ook de uiteindelijke beslissing die na discussie genomen werd, werd genoteerd. In totaal werden 48 ingezonden artikelen beoordeeld. Niet ieder redactielid beoordeelde alle artikelen. Over vier artikelen waren geen oordelen beschikbaar. Product-moment correlaties tussen de oordelen werden berekend.

## Resultaten en discussie

In *tabel 1* zijn per beoordelaar en per beoordelingscategorie percentages weergegeven. Daarnaast is ook het eindoordeel van de redactie als geheel opgenomen.

Uit deze tabel valt af te leiden dat de beoordelaars niet sterk uiteenlopen in de mate van strengheid waarmee geoordeeld wordt. Als we kijken naar de afwijzingspercentages dan blijkt dat Schroots en Wegman wat minder dan de andere redactieleden geneigd zijn om een stuk zonder meer te weigeren. Grotere verschillen zitten er in de geneigdheid van redactieleden een bijdrage zonder meer (of met kleine kanttekeningen) te accepteren. Wegman en Schmidt zijn daartoe meer bereid dan de overige redactieleden, die vaker vinden dat een stuk flink zou moeten worden gewijzigd.

Het eindoordeel blijkt een aardig gemiddelde van deze afzonderlijke oordelen te vormen.

In *tabel 2* zijn de intercorrelaties tussen beoordelaars onderling, en met het eindoordeel afgedrukt.

De gewogen gemiddelde correlatie tussen alle paren beoordelaars bedraagt 0,48. De gemiddelde correlatie tussen afzonderlijke beoordelaars en het uiteindelijke redactieoordeel is uiteraard hoger, namelijk 0,70. Deze resultaten duiden op redelijke consistente en tamelijk homogene beoordelingen, die, zeker in vergelijking met wat in ander onderzoek gebruikelijk is, enige reden tot tevredenheid met het beoordelingsproces geven.

Men kan zich uiteraard afvragen hoe

Tabel 1. *Beoordelingen van artikelen, ingezonden in 1984, per beoordelaar en voor de redactie als geheel, uitgedrukt als percentage*

| Beoordelaar          | Aantal beoordelingen | „Afwijzen” | „Wijzigen” | „Accepteren” |
|----------------------|----------------------|------------|------------|--------------|
| Wegman               | 47                   | 21,2       | 29,8       | 48,9         |
| Bonarius             | 44                   | 25         | 45,5       | 29,5         |
| Roe                  | 35                   | 28,6       | 48,6       | 22,9         |
| Schroots             | 39                   | 15,4       | 51,3       | 33,3         |
| Schmidt              | 42                   | 28,6       | 28,6       | 42,8         |
| Eindoordeel redactie | 48                   | 27,1       | 33,3       | 39,6         |

Tabel 2. *Correlaties tussen oordelen*

|          | Bonarius | Roe  | Schroots | Schmidt | Eindoordeel redactie |
|----------|----------|------|----------|---------|----------------------|
| Wegman   | 0,43     | 0,54 | 0,69     | 0,45    | 0,78                 |
| Bonarius |          | 0,34 | 0,46     | 0,28    | 0,58                 |
| Roe      |          |      | 0,42     | 0,55    | 0,69                 |
| Schroots |          |      |          | 0,52    | 0,79                 |
| Schmidt  |          |      |          |         | 0,61                 |

het komt dat deze resultaten enigszins afwijken van wat eerder in ander onderzoek gevonden werd. Een mogelijke verklaring is de volgende.

In het geciteerde onderzoek gaat het steeds om beoordelaars die schriftelijk hun oordeel geven en niet noodzakelijk de andere beoordelaars en hun wijze van beoordelen kennen. Het is niet uitgesloten dat de in dit onderzoek gevonden hogere interbeoordelaars overeenstemming het resultaat is van het feit dat de redactie van *De Psycholoog* veelvuldig intensief overleg pleegt over de inzendingen en dus langzamerhand elkaars voorkeuren, criteria, etc. heeft leren kennen. Dat laatste zou er dan weer toe kunnen leiden dat men zijn oordeel aanpast aan het te verwachten oordeel van de overigen. Tegen deze veronderstelling spreekt overigens dat de oudere redactieleden het niet meer met elkaar eens zijn, dan met diegenen die meer recent zijn togetreden.

Een andere verklaring zou kunnen zijn dat er bij redactieleden die veel artikelen moeten beoordelen een leereffect optreedt, dat bij beoordelaars die dat werk over het algemeen minder vaak doen, niet optreedt.

Hoe het ook zij: auteurs van niet-geaccepteerde bijdragen mogen, gezien hun reacties soms, het redactiebeleid als grillig ervaren; de hier gepresenteerde gegevens ondersteunen die visie niet.

## Literatuur

- Buiter, R. K. en J. Hoogstraten (1979). De poortwachtersfunctie van tijdschriftredacties. *De Psycholoog*, 14, 457-464.
- Marsh, H. W. en S. Ball (1981). Interjudgmental reliability of reviews for the Journal of Educational Psychology. *Journal of Educational Psychology*, 73, 872-880.

Drs. H. G. Schmidt is verbonden aan de afdeling Onderwijsontwikkeling en Onderwijsresearch van de Rijksuniversiteit Limburg.

---

## Summary

---

*Articles submitted for publication in De Psycholoog were independently judged by five editors. The judgments were made in one of three categories: „rejection”, „acceptation with minor editorial changes,” and „request to revise.” The mean interjudgmental reliability was equal to 0.48, which is higher than data normally produced in this kind of research. Reasons for this discrepancy are discussed.*