

## Suppressor variabelen in padmodellen: Definities en interpretaties

Gerard H. Maassen

(Methodenleer en Statistiek, Faculteit Sociale Wetenschappen,  
Universiteit van Utrecht)

en

Arnold Bakker

(Sociale en Organisatie Psychologie,  
Universiteit van Utrecht)

### Samenvatting

Het fenomeen van *suppressorvariabelen* is bekend in de context van multiële-regressie-vergelijkingen. Met voorbeelden wordt gedemonstreerd dat de varianten die in de literatuur zijn beschreven, zich niet alleen voordoen in voorspellingsvergelijkingen, maar ook in structurele-relatiesmodellen. De kans hierop is in modellen met latente variabelen, waarin wordt gecorrigeerd voor meetfouten in de *suppressed variable*, extra groot. Ruime aandacht wordt besteed aan een bijzondere kanshebber, het 2W2V-model. Het optreden van suppressie in structurele-relaties-modellen is meestal onvoorzien en levert de onderzoeker interpretatieproblemen. We bespreken definities van het suppressorfenomeen, geven aan hoe de argeloze onderzoeker gewaarschuwd zou kunnen worden en geven vingerwijzingen voor de interpretatie van de resultaten.

---

Correspondentieadres:

dr. G.H. Maassen

Capaciteitsgroep Methodenleer en Statistiek

Faculteit Sociale Wetenschappen, Universiteit van Utrecht

Postbus 80140

3508 TC Utrecht

e-mail: [g.maassen@fss.uu.nl](mailto:g.maassen@fss.uu.nl)

### 1. Typen en definities van suppressorvariabelen

Hoofdelementen in het basismodel van de multiële regressie zijn een afhankelijke variabele  $Y$  enerzijds en twee of meer onafhankelijke variabelen  $X_i$  anderzijds. Dit model werd aanvankelijk vooral toegepast in situaties waarin onbekende scores op  $Y$  moesten worden voorspeld vanuit bekende scores op de  $X_i$  (bijvoorbeeld het voorspellen van schoolprestaties of selectie van personeel). In zulke situaties wordt de afhankelijke variabele  $Y$  ook wel *criterium* genoemd en de onafhankelijke variabelen *predictoren*. De ideale situatie is die waarin men beschikt over predictoren die hoog correleren met het criterium (een hoge *predictieve validiteit* hebben) en onderling ongecorrleerd zijn. "Each predictor in this situation would be validly a different component of the criterion with a minimum of predictive redundancy," aldus Wiggins (1973). In de praktijk is het zelden mogelijk een verzameling van zulke predictoren te vinden. Men is al blij indien men predictoren met een goede predictieve validiteit gevonden heeft; dat ze onderling min of meer correleren neemt men dan op de koop toe.

Horst (1941) wees als eerste op het fenomeen dat ook variabelen die geen predictieve validiteit hebben, toch kunnen bijdragen aan een verbetering van de predictie. Hij gaf het klassieke voorbeeld, dat de selectie van vliegeniers in de Tweede Wereldoorlog kon worden verbeterd door in de voorspellingsvergelijking niet alleen een variabele die hun technische vaardigheid meet op te nemen, maar ook een variabele waarmee hun verbale vaardigheid wordt vastgesteld, ook al is de laatste op zichzelf niet met het criterium (vliegvaardigheid) gecorrleerd. Horst introduceerde voor een dergelijke variabele de term *suppressor variable*. McNemar (1945, 1969) gaf voor het fenomeen een verklaring in termen van wat hij *common elements* noemde. Volgens McNemar bevat een bruikbare predictor uiteraard veel elementen gemeenschappelijk met het criterium, maar daarnaast in de meeste gevallen ook irrelevante elementen. Een *suppressant* (dit is de term die McNemar bezigde) is een variabele die geen elementen gemeenschappelijk heeft met het criterium, maar wel irrelevante elementen met de predictor. Zijn predictor en suppressant positief gecorrleerd dan krijgt de suppressant bij opname in de regressievergelijking een negatief regressiegewicht. Dit weerspiegelt dat de irrelevante elementen uit de predictor worden weggepartialiseerd, waardoor de voorspelling wordt 'gezuiverd' en verbeterd. Kort samengevat: de klassieke verschijningsvorm van een suppressorvariabele is een variabele die niet correleert met de afhankelijke variabele maar wel met één of meer predictoren.

Het verschijnsel suppressorvariabele heeft veel wetenschappers die zich met onderwijskundig of psychologisch onderzoek bezighouden, gefraspeerd. Naarstig ging men op zoek naar suppressorvariabelen waarmee men ongewenste effecten uit de metingen die centraal stonden in het onderzoek, zou kunnen uitschakelen. Zo werden schalen voor sociale wenselijkheid ontworpen die als suppressorvariabele in een regressieanalyse wellicht het effect van sociale wenselijkheid uit de respons konden verwijderen. De resultaten van dergelijke pogingen bleken echter teleurstellend. Wiggins (1973, p.38) concludeert: "Considered as a whole, evidence

relating to the empirical effectiveness of suppressor variables in predicting socially relevant criterion measures is not impressive. In those instances in which suppressor variables have been developed successfully (...), the net predictive gain does not appear to be greater than that which would have been achieved by employing a more conventional variable as an additional predictor."

Behalve de praktische betekenis van suppressorvariabelen bleef ook de definitie ervan de gemoederen van onderzoekers bezighouden. Het werd duidelijk dat het voorbeeld van Horst niet meer was dan een grensgeval, dat in de praktijk alleen maar bij benadering voorkomt, en dat de verklaring van McNemar op nog meer situaties betrekking heeft. Er ontstond behoefte aan een scherpere, formele definitie, die ook andere situaties zou omvatten, waarin sprake is van het onderdrukken van irrelevante informatie. Ter inleiding van deze definities demonstreren we aan de hand van enkele voorbeelden tot welke typerende uitkomsten voor de schattingen van de parameters in de regressieanalyse het suppressorfenomeen kan leiden. De voorbeelden zijn te vinden in Tabel 1 en zijn ontleend aan Tabel 1 van Velicer (1978). Velicer beperkte zich hierbij tot 3-variabelen-situaties (met een criterium-variabele die we in dit bestek met  $Y$  en twee predictoren die we met  $X_1$  en  $X_2$  zullen aanduiden).

Tabel 1

Voorbeelden van suppressorsituaties (ontleend aan Velicer, 1978, Table 1)

| suppressor<br>situatie | 1<br>$r_{y_1}$<br>( $r^2_{y_1}$ ) | 2<br>$r_{y_2}$<br>( $r^2_{y_2}$ ) | 3<br>$r_{12}$ | 4<br>$b_{y_1.2}$ | 5<br>$b_{y_2.1}$ | 6<br>$r^2_{y(1.2)}$ | 7<br>$r^2_{y(2.1)}$ | 8<br>$R^2_{y.12}$ |
|------------------------|-----------------------------------|-----------------------------------|---------------|------------------|------------------|---------------------|---------------------|-------------------|
| klassiek               | .40<br>(.16)                      | .00<br>(.00)                      | .707          | .800             | -.566            | .320                | .160                | .320              |
| negatief               | .50<br>(.25)                      | .10<br>(.01)                      | .710          | .865             | -.514            | .371                | .131                | .381              |
| reciproke              | .50<br>(.25)                      | .30<br>(.09)                      | -.270         | .627             | .469             | .364                | .204                | .454              |

In de kolommen 1, 2 en 3 van Tabel 1 zijn de bivariate correlaties opgenomen van drie fictieve voorbeelden. In de kolommen 4 en 5 staan partiële regressiecoëfficiënten; in de kolommen 6 en 7 de semipartiële correlaties, dat wil zeggen de correlaties tussen  $Y$  en een onafhankelijke variabele waaruit de invloed van de andere onafhankelijke variabele is weggepartialiseerd. Een bekende relatie in de multiële-regressierekening is de opeelbaarheid van orthogonale



variantiecomponenten:

$$R_{y.12}^2 = r_{y1}^2 + r_{y(2.1)}^2 = r_{y2}^2 + r_{y(1.2)}^2. \quad (1)$$

Deze relatie zien we terug in: kolom 8 = kolom 1 + kolom 7 = kolom 2 + kolom 6.

Naar aanleiding van de tabel merken we het volgende op.

- 1) Het eerste voorbeeld illustreert de hierboven beschreven *klassieke* suppressorsituatie.
- 2) Het tweede voorbeeld is een situatie waarin twee predictoren een positieve predictieve validiteit hebben en onderling positief correleren; een van beide predictoren krijgt een negatief regressiegewicht. Deze situatie wordt met *negatieve suppressie* aangeduid. De suppressor heeft weliswaar relevante informatie met het criterium gemeen, maar minder dan de irrelevante informatie die het met de andere predictor gemeen heeft.
- 3) De derde situatie betreft twee variabelen die als goede voorspeller kunnen gelden. Ze bevatten beide echter ook variantie die irrelevant is voor het criterium, maar met een tegengestelde oriëntatie. Wanneer beide variabelen in de regressie worden opgenomen, onderdrukken zij een deel van elkaars irrelevante informatie. Men spreekt dan van *reciproke suppressie*.
- 4) In alle situaties van Tabel 1 zien we na opname van een tweede predictor de absolute waarde van de regressiecoëfficiënt stijgen.
- 5) We zien in alle situaties van Tabel 1, dat de gekwadrateerde semipartiële correlatie tussen het criterium en een variabele  $X_1$  waaruit de andere variabele is weggepartialiseerd, groter is dan de proportie variantie van het criterium die door  $X_1$  wordt verklaard.

In bovenstaande voorbeelden komt het suppressorfenomeen duidelijk in de resultaten tot uitdrukking. Het fenomeen kan echter door andere effecten versluierd worden. Zo zijn steekproeffluctuaties, meetfouten en de richting waarin variabelen zijn geschaald, uiteraard van invloed op de parameterschattingen en deze invloeden kunnen het onderkennen van het suppressorfenomeen compliceren. Gemakshalve hebben wij ons hierboven nog niet bekommerd om het onderscheid tussen populatie en steekproef, om meetfouten, of om de schaling van de variabelen. Wat dat laatste betreft, elk getalvoorbeeld wordt gekarakteriseerd door een patroon van correlaties en is ook representatief voor situaties waarin dat patroon na omschaling van variabelen ontstaat.

Onderzoekers die het suppressorverschijnsel wilden definiëren, waren zich van genoemde effecten bewust en probeerden hun definities daarvan te ontdoen. Darlington (1968) gaf de volgende definitie: Indien elke predictor, die in de populatie een correlatie ongelijk aan nul heeft met het criterium, zodanig wordt gescoord dat deze correlatie positief wordt, dan wordt een suppressorvariabele gedefinieerd als een variabele die bij opname in een regressievergelijking voor de populatie een negatief gewicht krijgt. Door deze definitie werd dus de

*negatieve suppressie* aan de klassieke verschijningsvorm van het suppressorfenomeen toegevoegd.

Conger (1974) nam de eigenschap vermeld onder 4) als uitgangspunt voor zijn definitie van suppressorvariabele: "A suppressor variable is defined to be a variable which increases the predictive validity of another variable (or set of variables) by its inclusion in a regression equation. This variable is a suppressor only for those variables whose regression weights are increased". In formulevorm: variabele  $X_2$  is suppressor voor predictor  $X_1$  (in relatie tot criterium  $Y$ ) indien:

$$\beta_1 \cdot r_{y1} > r_{y1}^2 \quad (2a)$$

Conger stipuleerde, aansluitend op de basisidee van McNemar (1945), het volgende onderliggende model: "there is nonerror variance in the predictor which is unrelated to the criterion (factor  $S$ ) as well as an uncorrelated common factor of that which the criterion is measuring (factor  $T$ ). The suppressor either has no relation to the criterion (a loading of zero for factor  $T$ ) or is measuring the criterion less than it is measuring the irrelevant variance.". Vervolgens onderzocht hij in welke situaties aan zijn definitie wordt voldaan en zo ontdekte hij nog een derde categorie van suppressorvariabelen: de *reciproke suppressors*.

Velicer (1978) merkte op dat een predictieve validiteit en een regressiecoëfficiënt niet goed vergelijkbaar zijn en nam eigenschap 5) als uitgangspunt van zijn definitie: Een suppressorvariabele is een predictor waarvoor geldt dat de gekwadrateerde predictieve validiteit kleiner is dan het kwadraat van de semipartiële correlatie tussen het criterium en de predictor waaruit de overige predictoren zijn uitgepartialiseerd.

Velicer geeft als commentaar dat deze definitie kan worden toegepast op elk van de drie typen suppressor variabelen die door Conger (1974) zijn besproken. De definitie heeft als voordelen dat ze berust op een vergelijking van proporties verklaarde variantie en consistent is met *stepwise* regressieprocedures. Smith, Ager en Williams (1992) benadrukken nog weer andere voordelen van Velicer's (1978) definitie. Als nadeel noemt Velicer zelf: "the definition will identify when a suppressor variable is present, but not specifically which variable is the suppressor. Designation of a variable as the suppressor would require knowledge of how a suppressor variable works." Tzelgov and Stern (1978) hebben aangetoond dat de definitie van Conger omvattender is dan die van Velicer. Zij lieten zien dat er situaties bestaan waarin een variabele volgens Conger worden aangemerkt als negatieve suppressor voor een andere variabele, maar niet voldoet aan de definitie van Velicer.

Het suppressorfenomeen treedt niet alleen op in 3-variabelen-situaties. De definities van Velicer en Conger zijn ook toepasbaar in situaties met meer dan twee predictoren. Stel dat in zo'n situatie variabele  $S$  voldoet aan Velicers definitie van suppressorvariabele. Met behulp van formule (1) kan men aantonen dat de overige predictoren, als lineaire combinatie bijeengenomen, dan ook voldoen aan deze definitie. De lineaire combinatie kan dus ook



worden gezien als de suppressor van variabele  $S$ . Met andere woorden, het is mogelijk dat meerdere predictoren in combinatie met elkaar een suppressorwerking hebben.

Ook Congers definitie kan worden uitgebreid naar situaties met meer dan twee predictoren (Conger, 1974; Tzelgov & Henik, 1991). Stel dat een onderzoeker de predictoren heeft verdeeld in twee disjuncte verzamelingen. Zij variabele  $P$  de lineaire combinatie van predictoren uit de ene set die volgens de multipele-regressietechniek criterium  $Y$  optimaal voorspelt; de correlatie tussen  $P$  en  $Y$  is gelijk aan de multipele correlatiecoëfficiënt  $R_{yp}$ . Zij  $S$  de lineaire combinatie van predictoren uit de andere set die  $Y$  optimaal voorspelt (met  $R_{ys}$  als multipele correlatiecoëfficiënt). Stel verder dat  $\beta_p$  en  $\beta_s$  de regressiecoëfficiënten zijn van respectievelijk de variabelen  $P$  en  $S$  in de multipele regressie van  $Y$  op  $P$  en  $S$ . Er is dan sprake van een suppressiesituatie en  $S$  is suppressorvariabele met betrekking tot  $P$ , indien:

$$\beta_p > R_{yp} . \quad (2b)$$

## 2. Suppressorvariabelen in padmodellen

Tot nu toe hebben we onze discussie toegespitst op een model waarin een aantal predictoren van invloed is op een afhankelijke variabele en men zoekt naar een optimalisering van de voorspelling (het algemene concept van multipele regressie). Dat wil zeggen, er wordt geen rekening gehouden met de wijze waarop de onafhankelijke variabelen elkaar beïnvloeden. In het kader van padmodellen of modellen voor structurele relaties tussen latente concepten staat dit laatste juist centraal en is het onderscheid tussen afhankelijke en onafhankelijke variabelen vervaagd. Natuurlijk kan padanalyse worden gezien als een sequentie van multipele regressies, waarbij eenzelfde variabele kan optreden als afhankelijke variabele in de ene en als onafhankelijke in een volgende regressievergelijking. Toch worden padanalyse en suppressorvariabelen zelden met elkaar in verband gebracht. Dat is niet verwonderlijk, want voor zover suppressorvariabelen praktische relevantie hebben, gaat dat veelal niet verder dan het verbeteren van de predictie. Dat is niet het oogmerk van padanalyse. In padanalyse toetst men in principe een model voor de verklaring van de relaties tussen variabelen of concepten. Suppressorvariabelen hebben meestal geen inhoudelijke betekenis en omdat men zich voor padanalyse gewoonlijk toch al moet beperken tot een selectie uit de variabelen die om hun verklarende waarde in aanmerking komen, zal men niet gauw een variabele opnemen die alleen maar de verklarende betekenis van een andere variabele ophoogt.

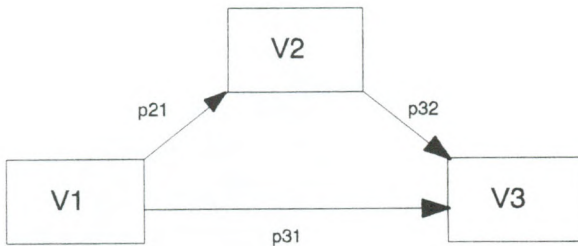
Het optreden van suppressorvariabelen in een padmodel lijkt dus een inhoudelijke tegenstrijdigheid. In dit artikel zullen we aan de hand van enkele voorbeelden demonstreren dat suppressorvariabelen in padanalyses, gewild maar meestal ongewild, toch een belangrijke rol kunnen spelen. Het is van belang ook in padanalyses te onderkennen wanneer men met een suppressorstructuur te maken heeft. Nadat men het fenomeen heeft herkend zal men er bij de interpretatie van de resultaten niet aan voorbij kunnen gaan. Aan deze aspecten zal eveneens

aandacht worden besteed.

Alvorens de voorbeelden te presenteren bespreken we eerst de consequenties van de definitie van suppressie voor de parameters in een padmodel. Dit zal ons later helpen bij de herkenning van het fenomeen. De definitie van Conger maakt een gemakkelijke vergelijking tussen regressie- en padcoëfficiënten mogelijk en vanwege dit voordeel nemen we deze definitie als uitgangspunt. We gaan voorlopig uit van een verzadigd 3-variabelen-model zoals weergegeven in Figuur 1. Omdat het onderscheid tussen afhankelijke en onafhankelijke variabelen in padanalyses minder duidelijk is, worden de variabelen aangeduid met  $V_1$ ,  $V_2$  en  $V_3$ ; regressiecoëfficiënten duiden we in het kader van padanalyse aan met  $p$ 's.

Figuur 1

Padmodel voor 3 variabelen



Verder nemen we aan dat beide verklarende variabelen  $V_1$  en  $V_2$  zodanig geschaald zijn dat  $r_{13}$  en  $r_{23} > 0$  zijn. Padanalyse leidt nu tot de volgende vergelijkingen (zie Pedhazur, 1982, p.594):

$$\begin{aligned}
 r_{12} &= p_{21} \\
 r_{13} &= p_{31} + p_{32} \cdot p_{21} \\
 r_{23} &= p_{32} + p_{31} \cdot p_{21}
 \end{aligned}
 \tag{3}$$

We bekijken eerst het geval waarin de *endogene* verklarende variabele  $V_2$  de rol van *suppressed variable* speelt. Volgens Conger geldt dan:  $p_{32} > r_{23} > 0$ . Hieruit volgt:

$r_{23} \cdot r_{13} \cdot r_{12} > r_{23}(1 - r_{12}^2)$ , ofwel  $r_{23} \cdot r_{13} \cdot r_{12} > r_{23} - r_{23} \cdot r_{12}^2$ , ofwel:

$$r_{13} \cdot r_{12} < r_{23} \cdot r_{12}^2
 \tag{4}$$

Uitgaande van de vergelijkingen (3) en (4) analyseren we de verschillende mogelijke combinaties van waarden voor  $p_{31}$  en  $p_{21}$ . Uit de laatste vergelijking van het stelsel (3) zal duidelijk zijn dat noch  $p_{31}$ , noch  $p_{21}$  gelijk kunnen zijn aan 0. In deze gevallen is  $r_{23} = p_{32}$  en is er dus geen suppressie. Ook mogen  $p_{31}$  en  $p_{21}$  niet hetzelfde teken hebben, want dan zou volgen:  $p_{32} < r_{23}$ . Dan blijven nog slechts de volgende twee mogelijkheden over.

a)  $p_{32} > 0$ ,  $p_{31} < 0$  en  $p_{21} > 0$ . In dit geval is sprake van *negatieve* suppressie. De correlatie tussen  $V_2$  en  $V_3$  berust gedeeltelijk op een negatief effect tengevolge van een schijnverband. Uit vergelijking (4) kan een verband tussen de correlatiecoëfficiënten worden afgeleid. Uit  $r_{12} > 0$  volgt nu:

$$r_{12} > \frac{r_{13}}{r_{23}} . \quad (4a)$$

In deze situatie is ook  $V_1$  *suppressed variable* indien:  $p_{31} < -r_{13}$ .  $V_1$  en  $V_2$  zijn dan voor elkaar zowel suppressor als *suppressed variable*.

In het specifieke geval dat  $r_{13} = 0$  is er sprake van *klassieke* suppressie.

b)  $p_{32} > 0$ ,  $p_{31} > 0$  en  $p_{21} < 0$ . Dit is een geval van *reciproke* suppressie. De correlatie tussen  $V_2$  en  $V_3$  berust dan weer gedeeltelijk op een negatief effect tengevolge van een schijnverband. Merk op dat door  $r_{12} < 0$  automatisch is voldaan aan vergelijking (4). In termen van correlatiecoëfficiënten kan deze situatie dus worden gekarakteriseerd als:

$$r_{13} > 0, r_{23} > 0 \quad \text{en} \quad r_{12} < 0 .$$

Merk verder op dat uit de tweede vergelijking van stelsel (3) volgt:  $p_{31} > r_{13}$ , dus dat ook variabele  $V_1$  *suppressed* is.

Is bovendien  $r_{13} = 0$  dan belanden we in het grensgeval van *klassieke* suppressie.

Vervolgens nemen we het geval onder de loop waarin *exogene* variabele  $V_1$  de rol van *suppressed variable* speelt. Volgens Conger geldt dan:  $p_{31} > r_{13} > 0$ . We analyseren nu de verschillende mogelijke waarden voor  $p_{32}$  en  $p_{21}$ . Ook hier behoeven we alleen te kijken naar de gevallen waarin  $p_{32}$  en  $p_{21}$  een verschillend teken hebben:

b')  $p_{31} > 0$ ,  $p_{32} > 0$  en  $p_{21} < 0$ . Dit geval valt samen met b), de situatie van *reciproke* suppressie. De correlatie tussen  $V_1$  en  $V_3$  berust nu gedeeltelijk op een negatief indirect effect.

Is bovendien  $r_{23} = 0$  dan belanden we in het grensgeval van *klassieke* suppressie.

c)  $p_{31} > 0$ ,  $p_{32} < 0$  en  $p_{21} > 0$ . In dit geval is sprake van *negatieve* suppressie. De correlatie tussen  $V_1$  en  $V_3$  berust gedeeltelijk op een negatief indirect effect. Op dezelfde wijze als voor ongelijkheid (4a) kan voor deze situatie worden afgeleid:



$$r_{12} > \frac{r_{23}}{r_{13}} . \quad (4b)$$

$V_2$  is eveneens *suppressed variable* indien geldt:  $p_{32} < -r_{23}$ .  $V_1$  en  $V_2$  zijn dan voor elkaar zowel suppressor als *suppressed variable*.

Is bovendien  $r_{23} = 0$ , dan is er sprake van *klassieke* suppressie.

#### *Verschillen tussen analyses met en zonder latente variabelen*

Veel padanalyses worden, indien mogelijk, uitgevoerd op latente concepten. Deze analyses hebben het voordeel dat de intercorrelaties van de latente concepten gecorrigeerd worden voor attenuatie ten gevolge van meetfouten, waardoor ze op een hoger niveau komen te liggen dan in situaties met alleen manifeste variabelen. De vraag is nu of er verschillen zijn te verwachten tussen analyses met en analyses zonder latente variabelen wat betreft het optreden van het suppressorfenomeen. Indien met  $\rho_{ii}$  de *reliability* van variabele  $X_i$  wordt aangegeven, kan de bekende formule voor attenuatiecorrectie wordt ingevoerd (McNemar, 1969, p.172) in de ongelijkheden (4a en b). In de 3-variabelen-situatie en het meest voorkomende geval van de negatieve suppressie betekent dit dat ongelijkheid (4a) gaat luiden:

$$\frac{r_{12}}{\rho_{22}} > \frac{r_{13}}{r_{23}} , \quad (5a)$$

indien verklarende variabele  $V_2$  de rol van *suppressed variable* speelt. De waarde van het linkerlid wordt daardoor verhoogd omgekeerd evenredig aan de *reliability* van de *suppressed variable*, terwijl het quotiënt in het rechterlid niet verandert. Indien variabele  $V_1$  de rol van *suppressed variable* speelt, is ongelijkheid (4b) van toepassing, die dan als volgt gaat luiden:

$$\frac{r_{12}}{\rho_{11}} > \frac{r_{23}}{r_{13}} . \quad (5b)$$

Indien in analyses met latente concepten wordt gecorrigeerd voor meetfouten in de *suppressed variable*, wordt de kans vergroot dat het suppressoreffect zal optreden.

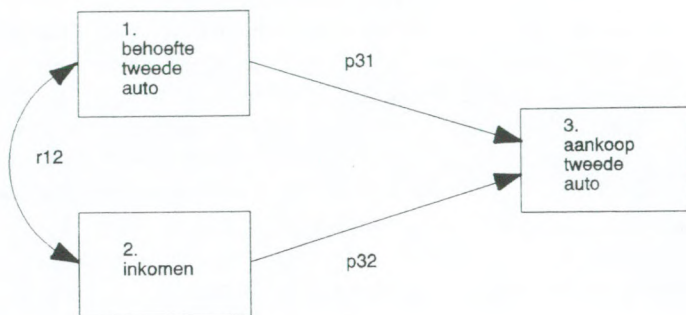
#### *Onderdrukkende of maskerende variabele?*

Verwant aan het begrip suppressorvariabele is het begrip *maskerende variabele*. We illustreren dit aan de hand van een (fictief) voorbeeld, dat is ontleend aan Nie, Hull, Jenkins, Steinbrenner and Bent (1975, p.305). Hierin is de afhankelijke variabele *aankoop van een tweede auto*, de onafhankelijke variabelen zijn *behoefte aan een tweede auto* en *inkomen*. Men wil onderzoeken in hoeverre de aankoop van een tweede auto wordt beïnvloed door de behoefte daaraan en men

veronderstelt uiteraard een positieve invloed. In Figuur 2 en het onderste paneel van Tabel 2 vindt men respectievelijk het model en de data bij dit voorbeeld.

Figuur 2

Padmodel bij het voorbeeld: *De aanschaf van een tweede auto*  
(ontleend aan Nie et al., 1975)



De lage correlatie tussen *behoefte aan een tweede auto* en *aankoop van een tweede auto* zal menigeen verbazen. Wordt echter ook de variabele *inkomen* in de analyse opgenomen, dan wordt de zaak duidelijker. De hoogte van het inkomen correleert positief met de aanschaf van een tweede auto maar blijkt negatief te correleren met de behoefte. Door het inkomen in de analyse te betrekken wordt informatie in de behoefte aan een tweede auto, die de correlatie met het wel of niet kopen van een tweede auto maskeert, weggefilterd. De partiële correlatie tussen de behoefte aan een tweede auto en de aankoop van een tweede auto, met uitschakeling van verschillen in inkomen, bedraagt .32. Algebraïsch gezien kan dit voorbeeld als klassieke-suppressiesituatie of als grensgeval van een reciproke-suppressiesituatie worden gezien. De variabele die volgens de klassieke optiek als de suppressor zou worden aangemerkt, is hier inhoudelijk gezien echter de belangrijkste verklarende variabele, terwijl de onafhankelijke variabele met de grootste predictieve validiteit hier de rol van suppressor speelt. Waarschijnlijk vanwege deze omkering hebben Nie et al. behoefte aan een een nieuwe term: *maskerende variabele*.

In dit verband is vermeldenswaard de enige ons bekende definitie van een suppressorvariabele, waarbij verband wordt gelegd met een padmodel, en die we zijn tegengekomen in de Nederlandse literatuur (Verschuren, 1991). Verschuren's definitie luidt in woorden: "Een suppressorvariabele is een variabele die een latent aanwezige correlatie tussen twee andere variabelen onderdrukt." Formeel luidt de definitie als volgt: " $V_s$  is een suppressorvariabele

voor  $V_k$  en  $V_l$  indien geldt dat  $r_{kl} = 0$  en  $r_{kl,s} \neq 0$ ." Als illustratie geeft Verschuren het model dat in Figuur 1 is weergegeven; voor de bijbehorende correlatiematrix en parameterwaarden zie het bovenste paneel van Tabel 2.

Tabel 2

Padanalyses in de 3-variabelen-situaties van Figuur 1 (voorbeeld van Verschuren, 1991) en Figuur 2 (voorbeeld van Nie et al., 1975)

| 1   | 2                          | 3        | 4                            | 5                            | 6              | 7              | 8            |
|---|----------------------------|----------|------------------------------|------------------------------|----------------|----------------|--------------|
| $r_{31}$<br>( $r^2_{31}$ )                                      | $r_{32}$<br>( $r^2_{32}$ ) | $r_{12}$ | $b_{31,2} =$<br>$\beta_{31}$ | $b_{32,1} =$<br>$\beta_{32}$ | $r^2_{3(1,2)}$ | $r^2_{3(2,1)}$ | $R^2_{3,12}$ |
| Voorbeeld van Verschuren  |                            |          |                              |                              |                |                |              |
| .00<br>(.00)  | .42<br>(.176)              | -.29     | .133                         | .459                         | .016           | .193           | .193         |
| Voorbeeld van Nie et al.  |                            |          |                              |                              |                |                |              |
| $V3 =$ Aankoop 2e auto, $V1 =$ Behoefte 2e auto, $V2 =$ Inkomen |                            |          |                              |                              |                |                |              |
| .08<br>(.006)   | .55<br>(.302)              | -.32     | .285                         | .641                         | .073           | .369           | .376         |

Deze definitie is curieus. Het gaat om het onderdrukken van een correlatie tussen twee andere variabelen. Blijkbaar is de richting van de causaliteit en het onderscheid in afhankelijke en onafhankelijke variabelen irrelevant. De definitie is in strijd met gangbare noties over het suppressorfenomeen indien  $V_3$  zelf de afhankelijke variabele is. In de illustratie treedt dit probleem niet op; daarin wordt  $V2$  als suppressorvariabele gedefinieerd en  $V3$  als afhankelijke variabele. Merk op dat we ook volgens de definities van Conger en Velicer in dit voorbeeld te maken hebben met een suppressorsituatie. De definitie van Verschuren is echter niet consistent met die van Conger en Velicer: volgens laatstgenoemden is ook  $V1$  suppressorvariabele, volgens Verschuren niet.  $V1$  moet in het voorbeeld als de inhoudelijk belangrijke verklarende variabele worden gezien. De definitie van Verschuren botst dan met de klassieke suppressorsituatie. Er is een opvallende overeenkomst tussen het voorbeeld van Verschuren en het voorbeeld van Nie et al. Het ziet er naar uit dat Verschuren het suppressorfenomeen heeft verward met het fenomeen *maskerende variabele*.



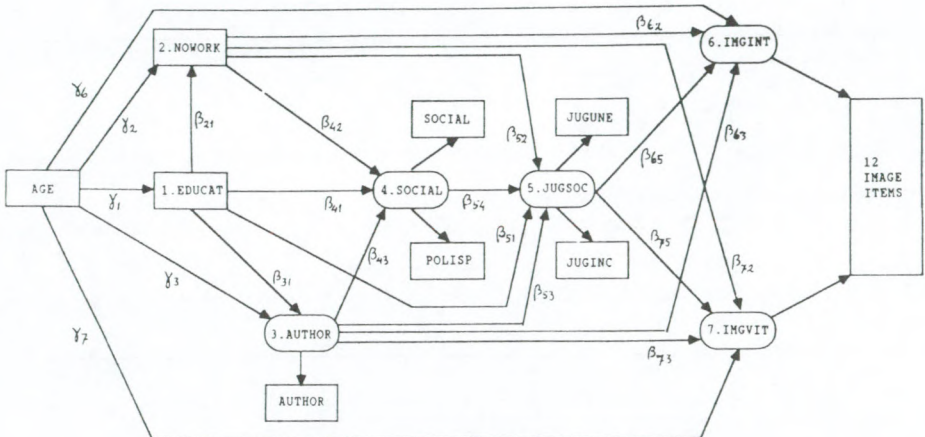
### 3. Voorbeelden

#### Voorbeeld 1: *Opinie over werklozen*

Het eerste voorbeeld is ontleend aan onderzoek naar de publieke opinie over niet-werkenden dat in de jaren zeventig en tachtig door de Maassen en De Goede is uitgevoerd (De Goede & Maassen, 1979; Maassen & De Goede, 1988 en 1991). In dit onderzoek is een model voor de factoren die opinies over niet-werkenden beïnvloeden getoetst (Maassen & De Goede, 1989; Maassen, 1997). Het volledige model is weergegeven in Figuur 3.

Figuur 3

Padmodel bij Voorbeeld 1: *Opinie over werklozen*



De afhankelijke variabelen in dit model zijn twee latente concepten: het beeld dat respondenten hebben van de vitaliteit van werklozen (IMGVIT) en het beeld dat respondenten hebben van de integriteit van werklozen (IMGINT). Beide concepten zijn geïndiceerd door 12 items; 10 daarvan hebben een substantiële lading op eerstgenoemd concept en de overige twee op het tweede concept. In het model is gehypothetiseerd dat deze twee beeldconcepten worden verklaard door drie geobserveerde variabelen (waarvan wordt verondersteld dat ze zonder meetfouten zijn gemeten), en drie latente concepten. Het eerste drietal wordt gevormd door: leeftijd (AGE), opleidingsniveau (EDUCAT) en de omstandigheid of respondent zelf een uitkering tengevolge van werkloosheid of arbeidsongeschiktheid geniet (NOWORK). De drie latente concepten zijn: autoritarisme (AUTHOR), gemeten door een somscore van 7 items uit Adorno's F-schaal; politieke voorkeur (SOCIAL), geïndiceerd door twee items; oordeel over

de kwaliteit van de sociale voorzieningen voor niet-werkenden (*JUGSOC*), eveneens geïndiceerd door twee items. Voor meer details over de opbouw van het model zie Maassen & De Goede, (1988); Maassen (1991, 1996, 1997).

Voor dit voorbeeld is onze aandacht gericht op de onderlinge relatie van *AGE*, *AUTHOR* en *IMGVIT*. De hypothesen die daaraan ten grondslag liggen, zijn de volgende:

1) Naarmate men ouder is, is men meer van mening dat burgers dienen te voldoen aan traditionele normen, zeker aan de norm dat men door werken in het eigen onderhoud voorziet. Naarmate men ouder is, is er een negatievere opinie over werklozen te verwachten.

2) Naarmate men een meer autoritaire instelling heeft, is men meer van mening dat burgers dienen te voldoen aan traditionele normen, zeker aan de norm dat men door werken in het eigen onderhoud voorziet. Er is een negatieve samenhang tussen autoritarisme en opinie over werklozen te verwachten.

3) Naarmate men ouder is, is men meer van mening dat burgers dienen te voldoen aan traditionele normen. Er is een positieve samenhang tussen leeftijd en autoritarisme te verwachten.

Dit voorbeeld kan als representatief beschouwd worden voor een uitgebreidere klasse van studies naar opinies of attitudes, want de meeste onderzoekers zullen wel geneigd zijn in padmodellen voor opinieonderzoek *leeftijd* en *autoritarisme* of verwante variabelen (zoals *rigiditeit* of *conservatisme*) op te nemen.

Ter vergelijking presenteren we ook de resultaten van een sterk vereenvoudigde analyse. Hierbij is het padmodel beperkt tot de variabelen *AGE*, *AUTHOR* en *IMGVIT*, waarbij wordt verondersteld dat behalve *AGE* ook *AUTHOR* en *IMGVIT* zonder meetfouten zijn gemeten, respectievelijk door de somscore van 7 items uit Adorno's F-schaal en de somscore van de 10 items die *IMGVIT* indiceren. Dit model wordt gerepresenteerd door Figuur 1, waarin  $V_1 = AGE$ ,  $V_2 = AUTHOR$  en  $V_3 = IMGVIT$ .

In het linkerpaneel van Tabel 3 zijn de correlaties (linksonder de hoofddiagonaal) van deze drie variabelen te vinden, alsmede de waarden die voor de padcoëfficiënten (rechtsboven de hoofddiagonaal) volgens een klassieke padanalyse worden geschat. In het rechterpaneel van Tabel 3 zijn de waarden voor de correlaties en de padcoëfficiënten van de drie variabelen te vinden zoals deze volgens een LISREL-analyse worden geschat. De tabel laat zien dat in beide gevallen de correlaties tussen de drie concepten conform de hypothesen zijn, dat wil zeggen de correlatie tussen *AGE* en *IMGVIT* is negatief. De uitkomsten voor de padcoëfficiënten tussen beide variabelen zijn echter verrassend. In het eenvoudige model is de waarde ongeveer 0 (en niet significant:  $t = .527$ ,  $p = .598$ ) en wordt de correlatie tussen *AGE* en *IMGVIT* geheel verklaard door de mediërende variabele *AUTHOR*. De LISREL-analyse levert voor het volledige model:  $\chi^2 = 238.6$  bij  $df = 141$  ( $p < .001$ ), hetgeen we beschouwen als een bevredigende fit, de steekproefomvang ( $N = 863$ ), het quotient van  $\chi$  en  $df (= 1.7)$ , de gefitte



residuen en de Q-plot in aanmerking genomen. In deze analyse krijgt de padcoëfficiënt tussen *AGE* en *IMGVIT* zelfs een positieve waarde. (De directe pijlen tussen *AGE* enerzijds en *IMGVIT* en *IMGINT* anderzijds zijn belangrijk voor de fit. Weglaten levert een significante verslechtering van de fit op:  $\chi^2 = 24.7$ , bij  $df = 2$ .)

Tabel 3

Correlaties (linksonder hoofddiagonaal) en padcoëfficiënten (rechtsboven hoofddiagonaal) in een onderzoek naar de opinie over werklozen (voorbeeld 1).

|                 | model met manifeste variabelen<br>(Figuur 1) |          |          | model met latente variabelen<br>(Figuur 3) |          |          |
|-----------------|--|----------|----------|--|----------|----------|
|                 | 1.AGE  | 2.AUTHOR | 3.IMGVIT | 0.Age                                      | 3.Author | 6.Imgvit |
| AGE             | -  | .317     | .017     | Age  | -        | .322     |
| AUTHOR          | .317   | -        | -.426    | Author                                     | .441     | -        |
| IMGVIT          | -.118  | -.421    | -        | Imgvit                                     | -.114    | -.601    |
| $r^2_{31}$      | .014   | .177     |          | $r^2_{61}$                                 | .013     | .361     |
| $r^2_{3(1..j)}$ | .000   | .164     |          | $r^2_{6(i..jkl)}$                          | .025     | .336     |
| $R^2_{3..12}$   | .178   |          |          | $R^2_{6..0235}$                            | .467     |          |

Het patroon van de padcoëfficiënten komt overeen met de in hoofdstuk 2 besproken situatie *a* (hetgeen blijkt na omschaling van *IMGVIT*). Er is hier dus sprake van negatieve suppressie volgens de definitie van Conger, waarbij *AGE* de rol van suppressor en *AUTHOR* de rol van *suppressed variable* spelen. Volgens de definitie van Velicer is *AGE* suppressor- en *suppressed variable* en is *AUTHOR* geen van beide.

*Conclusies en interpretatie.* De analyse met alleen manifeste variabelen en de analyse met latente concepten leiden dus tot verschillende resultaten: in de eerste analyse is niet of nauwelijks sprake van een suppressoreffect en in de tweede analyse wel. Dit verschil is op grond van de ongelijkheden (4a) en (5a) te verklaren: in de tweede analyse wordt gecorrigeerd voor meetfouten in de *suppressed variable* *AUTHOR*. We hechten uiteraard het meeste belang aan de resultaten van de analyse met latente concepten en we zullen ons alleen inspannen om deze resultaten te interpreteren. *IMGVIT* wordt dus door een lineaire combinatie van *AUTHOR* en *AGE* beter verklaard dan door *AUTHOR* alleen. In deze lineaire combinatie hebben de padcoëfficiënten van *AGE* en *AUTHOR* tegengesteld teken, terwijl hun validiteitscoëfficiënten hetzelfde teken hebben. Het teken van de coëfficiënt van *AGE* klappt om in combinatie met de verklarende variabele *AUTHOR*. Het lijkt niet zinvol de padcoëfficiënt van *AGE* te interpreteren als 'direct effect' in strikte zin. Ook in de interpretatie moeten *AGE* en *AUTHOR*



worden gecombineerd. Voor de bevinding dat de predictieve validiteit van *AUTHOR* door opname van *AGE* wordt verhoogd is een inhoudelijke verklaring te geven: Indien we de connectie tussen *AUTHOR* en *IMGVIT* beter willen begrijpen, moeten we het effect uitschakelen, dat respondenten een hogere score voor autoritarisme kunnen hebben alleen vanwege hun leeftijd.

*Voorbeeld 2: Burnout onder huisartsen*

Het volgende voorbeeld is een zogeheten 2W2V-situatie (twee herhaalde metingen, twee variabelen) en is ontleend aan een onderzoek naar burnout-verschijnselen onder huisartsen (Bakker, Schaufeli, Sixma, Bosveld & Van Dierendonck, 1999). Op twee tijdstippen ( $T_1 = 1991$  en  $T_2 = 1996$ ) zijn de mate van arbeidssatisfactie en mate van burnout van 270 huisartsen vastgesteld. De hypothesen die aan de te toetsen padmodellen ten grondslag liggen, luiden: "Het ontbreken van arbeidssatisfactie leidt tot burnout op korte en op langere termijn; burnout leidt op langere termijn tot vermindering van de arbeidssatisfactie."

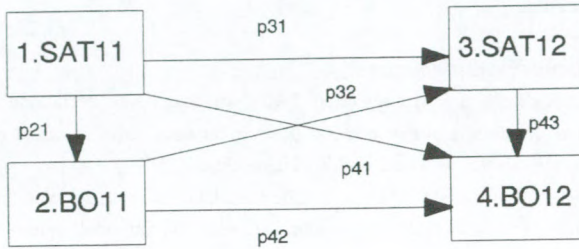
Er bestaat dus een negatief verband tussen arbeidssatisfactie en burnout. We hebben eerder gezien dat positieve predictieve validiteiten de analyse vergemakkelijken. Daarom worden de satisfactievariabelen omgeschaald, zodat ze nu gebrek aan arbeidssatisfactie indiceren. Ter illustratie zijn de gegevens op verschillende manieren geanalyseerd. In de meest geëlaboreerde variant worden beide concepten beschouwd als latente variabelen, die elk met behulp van manifeste variabelen worden gemeten. Als latent concept zullen we (gebrek aan) arbeidssatisfactie aanduiden met  $Sat_i$ , gemeten door de indicatoren  $SAT1i$  en  $SAT2i$  (de tijdstippen worden geïndiceerd door  $i = 1,2$ ). Burnout wordt als latent concept aangeduid met  $Bo_i$ , gemeten door de indicatoren  $BO1i$ ,  $BO2i$  en  $BO3i^1$ . De modellering van de hypothesen is zichtbaar gemaakt in Figuur 4b.

Tevens zullen we een gesimplificeerde variant presenteren, waarbij dezelfde hypothesen worden getoetst, maar op basis van een model dat alleen (een selectie uit de) geobserveerde variabelen bevat. Hierbij wordt verondersteld dat (gebrek aan) arbeidssatisfactie op de twee tijdstippen zonder meetfouten wordt gemeten door de variabelen  $SAT1i$  (dus  $SAT11$  en  $SAT12$ ); burnout door de variabelen  $BO1i$  (dus  $BO11$  en  $BO12$ ). De hypothesen zijn gemodelleerd zoals in Figuur 4a is aangegeven. De data en de resultaten zijn te vinden in de matrix in het linkerpaneel van Tabel 5. De benedenhelft van de matrix bevat de correlaties tussen de vier variabelen en de bovenhelft (gecursiveerd) bevat de padcoëfficiënten. Verder zijn in de Tabel de gekwadrateerde predictieve validiteiten, semipartiële correlaties en multiple-correlatiecoëfficiënten weergegeven.

In de meer geavanceerde variant (dus het model van Figuur 4b) zijn dezelfde modelparameters geschat met een LISREL-analyse. De uitkomsten staan vermeld in het rechterpaneel van Tabel 5. In deze analyse gaat het om correlaties en regressiecoëfficiënten tussen latente variabelen, die dus zijn ontdaan van attenuatie door meetfouten en daardoor hoger zijn dan de correlaties of de regressiecoëfficiënten in het linkerpaneel van de tabel.

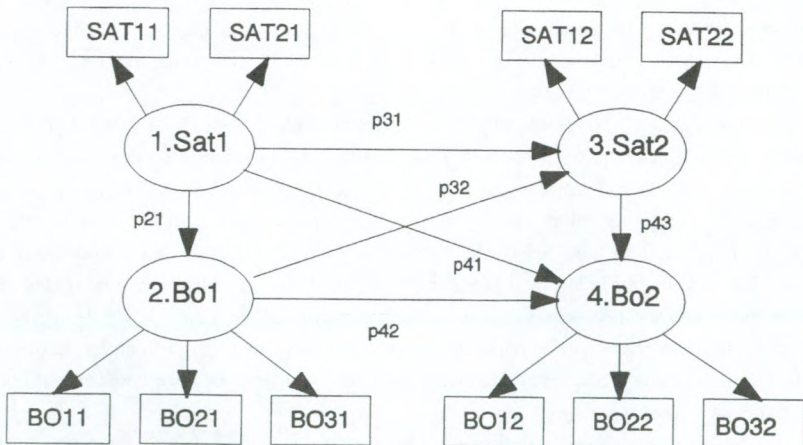
Figuur 4a

Padmodel met alleen manifeste variabelen bij *Voorbeeld 2: Burnout onder huisartsen*



Figuur 4b

Padmodel met latente en manifeste variabelen bij *Voorbeeld 2: Burnout onder huisartsen*



Tabel 4

Padanalyses in de 3-variabelen-situaties van voorbeeld 2 (Figuren 4a en 4b).

Correlaties (linksonder hoofddiagonaal) en padcoëfficiënten (rechtsboven hoofddiagonaal) in een onderzoek naar burnoutonder huisartsen.

| model met manifeste variabelen<br>(Figuur 4a)               |       |       |      | model met latente variabelen<br>(Figuur 4b)              |      |      |       |
|---|-------|-------|------|--|------|------|-------|
| $V_1 = \text{SAT11}, V_2 = \text{BO11}, V_4 = \text{BO12}$  |       |       |      | $V_1 = \text{Sat1}, V_2 = \text{Bo1}, V_4 = \text{Bo2}$  |      |      |       |
|   | SAT11 | BO11  | BO12 |  | Sat1 | Bo1  | Bo2   |
| 1. SAT11  | -     | .478  | .108 | 1. Sat1  | -    | .812 | -.227 |
| 2. BO11   | .478  | -     | .568 | 2. Bo1   | .812 | -    | .880  |
| 4. BO12   | .379  | .619  | -    | 4. Bo2   | .488 | .696 | -     |
| $r^2_{4i}$  | .144  | .384  |      | $r^2_{4i}$   | .238 | .484 |       |
| $r^2_{4(i,j)}$  | .009  | .249  |      | $r^2_{4(i,j)}$   | .017 | .264 |       |
| $R^2_{4,ij}$  | .393  |       |      | $R^2_{4,ij}$   | .501 |      |       |
| $V_1 = \text{SAT11}, V_3 = \text{SAT12}, V_4 = \text{BO12}$ |       |       |      | $V_1 = \text{Sat1}, V_3 = \text{Sat2}, V_4 = \text{Bo2}$ |      |      |       |
|   | SAT11 | SAT12 | BO12 |  | Sat1 | Sat2 | Bo2   |
| 1. SAT11  | -     | .354  | .214 | 1. Sat1  | -    | .496 | .154  |
| 3. SAT12  | .354  | -     | .469 | 3. Sat2  | .496 | -    | .672  |
| 4. BO12   | .379  | .544  | -    | 4. Bo2   | .488 | .749 | -     |
| $r^2_{4i}$  | .144  | .296  |      | $r^2_{4i}$   | .238 | .561 |       |
| $r^2_{4(i,j)}$  | .040  | .192  |      | $r^2_{4(i,j)}$   | .018 | .341 |       |
| $R^2_{4,ij}$  | .336  |       |      | $R^2_{4,ij}$   | .579 |      |       |
| $V_2 = \text{BO11}, V_3 = \text{SAT12}, V_4 = \text{BO12}$  |       |       |      | $V_2 = \text{Bo1}, V_3 = \text{Sat2}, V_4 = \text{Bo2}$  |      |      |       |
|   | BO11  | SAT12 | BO12 |  | Bo1  | Sat2 | Bo2   |
| 2. BO11   | -     | .393  | .479 | 2. Bo1   | -    | .476 | .439  |
| 3. SAT12  | .393  | -     | .356 | 3. Sat2  | .476 | -    | .540  |
| 4. BO12   | .619  | .544  | -    | 4. Bo2   | .696 | .749 | -     |
| $r^2_{4i}$  | .384  | .296  |      | $r^2_{4i}$   | .484 | .561 |       |
| $r^2_{4(i,j)}$  | .195  | .107  |      | $r^2_{4(i,j)}$   | .149 | .225 |       |
| $R^2_{4,ij}$  | .491  |       |      | $R^2_{4,ij}$   | .710 |      |       |



Tabel 5

Correlaties (linksonder hoofddiagonaal) en padcoëfficiënten (rechtsboven hoofddiagonaal) in een onderzoek naar burnout onder huisartsen (de 4-variabelen-situaties van voorbeeld 3).

|                 | model met manifeste variabelen<br>(Figuur 4a) |      |       |      | model met latente variabelen<br>(Figuur 4b) |      |      |      |       |
|-----------------|---|------|-------|------|---|------|------|------|-------|
|                 | SAT11   | BO11 | SAT12 | BO12 | Sat1  | Bo1  | Sat2 | Bo2  |       |
| SAT11           | -   | .478 | .215  | .033 | Sat1  | -    | .812 | .320 | -.419 |
| BO11            | .478  | -    | .291  | .466 | Bo1   | .812 | -    | .216 | .751  |
| SAT12           | .354  | .393 | -     | .349 | Sat2  | .496 | .476 | -    | .599  |
| BO12            | .379  | .619 | .544  | -    | Bo2   | .488 | .696 | .749 | -     |
| $r^2_{4i}$      | .144  | .384 | .296  |      | $r^2_{4i}$                                  | .238 | .484 | .561 |       |
| $r^2_{4(i,jk)}$ | .001  | .155 | .098  |      | $r^2_{4(i,jk)}$                             | .057 | .188 | .265 |       |
| $R^2_{4.123}$   | .491  |      |       |      | $R^2_{4.123}$                               | .767 |      |      |       |

Voorts zijn (weer ter vergelijking) nog analyses uitgevoerd, waarbij de twee varianten zijn gereduceerd tot 3-variabelen-situaties. Hierbij is burnout op T2 steeds als afhankelijke variabele en twee van de drie andere als verklarende variabelen genomen. Dit leidt tot in totaal zes 3-variabelen-situaties waarvan de resultaten in Tabel 4 zijn weergegeven.

Het gaat ons specifiek om de vraag hoe burnout op het tweede meetmoment wordt beïnvloed, dus we richten ons op de resultaten in Tabel 4 en in de laatste kolom van beide panelen van Tabel 5. In twee van de acht situaties lijkt het suppressorfenomeen op te treden volgens de definitie van Conger: in de 3-variabelen-situatie met latente concepten, waarbij satisfactie en burnout op T1 de onafhankelijke variabelen zijn, en in de 4-variabelen-situatie met latente concepten. In deze situaties is de padcoëfficiënt van burnout op T1 groter dan de validiteit; verder is het teken van de padcoëfficiënt van satisfactie op T1 tegengesteld aan het teken van de validiteit. Laten we in beide suppressorsituaties eens de opbouw van  $r_{24}^j$  (volgens de variabelenummering in Figuur 4) onderzoeken. In de 3-variabelen-situatie leidt dat tot:

$$r_{24}^j = p_{42} + p_{41} * r_{12} = .88 + -.23 * .81 = .70 .$$

We zien een direct effect, in getalwaarde groter dan de correlatie, en een negatief effect tengevolge van een schijnverband. Dit is de negatieve-suppressiesituatie  $a$ , besproken in hoofdstuk 2. Volgens Conger is satisfactie op T1 suppressant voor burnout op T1, maar niet omgekeerd. Volgens de definitie van Velicer blijkt er helemaal geen sprake te zijn van suppressie. In de 4-variabelen-situatie wordt de uitsplitsing:

$$\begin{aligned}
 r'_{24} &= P_{42} + P_{41} * r_{12} + P_{43} * P_{32} + P_{43} * P_{31} * r_{12} \\
 &= .75 + -.34 + .13 + .16 = .70.
 \end{aligned}$$

De correlatie is opgebouwd uit een direct effect, ook weer in getalwaarde groter dan de correlatie, vermeerderd met een negatief effect tengevolge van een schijnverband, een positief indirect effect en een positief effect tengevolge van het schijnverband. In beide gevallen wordt het negatieve effect veroorzaakt door een negatieve padcoëfficiënt van de satisfactie op T1. De correlatie tussen de satisfactie op T1 en de afhankelijke variabele is niet onaanzienlijk (.49), maar er vindt een omslag van het teken plaats als de regressiecoëfficiënt wordt berekend. De correlatie is dus niet hoog genoeg om een (verwacht) positief direct effect tussen beide variabelen waar te maken. De satisfactievariabele op T1 heeft informatie gemeen met de afhankelijke variabele, maar minder dan de informatie die deze met de burnout-variabele op T1 gemeen heeft en die irrelevant is voor de afhankelijke variabele. De satisfactie op T1 krijgt dus de rol van suppressorvariabele opgelegd; het gaat hier klaarblijkelijk om een negatieve-suppressorsituatie.

*Interpretaties.* Het hierboven besproken resultaat was onverwacht en in strijd met de vooraf geformuleerde hypothese. We moeten daaruit niet concluderen dat er sprake is van een negatief direct effect, dus dat het tegenovergestelde van de hypothese "gebrek aan arbeids-satisfactie heeft een positieve invloed op burnout" waar is. Toch is er een plausibele interpretatie in met name de 4-variabelen-situatie mogelijk. Beschouw eens de regressie-vergelijking:

$$Bo2' = .75 \cdot Bo1 + .60 \cdot Sat2 - .42 \cdot Sat1 .$$

De burnout op T1 en de satisfactie op T2 zijn goede voorspellers voor de burnout op T2, maar de voorspelling wordt geoptimaliseerd door in plaats van laatstgenoemde voorspeller een lineaire combinatie van beide satisfactievariabelen toe te voegen. Deze lineaire combinatie indiceert de *verandering in arbeidssatisfactie* tussen beide meetmomenten. Dat de gewichten van beide satisfactievariabelen in de lineaire combinatie in absolute waarde verschillend zijn, is geen beletsel om te spreken van een indicator voor verandering in satisfactie. We wijzen erop dat variabelen die een ruwe verschilscore constitueren, na standaardisering altijd verschillende gewichten zullen krijgen indien de standaarddeviaties verschillen. In de LISREL-analyse zijn de latente satisfactievariabelen gestandaardiseerd en een verschil in gewichten is dus niet in strijd met wat kon worden verwacht.

Opmerkelijk is het verschil met de interpretatie die aan de resultaten in de 4-variabelen-situatie met alleen manifeste variabelen moet worden gegeven. De regressievergelijking luidt hier:



$$BO12' = .47 \cdot BO11 + .35 \cdot SAT12 + .03 \cdot SAT11$$

In deze verklaringsvergelijking speelt de satisfactie op T1 geen rol, ook niet die van suppressorvariabele. Uit de ongelijkheid (4a) kunnen we een verklaring aflezen voor het verschil in resultaten van beide analyses. Omdat gecorrigeerd wordt voor meetfouten in de *suppressed variable(s)* bestaat in de analyse met latente concepten de grootste kans om het suppressorfenomeen aan te treffen. Uiteraard moet aan de resultaten van een analyse met latente variabelen de meeste waarde worden gehecht.

Het verband tussen verschillcores en suppressorvariabelen is al eerder besproken door Glasnapp (1984). Het probleem hoe men een correlatie tussen een verschillcore  $X$ - $Y$  en een derde variabele  $Z$  dient te berekenen was zijn uitgangspunt. Glasnapp adviseert om beide componenten van de verschillcore als aparte variabelen in een regressieanalyse op te nemen. Gegeven de aanname dat de variabelen zo zijn geschaald dat beide validiteitscoëfficiënten positief zijn, zullen, als een verschillcore inderdaad een goede voorspeller is, beide variabelen een regressiegewicht met verschillend teken krijgen. Dan is er dus sprake van een suppressor-situatie. In het algemeen zullen beide regressiegewichten ook verschillend van grootte zijn. Dat geeft dus informatie over welke variabele dominant is in de lineaire combinatie. Het toekennen van gelijke regressiegewichten (zoals in een verschillcore geschiedt) doet deze informatie verloren gaan. Bovendien is de echte verschillcore dan een minder goede voorspeller, aldus Glasnapp (1984). Hij besluit zijn artikel met: "Unfortunately, the literature indicates that the occurrence in data of suppression conditions of any magnitude is rare. This alone leads one to conclude that the search for highly meaningful change score composites as predictor or criterion variables will be unrewarding." In ons voorbeeld zien we het omgekeerde gebeuren. In principe gingen we niet uit van een verschillcore, maar bleek uit de analyse dat *a meaningful change composite* (onverwacht) de beste verklaring te geven. In het voorbeeld was dan ook sprake van een negatieve-suppressiesituatie.

Het optreden van negatieve suppressie in een 2W2V-situatie als ons voorbeeld moet overigens niet zeldzaam geacht worden. Om dat in te zien leiden we eerst een algemene formule af voor een 4-variabelen-situatie met één afhankelijke variabele  $Y$  en drie verklarende variabelen. De drie verklarende variabelen zijn door de onderzoeker in twee groepen verdeeld: variabelen  $P_1$  en  $P_2$  enerzijds en variabele  $S$  anderzijds. We veronderstellen dat alle variabelen een gemiddelde gelijk aan 0 en een variantie gelijk aan 1 hebben. Verder definiëren we  $P = \beta_1 \cdot P_1 + \beta_2 \cdot P_2$ , de lineaire combinatie van  $P_1$  en  $P_2$  die  $Y$  het beste voorspelt in de regressie van  $Y$  op  $P_1$  en  $P_2$ ; de bijbehorende multipele correlatiecoëfficiënt  $R_{y,12} = r_{Py}$ . Uit ongelijkheid (4a) volgt dat er sprake is van een negatieve-suppressiesituatie met variabele  $S$  als suppressor, indien:



$$r_{sp} > \frac{r_{sy}}{R_{y12}}, \quad (6)$$

waarin:

$$r_{sp} = \frac{\beta_1 \cdot r_{s1} + \beta_2 \cdot r_{s2}}{\sqrt{\beta_1^2 + \beta_2^2 + 2 \beta_1 \beta_2 \cdot r_{12}}}. \quad (7)$$

In het 2W2V-model als ons voorbeeld zijn  $S$  en  $P_2$  dezelfde variabele, respectievelijk gemeten op de tijdstippen T1 en T2;  $P_1$  en  $Y$  zijn een andere variabele, ook gemeten op respectievelijk T1 en T2. Volgens formule (7) is  $r_{sp}$  een gewogen som van een correlatie tussen twee variabelen die causaal met elkaar samenhangen ( $r_{s1}$ ) en een stabiliteitscoëfficiënt ( $r_{s2}$ ).  $r_{sp}$  neemt dus mogelijk een substantiële waarde aan. In het rechterlid van ongelijkheid (6) staat een quotiënt met in de noemer een multipele correlatiecoëfficiënt die zeker groter is dan de stabiliteitscoëfficiënt  $r_{1y}$  en in de teller een correlatie tussen twee verschillende variabelen die op verschillende momenten zijn gemeten. Het rechterlid zal in het algemeen beduidend lager dan 1 uitvallen. De kans dat zich een negatieve-suppressiesituatie in een 2W2V-model voordoet, moet dus reëel worden geacht.

Om ongelijkheid (6) te kunnen verifiëren in ons voorbeeld, bepalen we eerst de lineaire combinatie van Bo1 en Sat2 die Bo2 het best voorspelt:  $P = .44 \cdot \text{Bo1} + .54 \cdot \text{Sat2}$ , met bijbehorende multipele correlatiecoëfficiënt gelijk aan  $r(P, \text{Bo2}) = .710^{1/2} = .843$ . Verder berekenen we met vergelijking (7):  $r(\text{Sat1}, P) = .76$ . Dus volgens ongelijkheid (6) is in de regressie van Bo2 op Sat1 en  $P$  sprake van negatieve suppressie volgens Conger indien:  $r(\text{Sat1}, P) > [r(\text{Sat1}, \text{Bo2})]/[r(P, \text{Bo2})]$ . Dit is het geval want:  $.76 > .488/.843 = .58$ .

In dit voorbeeld was de suppressiesituatie met behulp van de definitie van Conger gemakkelijk te herkennen. Dat is in situaties met meer dan twee verklarende variabelen niet altijd het geval. Stel eens dat de stabiliteitscoëfficiënt van burnout gelijk is aan .65 (in plaats van .749). Dan vinden we de volgende regressievergelijkingen:

$$\begin{aligned} \text{Bo2}^I &= .61 \cdot \text{Bo1} + .61 \cdot \text{Sat2} - .31 \cdot \text{Sat1} \\ \text{Bo2}^{II} &= P = .38 \cdot \text{Bo1} + .57 \cdot \text{Sat2} \\ \text{Bo2}^{III} &= .82 \cdot P \\ \text{Bo2}^{IV} &= .97 \cdot P - .21 \cdot \text{Sat1} \end{aligned}$$

In de bovenste vergelijking valt op dat het teken van de padcoëfficiënt van Sat1 tegengesteld is aan het teken van  $r(\text{Sat1}, \text{Bo2})$ , maar geen van de overige verklarende variabelen speelt op zichzelf de rol van *suppressed variable*. Uit de onderste twee vergelijkingen blijkt Sat1 echter wel suppressor te zijn voor  $P$ , de lineaire combinatie van Sat2 en Bo1 die Bo2 optimaal verklaart. Alleen de tekenomkering maakt de onderzoeker hier nog attent op een mogelijke

suppressorsituatie.

*Conclusies.* Uit het voorgaande kunnen we concluderen dat in een model waarin twee variabelen  $X$  en  $Y$  elk op twee tijdstippen worden gemeten, en waarin de kruislingse verbanden  $r(X_1, Y_2)$  en  $r(X_2, Y_1)$  beduidend lager zijn dan de stabiliteitscoëfficiënten  $r(X_1, X_2)$  en  $r(Y_1, Y_2)$  en de verbanden tussen de op hetzelfde tijdstip gemeten variabelen  $r(X_1, Y_1)$  en  $r(X_2, Y_2)$  grote kans bestaat op een suppressorsituatie. Een verklarende variabele die inhoudelijk interessant lijkt, loopt dan gevaar de rol van suppressorvariabele opgelegd te krijgen. Met name in modellen met latente concepten, waarin de intercorrelaties in het algemeen op een hoger niveau liggen kan het suppressoreffect pregnant naar voren komen. Een interpretatie die gebruik maakt van verschijscores ligt dan voor de hand.

#### 4. Samenvatting en discussie

##### *Het determineren van suppressoreffecten*

Een onderzoeker die een padmodel opstelt, zal in het algemeen alleen variabelen opnemen die een inhoudelijke, verklarende betekenis hebben. Zelden zal hij een variabele uitkiezen vanwege de suppressoreigenschap. Indien het suppressorfenomeen zich in een padanalyse voordoet, is dat dus meestal onverwacht. Een variabele waarvan de onderzoeker meende dat deze een verklarende betekenis zou hebben, blijkt plotseling slechts de rol van suppressorvariabele toebedeeld te krijgen. Uit bovenstaande voorbeelden mag blijken dat zulke situaties bepaald niet zeldzaam zijn.

Wanneer het suppressorfenomeen optreedt wordt de onderzoeker in zoverre gewaarschuwd dat bij één of meer variabelen wonderlijke uitkomsten worden gevonden: een regressiecoëfficiënt krijgt opeens een veel hogere waarde of een ander teken dan verwacht. De onderzoeker die van het bestaan van het suppressorfenomeen op de hoogte is, zal het verschijnsel in het algemeen wel herkennen. Hij of zij zal dan ook wel een variabele kunnen aanwijzen die de rol van suppressor speelt, daarbij eventueel geholpen door de onderzoeksvraag. We hebben gezien dat dat algebraïsch gezien echter minder duidelijk kan liggen. Zo kwam in het voorbeeld van Nie et al. (Tabel 2 en Figuur 2), los van de onderzoeksvraag, meer dan één onafhankelijke voor de titel van suppressorvariabele in aanmerking. De definities van Conger en Velicer, die dit met opzet toelaten, zijn daar uiteraard niet vreemd aan.

De onderzoeker die het verschijnsel niet kent, zal er waarschijnlijk geen raad mee weten. De meest argeloze onderzoeker is misschien zelfs geneigd de data als waardeloos en de resultaten als niet-publiceerbaar te beschouwen. Hij of zij zou in elk geval gebaat zijn bij een waarschuwing op de computeroutput, die hem of haar bij een statistisch adviseur doet belanden. Daarvoor moet dus antwoord worden gegeven op de vraag hoe een suppressorvariabele in het kader van padmodellen in abstracto kan worden gedefinieerd. Welke van de hierboven gegeven definities zou daarbij als uitgangspunt kunnen dienen?

Nadere bestudering van de Tabellen 2 en 3 leert dat de definitie van Velicer toepasbaar blijkt in voorbeeld 1, maar niet in de twee suppressorsituaties van voorbeeld 2. In die situaties



zijn de semipartiële correlaties niet hoger dan de corresponderende predictieve validiteiten. De definitie van Conger past daarentegen op alle situaties die wij in dit artikel als suppressorsituatie hebben benoemd. Tzelgov en Stern (1978) hebben voor 3-variabelen-situaties (met een criterium  $Y$ , predictor  $P$  en suppressor  $S$ ) aangegeven wanneer wel is voldaan aan de definitie van Conger, maar niet aan die van Velicer. Tzelgov en Stern spreken van "a marginal situation in which  $S$  is a suppressor according to Conger, but fails to satisfy the intuitively reasonable demand that it increases the validity of  $P$ ." (De notaties zijn ten behoeve van de context van dit artikel aangepast.) We zullen hier niet de ongelijkheden van Tzelgov en Stern toepassen, maar andere relaties afleiden die geschikt zijn voor de 2W2V-situatie. In deze ongelijkheden zal de correlatie  $r_{sp}$  een centrale rol spelen. In een 2W2V-model weerspiegelt deze correlatie een causaal verband en neemt dus mogelijk een substantiële waarde aan. De kans op een suppressorsituatie is groter naarmate de waarde van  $r_{sp}$  groter is. We zullen nu de waarden van  $r_{sp}$  proberen te vinden waarbij wel sprake is van een suppressie-situatie volgens Conger, maar niet volgens Velicer. We beschouwen het 2W2V-model als een uitbreiding van een 3-variabelen-situatie met afhankelijke variabele  $Y$  en verklarende variabelen  $P$  en  $S$ . In deze situatie is variabele  $S$  geen suppressor volgens Velicer indien:

$$\frac{(r_{py} - r_{sp} \cdot r_{sy})^2}{1 - r_{sp}^2} < r_{py}^2.$$

Vermenigvuldigen van beide leden met  $1 - r_{sp}^2$ , uitwerken van het kwadraat en isoleren van  $r_{sp}$  leiden tot:

$$r_{sp} < \frac{2 r_{py} r_{sy}}{r_{sy}^2 + r_{py}^2}.$$

Wanneer we definiëren:  $k = r_{sy} / r_{py}$  en gebruik maken van ongelijkheid (4a), geldt dus dat er volgens Conger, maar niet volgens Velicer, sprake is van negatieve suppressie met variabele  $S$  als suppressor, als voldaan is aan de volgende ongelijkheden:

$$k < r_{sp} < \frac{2k}{k^2 + 1} \quad \text{waarin} \quad k = \frac{r_{sy}}{r_{py}}. \quad (8)$$

We mogen aannemen:  $r_{py} > r_{sy}$ , want  $r_{py}$  is een stabiliteitscoëfficiënt (van burnout in ons voorbeeld) en  $r_{sy}$  is de correlatie tussen twee verschillende variabelen die op verschillende momenten zijn gemeten. Variabele  $S$  is ook een inhoudelijk interessante variabele, dus enerzijds zal  $k$  beduidend kleiner zijn dan 1, maar anderzijds zal  $k$  niet al te klein zijn. Voor  $k = 1/2$  gelden bijvoorbeeld de volgende grenzen:  $0.50 < r_{sp} < 0.80$  en voor  $k = 3/4$ :  $0.75 < r_{sp} < 0.96$ . De kans dat  $r_{sp}$  tussen zulke grenzen ligt is reëel.



Een 2W2V-model telt behalve suppressorvariabele  $S$  nog twee andere verklarende variabelen  $P_1$  en  $P_2$  voor  $Y$  (Bo1 en Sat2 in ons voorbeeld). Definiëren we nu  $P$  als de lineaire combinatie van  $P_1$  en  $P_2$  die  $Y$  het best verklaart (met multiële correlatiecoëfficiënt  $R_{y.12} = r_{py}$ ), dan is  $S$  suppressor volgens Conger maar niet volgens Velicer, indien  $r_{sp}$  voldoet aan bovenstaande ongelijkheden, waarbij nu geldt:  $k = r_{sy}/r_{py} = r_{sy}/R_{y.12}$  en waarbij  $r_{sp}$  wordt gegeven door vergelijking (7). Omdat in het algemeen  $R_{y.12} > r_{1y}$  is, is  $k$  in het algemeen kleiner geworden dan in de 3-variabelen-situatie die als uitgangspunt diende. Hoe  $r_{sp}$  in beide situaties zich tot elkaar verhouden valt niet te zeggen, maar de kans dat nog steeds voldaan is aan de ongelijkheden (8) blijft reëel.

Ter illustratie passen we de ongelijkheden (8) toe in ons voorbeeld. Dat leidt tot (1) in de 3-variabelensituatie:  $.701 = .488/.696 < .812 < 2 \cdot .701/(\cdot 701^2 + 1) = .940$  en (2) in de 4-variabelen-situatie:  $.579 = .488/.843 < .76 < 2 \cdot .843/(\cdot 843^2 + 1) = .986$ . Wij concluderen hieruit dat Velicers definitie niet bruikbaar is voor het determineren van suppressorsituaties in 2W2V-modellen.

Het spreekt vanzelf dat wij moeite hebben met de kwalificatie van Tzelgov en Stern ("a marginal situation"). Onze voorkeur gaat uit naar de definitie van Conger. Deze definitie is aantrekkelijk, omdat daarin een vergelijking wordt getroffen tussen validiteiten en padcoëfficiënten, welke in structurele-vergelijkingen-analyse uiteraard een centrale rol spelen. Verder hebben we gezien dat de definitie van Conger wel en die van Velicer niet toepassing is op een belangrijke categorie van situaties die men gewoonlijk als suppressorsituatie zal beschouwen.

Congers definitie is ook geschikt als grondslag voor een eventuele waarschuwing in de computeroutput voor suppressorsituaties in een analyse van structurele relaties. De definitie berust op het aanwijzen van een *suppressed variable*. Dit kan in een 3-variabelen-situatie op eenvoudige wijze gebeuren. De waarschuwing kan verder worden gespecificeerd op basis van de gevallen die op pagina 8 zijn onderscheiden. In situaties met meer variabelen is dit ingewikkelder als de *suppressed variable* verscholen is in een lineaire combinatie. Is de suppressor een enkelvoudige variabele, dan is de variant op de definitie van een suppressorsituatie, gegeven door Cohen en Cohen (1975), bruikbaar: Er is sprake van suppressie indien een padcoëfficiënt van een verklarende variabele (1) groter is dan zijn validiteitscoëfficiënt, of (2) kleiner is dan nul. De onderzoeker wordt dan gewaarschuwd door de negatieve padcoëfficiënt. Een situatie waarin ook de suppressorvariabele verscholen in een lineaire combinatie, een situatie die we overigens nog nooit zijn tegengekomen, kan waarschijnlijk alleen worden ontdekt door de regressiecoëfficiënten van alle lineaire combinaties van deelverzamelingen van verklarende variabelen te vergelijken met de corresponderende multiële correlatiecoëfficiënten.

*Wat te doen in een suppressorsituatie?*

Als een onderzoeker (bijvoorbeeld met behulp van de definitie van Cohen en Cohen) heeft vastgesteld dat het suppressorfenomeen in het spel is, zal hij of zij daar meestal niet gelukkig mee zijn. Met name wanneer men in het structurele-relatiesmodel met een klassieke of negatieve suppressorsituatie wordt geconfronteerd, kan men inhoudelijk interpretatief in de problemen raken: men vindt iets dat tegengesteld is aan wat men verwachtte. De suppressorvariabele kan substantieel samenhangen met de afhankelijke variabele, maar hangt ook sterk samen met informatie van de andere verklarende variabele, die niet relevant is voor de afhankelijke variabele. Indien suppressorvariabele en verklarende variabele sterk verwant zijn (bijvoorbeeld *conservatisme* en *autoritarisme*) dan kan men aansluitend op het beginsel van *parsimony* één van beide variabelen laten vallen. Gaat het niet om verwante variabelen, dan zal het vaak op theoretische gronden niet mogelijk zijn het probleem eenvoudigweg op te lossen door variabelen weg te laten, zoals het voorbeeld van het verband tussen arbeids-satisfactie en burnout laat zien. De analyse zal dan uitwijzen dat de hypothesen en het hierop gebaseerde theoretische model waarvan men uitgaat, partieel niet kloppen, c. q. moeten worden verworpen. De onderzoeker zal geneigd zijn in aansluiting op het initiële theoretisch model de uitkomsten te presenteren van een aangepast model dat beter bij de data past. In het aangepaste model zullen de suppressorvariabele en de afhankelijke variabele door een direct pad zijn verbonden, dat is geassocieerd met een regressiecoëfficiënt die groter is of een ander teken heeft dan de onderzoeker verwachtte. Het is duidelijk dat men dat vergezeld moet laten gaan van de verklaring dat het theoretische model door het optreden van het suppressorfenomeen niet fit.

Verder is in bovenstaande voorbeelden al aangestipt dat de interpretatie van de effecten van een onafhankelijke variabele op een afhankelijke variabele in een padmodel (Pedhazur, 1982) nadere beschouwing vereist, indien er een suppressorvariabele in het spel is. Indien men een variabele als suppressor heeft gediagnostiseerd en tussen deze variabele en de afhankelijke variabele een padcoëfficiënt vindt met een teken tegengesteld aan wat men verwachtte, moet men niet concluderen dat het directe effect dan ook een ander teken heeft dan men verwachtte. Een variabele treedt alleen op als suppressor in combinatie met één of meer andere verklarende variabelen. Ook bij de interpretatie zal men de suppressor en deze andere verklarende variabelen moeten combineren en de gevonden lineaire combinatie op een zinvolle wijze zien te interpreteren. In het geval dat een suppressor en een andere verklarende variabele dezelfde variabele zijn, maar op verschillende tijdstippen gemeten, kan een interpretatie in termen van verandering zinvol zijn.



### Noot

1. Terwille van de leesbaarheid is, vergeleken met het geciteerde artikel, de naamgeving van de variabelen gestroomlijnd. De drie burnout-indicatoren betreffen de drie subschalen van de Maslach Burnout Inventory (MBI; Maslach & Jackson, 1986): *Emotionele uitputting* (gemeten met 8 items), *depersonalisatie* (5 items), en *verminderde persoonlijke bekwaamheid* (7 items). Zie Bakker et al. (1999) voor een uitgebreide beschrijving van deze schalen. De indicatoren van (gebrek aan) arbeidssatisfactie zijn items van een zelfgeconstrueerde schaal. Een voorbeeld-item is: "Ik heb echt plezier in mijn werk".

### Literatuur

- Bakker, A.B., Schaufeli, W.B., Sixma, H., Bosveld, W., & Van Dierendonck, D. (1999). Patient demands, lack of reciprocity, and burnout: A five-year longitudinal study among general practitioners. *Journal of Organizational Behavior* (in druk).
- Cohen, J. & Cohen, P. (1975). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Ass.
- Conger, A.J. (1974). A revised definition for suppressor variables: a guide to their identification and interpretation. *Educational and Psychological Measurement*, 34, 35-46.
- Darlington, R.B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161-182.
- De Goede, M.P.M. & Maassen, G.H. (1979). *De publieke opinie over niet-werken. Analyse van een onderzoek naar de beeldvorming over werklozen en arbeidsongeschikten*, 126 pgs. Lisse: Swets en Zeitlinger.
- Glasnapp, D.R. (1984). Change scores and regression suppressor conditions. *Educational and Psychological Measurement*, 44, 851-867.
- Horst, P. (1941). The prediction of personnel adjustment. *Social Science Research Council Bulletin*, 48.
- Maassen, G.H. (1996). Handling of missing values in path models for opinions or attitudes. *European Journal of Social Psychology*, 26, 1-13.
- Maassen, G.H. (1997). An application of structural modeling in public opinion research: conceptualising public and opinions. *International Journal of Public Opinion Research*, 9, 145-169.
- Maassen, G.H. & De Goede, M.P.M. (1988). *Publieke opinie over werklozen en arbeidsongeschikten. Een trendstudie*, 289 pgs. Utrecht: Rijksuniversiteit Utrecht (dissertatie).
- Maassen, G.H. & De Goede, M.P.M. (1989). Public opinion about unemployed people in the period 1975-1985; the case of the Netherlands. *The Netherlands' Journal of Social Sciences*, 25, 97-113.



- Maassen, G.H. & De Goede, M.P.M. (1991). Changes in public opinion on the unemployed; the case of the Netherlands. *International Journal of Public Opinion Research*, 3, 182-194.
- Maslach, C., & Jackson, S. (1986). *MBI: Maslach Burnout Inventory, manual research edition*. Palo Alto, CA: University of California.
- McNemar, Q. (1945). The mode of operation of suppressant variables. *American Journal of Psychology*, 58, 554-555.
- McNemar, Q. (1969). *Psychological Statistics*. London: John Wiley.
- Nie, H.N., Hadlai Hull, C., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. (1975). *SPSS: Statistical Package for the Social Sciences* (2nd ed.). New York: McGraw-Hill.
- Pedhazur, E.J. (1982). *Multiple regression in behavioral research; explanation and prediction*. New York: Holt, Rinehart & Winston.
- Smith, R.L., Ager, jr., J.W., & Williams, D.L. (1992). Suppressor variables in multiple regression/ correlation. *Educational and Psychological Measurement*, 52, 17-29.
- Tzelgov, J. & Henik, A. (1991). Suppression Situations in Psychological Research: Definitions, Implications, and Applications. *Psychological Bulletin*, 109, 524-536.
- Tzelgov, J. & Stern, I. (1978). Relationships between variables in three variable linear regression and the concept of suppressor. *Educational and Psychological Measurement*, 38, 325-335.
- Velicer, W.F. (1978). Suppressor variables and the semipartial correlation coefficient. *Educational and Psychological Measurement*, 38, 953-958.
- Verschuren, P.J.M. (1991). *Structurele modellen tussen theorie en praktijk*. Utrecht: Het Spectrum.
- Wiggins, J.S. (1973). *Personality and Prediction: Principles of Personality Assessment*. Reading MA: Addison-Wesley Publishing Company.

Ontvangen: 1 april 1999

Geaccepteerd: 25 april 2000

