

Goal Attainment Scaling

door C. A. M. de Wit*

Goal Attainment Scaling is een methode om het effect van psychotherapie te evalueren. Deze methode is ontwikkeld door Kiresuk en Sherman in 1968. Zij lijkt grote mogelijkheden te bieden en vele voordelen te hebben boven andere methodes. In dit artikel wil ik de methode beschrijven; een uiteenzetting geven over het theoretisch kader, waar de methode in te plaatsen is; kritische kanttekeningen plaatsen bij enkele aspecten van de methode van Kiresuk en Sherman en bij wat anderen daar over gezegd hebben; enkele problemen bespreken, die zich kunnen voordoen bij konstruktie en gebruik van G.A.S.-schalen; de voordelen van de G.A.S.-methode beschrijven.

1 Opzet van de methode

De G.A.S.-methode steunt op het principe dat de individuele problemen, waarmee een cliënt zich aanmeldt, alsook de daarvoor opgestelde therapie-doelcinden, in een schaal worden uitgezet. Goal Attainment Scaling is een individualiserende evaluatiemethode.

Men gaat bij het opstellen van de schalen als volgt te werk. Getracht wordt om zoveel mogelijk informatie te verkrijgen over de problemen en omstandigheden van de cliënt. Door een 'Goal-selector' of een 'goal selection committee' worden op grond van deze informatie een reeks reëel bereikbare doelcinden opgesteld. Voor elk van deze problemen en doelcinden wordt een schaal gekonstrueerd: een vijf-puntsschaal. Op het -2 niveau wordt het minst gunstige behandelingsresultaat uitgezet, dat men kan bereiken, en op het +2 niveau het meest gunstige. Het centrum van de schaal is het 0 niveau; hierop wordt (in de opzet van Kiresuk en Sherman), het behandelingsresultaat uitgezet, dat men het meest waarschijnlijk te bereiken acht. Op het -1 en het +1 niveau worden resultaten omschreven, die tussen het 0 en -2, respectievelijk 0 en +2 niveau liggen. Kiresuk en Sherman stellen hierbij als eis, dat minstens twee van deze niveau's voldoende exact en objectief omschreven worden in termen van gebeurtenissen, waarvan de aan- of afwezigheid voldoende objectief vastgesteld kan wor-

* Psycholoog, werkzaam op De Mark, psychotherapeutisch centrum voor kinderen, jeugdigen en het gezin, te Breda (hoofd G. Bonroy, psychiater).

den. Dit betekent, dat omschrijvingen mogelijk zijn in termen van observeerbaar gedrag, test- en meetuitslagen, zelfbeschrijvingen door de cliënt etc. Kiresuk en Sherman stellen voor, om de G.A.S.-schalen te konstrueren vóórdat de cliënt in therapie gaat, om ze te laten scoren door een onafhankelijke beoordelaar en om de methode te gebruiken ter evaluatie bij een follow-up-onderzoek.

Een voorbeeld van een G.A.S.-schaal vinden we bij Kiresuk en Sherman. Het probleem is: angst voor seksueel contact. Voor dit probleem stellen zij de volgende schaal op:

minst gunstige behandelings- resultaat	minder dan verwacht succes	verwacht resultaat van behandeling	meer dan verwacht succes	meest gunstige behandelings- resultaat
- 2	- 1	0	+ 1	+ 2
afwerend, geen af- spraakjes, geen seksueel contact		afpraakjes, vrijen	enige bevredigende seksuele gemeenschap	regelmatige afpraakjes, regelmatige bevredigende seksuele gemeenschap, huwelijk

2 Het theoretisch kader

In het eerste artikel over de G.A.S.-methoden beschrijven Kiresuk en Sherman enige redenen, die hen tot deze geïndividualiseerde methode gebracht hebben. Zij stellen: 'de tendens bestond om, in plaats van een groot aantal verschillende technieken, een vaste batterij van evaluatieinstrumenten te gebruiken, ongeacht de kenmerken of problemen van de cliënt' (10, pag. 444). Daarna zeggen zij: 'het gebruik van hetzelfde psychometrische of beoordelingsinstrument voor alle patienten leidde onvermijdelijk tot het evalueren van bepaalde patienten op variabelen, die niet relevant waren voor een specifieke dilemma of voor hun speciale omstandigheden. Volgens Kiresuk en Sherman lijkt het 'onredelijk en inefficiënt' om alle cliënten op eenzelfde reeks van variabelen te evalueren, zonder zich daarbij af te vragen, of die variabelen wel van toepassing zijn op deze cliënt met zijn eigen individuele problematiek. Algemene, niet op het individu afgestemde tests of evaluatie-instrumenten kunnen aan de eigen problematiek van een bepaalde cliënt voorbij gaan. Dit lijkt hun onaanvaardbaar: hun alternatief is een geïndividualiseerde evaluatiemethode.

Deze beweegredenen zijn te plaatsen in de bredere context van een ontwikkeling, die zich steeds meer in de huidige psychotherapeutische evaluatie-theorie begint af te tekenen (4, 5, 8, 9). Het evaluatie-onderzoek dient te specificeren naar cliënt-, therapeut- en therapie-variabelen. Evaluatie van het effect van psychotherapie wordt dan: nagaan 'welke therapie, toegepast door wie en onder

Goal Attainment Scaling

welke omstandigheden, het meeste effect heeft bij welke cliënt met welke problemen' (5, pag. 253). Evaluatie-onderzoeken, met een steekproef van cliënten, bij wie het therapie-effect wordt nagegaan met behulp van een voor iedere cliënt toepasbaar geacht meet-instrument, geven een 'gemiddeld resultaat'. Hiertegen zijn zeker drie bezwaren in te brengen.

1 – Het lijkt niet juist, om bij evaluatie-onderzoeken te werken met een steekproef van individuele, verschillende cliënten, met individueel verschillende problemen, die behandeld worden door verschillende therapeuten, die verschillende therapeutische technieken hanteren. Al deze gevallen kunnen immers zo zeer verschillen wat betreft diagnose, behandelingsdoeleinden, criteria, soorten therapeuten en therapieën, dat ze niet als vergelijkbaar in één steekproef samen te brengen zijn.

2 – Het lijkt niet juist om in het geval, dat men toch met een dergelijke steekproef werkt, deze te onderzoeken met behulp van een algemeen meetinstrument. Alle leden van de steekproef verschillen wat betreft persoonlijkheid, achtergronden en problematiek, en daarmee wat betreft doel en criteria van de behandeling, alsook wat betreft de aard, mate en richting van de verandering die zij vertonen na psychotherapeutische behandeling. Al deze verschillen zijn niet te vangen binnen het kader van één of enkele algemene tests.

3 – Het lijkt niet juist om een 'gemiddeld effect' te berekenen uit de veranderingen van verschillende cliënten. Berekening van een gemiddelde leidt tot nivellering van evaluatie-uitkomsten, en dient dus beter achterwege te blijven.

De waarde van deze evaluatiepraktijk lijkt dan ook gering. Om het effect van psychotherapie te meten is dus een alternatieve werkwijze gezocht. De geïndividualiseerde schaal lijkt akseptabel. Het is evenwel belangrijk om dan rekening te houden met de hierna volgende kritiek.

3 Kritische kanttekeningen en alternatieven

1 Het nulpunt — Ter evaluatie van het effect van de residentiële behandeling op De Mark wordt momenteel een onderzoek uitgevoerd met behulp van G.A.S.-schalen. In dit onderzoek wordt als definiëring van het nulpunt gehanteerd: 'de toestand bij opname'. Een dergelijke omschrijving van het nulpunt wijkt principieel af van die welke Kiresuk en Sherman hanteren, namelijk 'verwacht resultaat van behandeling'. Tegen een dergelijke nulpunt-omschrijving heb ik het bezwaar dat zij aanleiding geeft tot onnodige subjectiviteit, omdat men gedwongen wordt tot het formuleren van verwachtingen, en tevens dat zij een extra beoordeling (namelijk een beginbeoordeling) noodzakelijk maakt (zie 13, voor een uitgebreide bespreking van deze bezwaren).

Dezelfde bezwaren met betrekking tot het alternatieve nulpunt dat Anderson, Dutszka en Ward, en Van der Hart en Lange (overigens op grond van andere bezwaren dan hierboven) voorstellen, namelijk 'verwacht resultaat zonder behandeling'. Ook dit nulpunt geeft

aanleiding tot onnodige subjektiviteit en maakt een extra beoordeling noodzakelijk. Er is echter nog een ander bezwaar tegenin te brengen. Genoemde auteurs hopen met dit nulpunt de mogelijkheid te openen tot het vergelijken van verschillende (soorten) behandelingen, omdat alle behandelingen dan eenzelfde nulpunt hebben: 'aangezien ik de in paragraaf 2 omschreven theoretische trend onderschrijf (namelijk om per individueel behandingsgeval te evalueren) lijkt mij een dergelijke vergelijking niet toelaatbaar (zie ook 13).

Dus, zowel het nulpunt van Kiresuk en Sherman, als dat van Anderson, Dutszka en Ward en van Van der Hart en Lange, lijkt mij minder aan te bevelen. Ik stel daarom voor om als nulpunt te kiezen 'de toestand bij opname' (c.q. bij begin van de behandeling): dit geeft géén aanleiding tot onnodige subjektiviteit, het maakt géén vergelijkingen mogelijk en het maakt géén beginbeoordeling noodzakelijk.

2 Statistische assumpties en berekeningen – In het onderzoek, dat op De Mark gaande is, worden de statistische assumpties, die Kiresuk en Sherman maken, achterwege gelaten. De assumpties, die zij maken met betrekking tot de statistische verdeling van de scores op de schalen, tot de verwachte waarde van de scores en tot de interkorrelatie van de scores op de afzonderlijke schalen van een cliënt, zijn niét bewezen en behoeven daarom ook niet overgenomen te worden. Bovendien is het de vraag of de functie van deze assumpties wel zo groot is. Zij zijn alleen nodig om de door Kiresuk en Sherman voorgestelde T-score-berekeningen (zie 11) mogelijk te maken. Deze scores willen Kiresuk en Sherman gebruiken, om vergelijkingen tussen verschillende (soorten) behandelingen mogelijk te maken. Tegen dergelijke vergelijkingen heb ik echter bezwaar, omdat ik vind dat per individueel behandingsgeval geëvalueerd dient te worden. Daarom moeten T-score berekeningen achterwege blijven, waardoor ook de functie van genoemde statistische assumpties wegvalt (zie 13).

3 Omschrijving van de schaalniveau's – In ons onderzoek trachten wij elk schaalniveau te voorzien van een zo exakt en objectief mogelijke omschrijving. Hierdoor wordt de kans op subjektiviteit in de beoordelingen gereduceerd.

Deze werkwijze is een uitbreiding op het voorstel van Kiresuk en Sherman, die zeggen dat het aanbeveling verdient om tenminste 2 punten op de schaal exakt te omschrijven. Het is ook een uitbreiding op het voorstel van Van der Hart en Lange, die stellen, dat 'minimaal het nulpunt, plus in ieder geval één positief en één negatief punt' omschreven dienen te worden. Overigens, beide groepen auteurs spreken van 'tenminste', respectievelijk 'minimaal', waarmee zij lijken te zeggen, dat omschrijving van alle niveau's de uiteindelijke voorkeur heeft (13).

4 Differentiatievermogen – In de praktijk blijkt, dat het differen-

tatievermogen van de G.A.S.-schalen gering is. Alle gradaties in verbetering, respectievelijk verslechtering, die een cliënt door-
maakt, voordat hij het maximale (positieve of negatieve) niveau
bereikt heeft, kunnen slechts ondergebracht worden in het +1,
respectievelijk -1 niveau. (Hierbij dient nog wel aangetekend te
worden, dat het differentiatievermogen ietwat groter is, echter niet
veel, wanneer het nulpunt van Kiresuk en Sherman gehandhaafd
wordt. Zij betrekken immer de verwachte verandering in hun
schaal, hetgeen impliceert dat er één veranderingsniveau meer is
dan wanneer men het nulpunt omschrijft als de toestand bij op-
name).

Door het geringe differentiatievermogen vallen verschillen in de
beoordelingen van verschillende beoordelaars weg. Ook gradaties
in de beoordelingen van één beoordelaar op verschillende momen-
ten van de behandeling, zijn niet uit te drukken. Een *proces-eval-
uatie* wordt daardoor wel zeer moeilijk, zo niet onmogelijk. Wan-
neer gradaties in niet maximale veranderingen wegvallen, kan
nooit een beeld gegeven worden van geleidelijke verbetering of
verslechtering of van schommelingen, die de cliënt vertoont tijdens
zijn veranderingsproces. (Dit geldt overigens niet in de opzet van
Kiresuk en Sherman, immers zij stellen voor om de G.A.S.-me-
thode te gebruiken als een evaluatie-instrument bij follow-up-
onderzoeken, zodat van een procesevaluatie bij hen geen sprake
is).

Het differentiatietekort kan men oplossen, indien men met meer-
dere schaalniveau's werkt (bijvoorbeeld een 7-, 9-, of 11-punts-
schaal). Het gevaar daarbij is dan mogelijk weer, dat de schaal on-
scoorbaar wordt, doordat het differentiatievermogen, dus ook de
keuzemogelijkheid, te groot wordt. Eventueel kan men gelijktijdig
met twee sets schalen werken, een ongedifferentieerde schaal (een
5-puntsschaal), en een gedifferentieerde schaal, waarbij elk van de
5 niveau's onderverdeeld wordt in meerdere subniveau's. Wanneer
de beoordelaars beide schalenpakketten ter beoordeling krijgen
voorgelegd, en zij vertonen onderling op de gedifferentieerde scha-
len verschillen, terwijl zij op de ongedifferentieerde schalen gelijk
scoren, dan is er sprake van de invloed van het differentiatie tekort
en dienen deze verschillen nader besproken te worden.

5 Beoordelaarseffekten — In het 'Mark-project' trachten wij de
G.A.S.-schalen te laten scoren door méerdere beoordelaars, om
de invloed van beoordelaarseffekten te verminderen. Ideaal zou
zijn om de schalen te laten scoren door meerdere beoordelaars,
die ónafhankelijk staan ten opzichte van de behandeling, dat wil
zeggen, dat zij er niet rechtstreeks bij betrokken zijn. Wanneer
men echter werkt met meerdere ónafhankelijke beoordelaars, is
het gevaar erg groot, dat die verschillende beoordelaars de te be-
oordelen cliënt in een verschillende mate zullen kennen. Wij wer-
ken daarom met een groep beoordelaars, waarvan een deel zeer
dicht bij de behandeling betrokken is (pedagogische leiding) en een
deel minder (hoofden van de pedagogische leiding). Dit om de

invloed van beoordelaarseffekten zoveel mogelijk te verminderen. De volgende beoordelaarseffekten deden hun invloed in een vooronderzoekje gelden:

- Het 'halo-effekt'. Een beoordelaar, die met betrekking tot een eerste schaal, die hij van een cliënt te beoordelen krijgt, van mening is, dat de cliënt daar een zeer goede vooruitgang vertoont, kan daardoor ten aanzien van de veranderingen van die cliënt op andere schalen (mogelijk ten onrechte) positief gestemd raken en deze stemming in positieve beoordelingen omzetten.

- Het 'welwillendheidseffekt': een beoordelaar, die de schalen van een cliënt te scoren krijgt, bij wiens behandeling hij meer betrokken is, is mogelijk ten aanzien van zijn eigen behandelwijze zo positief gestemd, dat hij het gedrag van de cliënt niet anders dan positief zal beoordelen;

- Het 'pas-twee (drie)-maanden-effekt' en het 'al-een (twee)-jaar-effekt': indien een beoordelaar weet, dat een cliënt 'pas' een paar maanden in behandeling is of 'al' een paar jaar, kan hij daardoor negatief ingesteld zijn ten aanzien van de eventuele verandering van die cliënt: hij zal nog wel niet, of wel niet meer veranderd zijn;

- Wanneer de beoordelaars in de loop van de behandeling geregeld, onderling contact met elkaar hebben over de betreffende cliënt, kan de overeenstemming in de beoordelingen met het verstrijken van de tijd groter worden: tijdsinvloeden bepalen zo mede de waarnemingen en beoordelingen.

Het is niet uitgesloten dat, naast deze vier, ook nog andere beoordeelaarseffekten werkzaam zijn.

4 Konstruktieproblemen

De hierna volgende paragrafen bevatten een beschrijving van moeilijkheden, waarmee men te maken kan krijgen, wanneer men G.A.S.-schalen gaat konstrueren en gebruiken bij evaluaties.

1 Doelselektie en schaalkonstruktie – Wanneer het selekteren van de probleemgebieden en de therapiedoeleinden en het konstrueren van de schalen in handen gelegd worden van één persoon, bestaat de kans, dat de inhoud van de doeleinden en schalen subjektief bepaald zal zijn. Om de invloed van deze subjektiviteit te reduceren, worden in het Mark-project deze activiteiten door meerdere personen gedaan. Elk van hen selekteert eerst afzonderlijk de probleemgebieden en therapiedoeleinden, waarna in onderling overleg een definitieve keuze gemaakt wordt. Op dezelfde wijze worden de schalen gekonstrueerd. Het is van belang, dat degene, die de probleemgebieden en therapiedoeleinden selekteert en de schalen konstrueert, uitgesloten wordt van de beoordeling. Dit voorkomt, dat de scores bevooroordeeld zullen zijn.

2 Onafhankelijke verandering van problemen en variabelen – Men dient zich bij de konstruktie van de schalen wel bewust te zijn van het feit, dat verschillende problemen (bijvoorbeeld angst en agressiviteit) en verschillende variabele aspecten van één probleem (bij-

voorbeeld frequentie en intensiteit van agressiviteit), onafhankelijk van elkaar kunnen veranderen. Een mogelijkheid om deze mogelijkheid op te vangen is, om voor elk probleem en voor elk variabel aspect van een probleem afzonderlijke schalen te creëren. De kans bestaat dan wel, dat het te scoren schalenpakket erg omvangrijk wordt. Hiertegen bestaat in principe geen bezwaar, maar het is niet uitgesloten dat de beoordelaar daardoor moe of geïrriteerd zal worden.

De verschillende problemen moeten zeker in afzonderlijke schalen ondergebracht worden. Om te voorkomen, dat de onafhankelijk van elkaar veranderende variabele aspecten van één probleem ook in verschillende schalen onderverdeeld moeten worden, handelen wij als volgt. In het +1 niveau brengt men een onderverdeling aan in +1a, +1b, +1c, etc.: in niveau +1a laat men dan variabele X variëren, terwijl men Y konstant houdt; in +1b laat men Y variëren, en houdt men X konstant; in +1c worden X en Y samen gevarieerd.

Bijvoorbeeld: het nul-niveau luidt: het kind heeft de neiging om iedere avond (frequentie) gedurende een halfuur (duur) rituelen te vertonen voor het naar bed gaan; +1a luidt dan: het kind vertoont de rituelen niet meer iedere avond, maar wel nog gedurende een halfuur; +1b: hij vertoont de rituelen nog wel iedere avond, maar korter dan een halfuur; +1c: hij vertoont de rituelen niet meer iedere avond, bovendien korter dan een halfuur; het +2 niveau luidt dan: hij vertoont zijn rituelen niet meer.

Dit voorstel is een correctie op de oplossing, die Bensing en Dekker suggereren. Zij variëren in +1 bijvoorbeeld de frequentie en houden de intensiteit konstant, terwijl zij in +2 de intensiteit variëren en de frequentie konstant houden. Hiermee leggen zij de volgorde van de verandering vast in verschillende gewaardeerde niveau's, terwijl in de praktijk zou kunnen blijken, dat die volgorde juist andersom gewaardeerd zou moeten worden (13). Onze onderverdeling in +1a, +1b en +1c betekent niet, dat deze sub-niveau's in toenemende mate hoger gewaardeerd worden: al deze niveau's vallen immers samen binnen het +1 niveau.

3 Selectie, konstruktie en beoordeling door de cliënt – Van der Hart en Lange bespreken in hun artikel de mogelijkheid, om doelstellingen en schalen door de cliënt zelf op te laten stellen. Zij stellen ook voor, om de cliënt zelf te betrekken in de scoring van de schalen. Deze procedure lijkt me wenselijk, omdat het gevoel van verantwoordelijkheid en de mate van betrokkenheid voor het verloop van de behandeling bij de cliënt hierdoor vergroot zullen worden (7). Ook kunnen de formuleringen van doelstellingen en schalen en de beoordelingen door de cliënt zelf, een belangrijke controle zijn op de formuleringen en beoordelingen, die door de therapeut of andere beoordelaars gegeven worden.

Hoewel een dergelijke werkwijze wenselijk is, zal ze echter niet altijd te realiseren zijn. Bijvoorbeeld wanneer de cliënt een kind is, die zijn problemen niet als zodanig kent, of niet wil kennen, of niet

kan verwoorden. Of wanneer de cliënt een ernstig gestoorde volwassene is, die zijn problemen niet kan of niet wil verwoorden. Of wanneer de cliënt 'tot cliënt gemaakt is' door een vonnis van de rechter, bij welke cliënten vaak een grote weerstand bestaat om mee te werken.

4 Symptoomverschuiving – Wanneer men een cliënt in behandeling heeft, is het niet uitgesloten, dat zich het verschijnsel van symptoomverschuiving voordoet. Indien dit gebeurt in een behandeling, die men evalueert met behulp van G.A.S.-schalen, is het noodzakelijk om één of meerdere nieuwe schalen te konstrueren voor dit nieuwe probleemgedrag. Dit dient gedaan te worden meteen op het moment, dat dit probleemgedrag zich voordoet en vóórdat men aan de behandeling ervan begint, precies zoals men G.A.S.-schalen ontwikkelt, voordat men aan de behandeling van een cliënt begint.

5 Voordelen van de G.A.S.-methode

Naar mijn mening is de G.A.S.-methode een zinvol evaluatie-instrument, zeker wanneer men zich bewust is van de hierboven genoemde punten van kritiek en wanneer men in de G.A.S.-methode de voorgestelde (of misschien andere) alternatieven in praktijk brengt. De betekenis en bruikbaarheid wil ik hieronder in het kort samenvatten.

1 – De G.A.S.-methode is een evaluatie-instrument, waarbij elke cliënt en elke behandeling afzonderlijk, als verschillend van andere cliënten en behandelingen, geëvalueerd worden. Deze methode dwingt met andere woorden niet om gebruik te maken van één grote steekproef van cliënten, onderzocht met hetzelfde evaluatie-instrument, waarbij de verkregen scores in een soort 'gemiddeld effect' worden omgezet.

2 – De G.A.S.-methode dwingt de therapeut en de onderzoeker om de doeleinden van de te evalueren psychotherapie, als ook de criteria voor die therapie, vooraf te specificeren. Een procedure, die trouwens bij elke therapie noodzakelijk is, of men deze nu evalueert of niet.

3 – Wanneer men de problemen van een cliënt na een bepaalde tijd beoordeelt op G.A.S.-schalen, krijgt men een concrete weergave van de voor- of achteruitgang van die cliënt of van zijn niet-veranderd-zijn. Dergelijke concrete weergaven ontbreken vaak, wanneer men geen gebruik maakt van G.A.S.-schalen. Men verkrijgt hierdoor duidelijke informatie met betrekking tot de vraag of de betreffende therapie wel of niet beëindigd kan worden. Men vindt ook aanknopingspunten om (eventueel in overleg met de cliënt) te komen tot een herziening (of continuering) van het behandelingsplan. Ook kan een discussie in gang gezet worden tussen de beoordelaars, naar aanleiding van gekonstateerde verschillen in de beoordeling.

4 – De G.A.S.-methode is bruikbaar in iedere situatie, waar men tracht na te gaan of de doeleinden, die men zich tevoren gesteld

heeft, bereikt zijn (2).

Tot slot wil ik opmerken, dat men uit een op een G.A.S.-schaal gebleken vooruitgang of verslechtering van een cliënt, niet zonder meer mag afleiden, dat die verandering veroorzaakt is door de verkregen psychotherapie of behandeling (in de Mark evalueren wij de totale residentiële behandeling, dat wil zeggen pedagogische situatie, psychotherapie, oudertherapie, etc.). Hoogstens mag men daaruit afleiden, dat zich tijdens de therapie of behandeling veranderingen hebben voorgedaan, zonder dat men zekerheid heeft over een eventueel causaal verband. Zekerheid over een eventuele causale relatie tussen de verkregen behandeling en de veranderingen, die een cliënt vertoont, lijkt geen enkele methode te kunnen verschaffen, ook de G.A.S.-methode niet.

Summary

Goal Attainment Scaling: a method for evaluating psychotherapies

Goal Attainment Scaling is a method, where every client, every problem and every variable aspect of those problems are evaluated as different from other clients, other problems and other variable aspects: it is an individualized evaluating method. In this article a short description is given of the Goal Attainment Scaling method. It is treated as related to modern trends in psychotherapeutic theories and research. Some aspects of the method as developed by Kiresuk and Sherman, and as described, criticized and alternatively changed by Anderson, Dutszka and Ward and by Van der Hart and Lange and Bensing and Dekker, are criticized: the definition of the zero-level, the statistical assumptions made by Kiresuk en Sherman, the use of T-scores, the description of the scale points, the possibility to differentiate between gradual changes when Goal Attainment Scales are used, and the change for biases induced by the judge(s); alternatives are proposed for the criticized aspects. Also a description is given of problems, that one can meet in constructing and using Goal Attainment Scales, and some solutions are formulated. The last paragraph is dedicated to the qualities and the benefits of the Goal Attainment Scaling method as compared to other methods.

Literatuur

- 1 Anderson, D. H., J. R. Dutszka, L. G. Ward, *Individualized Goal Scaling: a research proposal*, Intern rapport, Toronto, 1969
- 2 Bensing, J., M. Dekker, Goal Attainment Scaling. *Maandblad voor de Geestelijke Volksgezondheid*, 1973, 342-356
- 3 Bergin, A. E., The effects of psychotherapy: negative results revisited, *Journal of Counseling Psychology*, 1963, 10, 244-250
- 4 Bergin, A. E., Some implications of psychotherapy research for therapeutic practice, *Journal of Abnormal Psychology*, 1966, 71, 235-246
- 5 Bergin, A. E., The evaluation of therapeutic outcomes, in: Bergin, A. E. and Garfield, S.: *Handbook of psychotherapy and behavior change* 1971, Wiley and Sons, New York
- 6 Cartwright, D. S., Note on 'Changes in psychoneurotic patients with and without psychotherapy', *Journal of Consulting Psychology*, 1956, 20, 403-404
- 7 Hart, O. van der, F. Lange, Goal Attainment Scaling, *Maandblad voor de Geestelijke Volksgezondheid*, 1972, 502-510
- 8 Kiesler, D. J., Some myths of psychotherapy and the search for a para-

- digs, *Psychological Bulletin*, 1966, 65, 110-136
- 9 Kiesler, D. J., Experimental designs in psychotherapy research, in A. Bergin and S. Garfield: *Handbook of psychotherapy and behavior change* 1971, Wiley and Sons, New York
- 10 Kiresuk, T. J., R. E. Sherman, Goal Attainment Scaling: a general method for evaluating comprehensive community mental health programs, *Community Mental Health Journal*, 1968, 4, 443-453
- 11 Spitz, J. C., *Statistiek voor psychologen, pedagogen, sociologen*, Noord-Hollandsche Uitgeversmaatschappij, 2e druk, 1967
- 12 Wit, C. de, *Het stellen van criteria bij evaluatie van het effect van psychotherapie*, Doctoraal Scriptie, Universiteit van Amsterdam, Afd. Ontwikkelingspsychologie, 1973
- 13 Wit, C. de, *Goal Attainment Scaling*, Interne publicatie, De Mark, Breda, 1974