

Methodologische aspecten van het IQ-debat

J. de Leeuw*

1. Inleiding

We gaan in dit artikel niet in op de vraag, of intelligentie en IQ iets met elkaar te maken hebben of niet, niet op de vraag naar de kwaliteit van de in het IQ-debat gebruikte gegevens, en niet op de maatschappelijke implicaties van de diverse mogelijke conclusies. We doen dit met bloedend hart, omdat deze vragen zowel fundamenteeler als interessanter zijn dan de zuiver methodologische problemen die we de revue laten passeren. Niettemin is het noodzakelijk om de aandacht te vestigen op de in het IQ-debat gebruikte methodologie, omdat een groot deel van de conclusies respectabiliteit ontleent aan de gebruikte methoden. Dit is volgens ons ten onrechte. De psychometrische genetica maakt gebruik van eenvoudige statistische technieken met een voornamelijk descriptief karakter, op basis waarvan geen causale conclusies mogelijk zijn. Deze stelling is niet nieuw. In vele varianten is zij al verdedigd door Hogben (1939), Loevinger (1940), Anastasi en Foley (1948), Anastasi (1958), Cavalli-Sforza en Feldman (1973), Kamin (1974), Layzer (1974), Lewontin (1975), Kempthorne (1978), Vroon (1980), en vele anderen. Aan de andere kant staan hier de claims tegenover van bijvoorbeeld Burt en Howard (1956), Jensen (1969), Jinks en Fulker (1970), Herrnstein (1973), Eaves (1977), Eaves e.a. (1977), Eysenck (1979), en alweer vele anderen. Het is weinig zinvol om het gehele begrippenapparaat van de psychometrische genetica aan te vallen; het lijkt effectiever om dezelfde theoretische uitgangspunten te kiezen en dezelfde methoden te hanteren, om dan vervolgens de beperkingen daarvan te laten zien. In termen van Overton (1973) is 'within-paradigm' kritiek in deze context minstens zo nuttig als 'between-paradigm' kritiek. Uitvoeriger voorbeelden van dit soort kritiek hebben we eerder gepubliceerd in De Leeuw (1978), De Leeuw en Stoop (1980), Jaspars en De Leeuw (1980), en De Leeuw (1981). Dit artikel is voor een groot deel gebaseerd op aan deze eerdere publikaties ontleende argumenten.

2. Korte geschiedenis van de psychometrische genetica

Psychometrische genetica is al heel oud. De eerste uitvoerige publikatie is die van Galton (1869), zijn methoden werden verbeterd door Pearson (1904, vergelijk ook

* Rijksuniversiteit Leiden, Faculteit der Sociale Wetenschappen, Afdeling Datatheorie, Breestraat 70, 2311 CS Leiden.

Elderton, 1923). De uitgangspunten van Galton en Pearson waren descriptief, en niet gebonden aan de een of andere biologische theorie over erfelijkheid. Met de herontdekking van Mendel rond 1900 ontstond er een controverse tussen biometrische genetici en Mendeliaanse genetici, die pas uit de wereld geholpen werd door een befaamd artikel van Fisher (1918). Daarin werd aangetoond dat continue variatie verklaard kan worden door aannemen van een groot aantal genen die de desbetreffende eigenschap (bijvoorbeeld IQ) beïnvloeden, en dat de door de biometrici geobserveerde correlaties in overeenstemming gebracht konden worden met de Mendeliaanse theorie. Het artikel van Fisher is op een aantal kritieke punten onbegrijpelijk, maar heeft de ontwikkeling van vooral de psychometrische genetica ingrijpend beïnvloed. In de eerste plaats door het werk van Burt en zijn medewerkers (bijvoorbeeld Burt en Howard, 1956). En in de tweede plaats via de biometrische genetica, met name de zogenaamde 'Birmingham School' (Mather en Jinks, 1971). Jensen, een leerling van Burt, bracht voor het eerst na de oorlog het IQ-debat weer op gang met zijn controversiële artikel (Jensen, 1969). Jinks, van de Birmingham School, publiceerde een jaar later tezamen met de psycholoog Fulker een invloedrijk artikel (Jinks en Fulker, 1970). Eysenck, een andere leerling van Burt, mengde zich in het debat met een paar veelgelezen boeken (Eysenck, 1971, 1973). De laatste resultaten van de samenwerking tussen laten we zeggen Burt en Birmingham zijn de belangrijke artikelen van Eaves (1977), Eaves e.a. (1977), en het boek van Eysenck (met bijdragen van Fulker) (1979).

In de inleiding hebben we al aangestipt dat zowel de oude als de nieuwe methoden van de psychometrische genetica dikwijls bekritiseerd zijn. Omdat we ons in dit artikel concentreren op de meest moderne methoden, zijn een aantal van de oude tegenwerpingen niet meer relevant. Een groot aantal bezwaren blijft evenwel overeind. Het duidelijkst worden ze uiteengezet in Lewontin (1975), Kempthorne (1978), Jaspars en De Leeuw (1980). We zullen eerst, onder meer aan de hand van een eenvoudig voorbeeld, de uitgangspunten en methoden van de moderne psychometrische genetica bespreken, en vervolgens een aantal kritische kanttekeningen plaatsen.

3. Uitgangspunten van de biometrische genetica

Het belangrijkste uitgangspunt van de biometrische genetica is de vergelijking $P_X = f(G_X, E_X)$, in woorden: het *fenotype* van organisme X is een functie van zijn *genotype* en van zijn *omgeving*. Het fenotype is datgene wat we waarnemen, in de psychometrische genetica dus een of andere vorm van gedrag; het genotype is de erfelijke aanleg, gecodeerd in de chromosomen; en de omgeving is de totaliteit van alle stimuli die ingewerkt hebben op het individu. Op zichzelf is deze formule vaag, misschien zelfs triviaal, maar gebleken is dat met deze formule als uitgangspunt een efficiënte organisatie van research mogelijk is. We kunnen het effect van omgeving op fenotype vaststellen door organismen met een zelfde genotype in verschillende omgevingen te brengen en het effect van genotype vaststellen door personen met een verschillend genotype in dezelfde omgeving te brengen. Dit research-programma gaat

volgens het paradigma van de fysische wetenschappen; inferenties omtrent causale of functionele verbanden volgen uit experimentele manipulatie van condities.

De biometrische genetica, zoals uiteengezet in het boek van Mather en Jinks (1971), heeft verder als belangrijkste uitgangspunt de zogenaamde 'scaling test'. In feite betekent dit, dat een bepaalde nadere specificatie van het model $P_X = f(G_X, E_X)$ getoetst wordt met de methoden van de klassieke inferentiële statistiek, in het bijzonder met aan de variantie-analyse ontleende technieken.

Het is belangrijk in te zien, hoe we van de 'tautologie' $P_X = f(G_X, E_X)$ overgaan op variantie-analytische modellen en technieken. In de biometrisch-genetische literatuur wordt aan deze fundamentele vertaling nauwelijks aandacht besteed. Een eerste stap is, dat men overgaat van organismen of individuen op populaties waarover waarschijnlijkheidsmaten gedefinieerd zijn. We pretenderen niet meer, dat we kunnen voorspellen wat er met ieder afzonderlijk individu gebeurt; we doen alleen nog uitspraken over de structuur van stochastische grootheden in een aantal populaties.

Fenotype in populatie k is nu een stochastische grootheid \underline{P}_k , hetzelfde geldt voor genotype \underline{G}_k en omgeving \underline{E}_k (stochastische grootheden worden aangeduid met een streep onder het desbetreffende symbool). Een tweede vereenvoudigende stap is, dat we aannemen dat \underline{P}_k , \underline{G}_k , en \underline{E}_k één-dimensionale stochastische grootheden zijn, dat wil zeggen *schalen*. We schrijven ons model nu in de vorm

$$\underline{P}_k = \underline{G}_k + \underline{E}_k + (\underline{GE})_k + \underline{R}_k.$$

Hierin is $(\underline{GE})_k$ de *interactie* tussen genotype en omgeving, ook al gemeten op een enkelvoudige schaal. Meetfouten en aspecten van omgeving, genotype, en interactie die niet in de schalen te vangen zijn, nemen we samen tot het residu \underline{R}_k , ook een schaal. Zolang er geen verdere aannamen gemaakt worden, is het model nog steeds tautologisch, omdat \underline{R}_k gedefinieerd wordt als $\underline{P}_k - f(\underline{G}_k, \underline{E}_k)$.

Met de verdere specificaties kunnen we twee kanten uit. Een eerste manier concentreert zich op de *verwachte waarden*. Stel $P_k, G_k, E_k, (GE)_k$, en R_k zijn de verwachte waarden van de overeenkomstige variabelen in populatie k . We kunnen nu schrijven

$$\underline{P}_k = \underline{G}_k + \underline{E}_k + (\underline{GE})_k + \underline{S}_k,$$

waarbij

$$\underline{S}_k = (\underline{G}_k - G_k) + (\underline{E}_k - E_k) + ((\underline{GE})_k - (GE)_k) + \underline{R}_k$$

als nieuw residu opgevat wordt. Dit is nog steeds een eenvoudige herschrijving van het algemene model; we maken het model nu restrictief door de additionele aannamen $AVE(\underline{S}_k) = 0$ en $VAR(\underline{S}_k) = \sigma^2$, waarbij AVE en VAR gebruikt worden voor verwachte waarde en variantie. Merk op dat de aanname $AVE(\underline{S}_k) = 0$ equivalent is aan $R_k = AVE(\underline{R}_k) = 0$, met woorden dus aan de aanname dat meetfouten en niet-schaalbare componenten gemiddeld niet bijdragen aan fenotype. De aanname dat \underline{S}_k in alle populaties dezelfde variantie heeft, is buitengewoon sterk, en zal op de een of andere manier theoretisch aannemelijk gemaakt moeten worden. Deze eerste specificatie van het algemene model zegt iets over de structuur van de gemiddelden, alle variatie binnen populaties is meetfout. We noemen dit wel het 'fixed

effects' model van de variantie-analyse, waarbij we moeten aantekenen dat de naam 'variantie-analyse' hierbij nogal misleidend is.

Bij een tweede specificatie van het algemene model nemen we aan, dat $P_k = G_k = E_k = (GE)_k = R_k = 0$, dat wil zeggen alle in het model voorkomende componenten hebben verwachte waarde gelijk aan nul. Alle gemiddelden worden hierdoor direct aan het begin al uit de analyse verwijderd, eventuele structuur kan alleen nog uit aannamen over de hogere orde momenten afgeleid worden. We noemen dit het 'random effects' model; in het algemeen zullen we hierbij aannamen maken over varianties en covarianties van de variabelen. Specificatie van hogere dan tweede orde momenten is ongebruikelijk, omdat statistische procedures voor het random effects model gebaseerd zijn op de additionele assumptie dat de verdeling van de variabelen normaal is.

In het model met *vaste* effecten is het vervolgens gebruikelijk om de populaties te structureren door gebruik van twee indices, een index i voor genotypen of 'rijen' en een index j voor omgevingen of 'kolommen'. We krijgen hierdoor

$$\underline{P}_{ij} = G_{ij} + E_{ij} + (GE)_{ij} + \underline{S}_{ij},$$

waarbij $i = 1, \dots, n$ en $j = 1, \dots, m$. De additionele aannamen zijn $G_{ij} = G_i$ en $E_{ij} = E_j$. Dat wil dus zeggen dat het gemiddelde genotype constant is binnen rijen, en de gemiddelde omgeving is constant binnen kolommen. Zoals we gezien hebben, is bij het vaste effecten model de aanname $\text{VAR}(\underline{S}_{ij}) = \sigma^2$ cruciaal. Vrijwel de enig denkbare manier om ervoor te zorgen dat deze aanname inderdaad opgaat, is de proefopzet zo te kiezen dat $\text{VAR}(\underline{G}_{ij}) = \text{VAR}(\underline{E}_{ij}) = \text{VAR}((\underline{GE})_{ij}) = 0$. We hoeven dan alleen nog maar te geloven dat $\text{VAR}(\underline{R}_{ij}) = \sigma^2$, en dit is gemakkelijker aannemelijk te maken. In cel (i, j) hebben dus alle individuen hetzelfde genotype en dezelfde omgeving, terwijl de storende variabelen allemaal dezelfde variantie hebben.

Met lagere organismen is het mogelijk door inteelt, zelfbestuiving, of direct chromosomenonderzoek individuen te vinden met een identiek genotype. We brengen deze individuen dan in een aantal verschillende omgevingen (temperaturen, bemestingen, vochtigheidsgraden) en gaan het effect op genotype na via variantie-analyse. Op dit eenvoudige design zijn vele variaties mogelijk door gebruik te maken van speciaal ontwikkelde proefopzetten of door omgeving en/of genotype multivariaat op te vatten; het belangrijkste blijft echter dat we het design pas in elkaar kunnen zetten, als zowel genotype als omgeving eenduidig gedefinieerd en operationeel vaststelbaar zijn. Hierdoor is de indeling in rijen en kolommen mogelijk, en hierdoor ook kunnen we randomisatie gebruiken door de genotypen willekeurig over omgevingen te verspreiden.

Bij hogere organismen is direct chromosomenonderzoek onmogelijk, bij mensen is ook gecontroleerde teeltkeuze niet te verwezenlijken. De enige personen waarvan we met een zeker vertrouwen kunnen zeggen dat ze een identiek genotype hebben, zijn ééneiige meerlingen, hoewel ook deze stelling ter discussie staat. Theoretisch zouden we dus n monozygotische meerlingen willekeurig over m pleeggezinnen kunnen verdelen, of over pleeggezinnen uit m sociaal economische klassen. Dit is het

psychometrisch genetische analogon van het klassieke vaste effecten design. Het enige 'experiment' dat pretendeert deze condities te benaderen, is Burt (1966); men is het er tegenwoordig over eens, dat de daar gerapporteerde resultaten frauduleus zijn. De problemen hangen natuurlijk ook samen met het feit dat psychometrische genetici 'interessante' verschijnselen als schoolprestaties, intelligentie, criminaliteit, en schizofrenie bestuderen, en geen enkelvoudig bepaalde Mendeliaanse karakteristieken die men in de medische genetica bestudeert.

Het is daarom niet verwonderlijk, dat in het boek van Mather en Jinks weinig aandacht besteed wordt aan natuurlijke populaties, maar in de psychometrie en in de latere publikaties van de Birminghamse kwantitatieve genetici is er wel degelijk naar een uitweg uit deze lastige situatie gezocht. We gaan over van het model met *vaste* effecten, waarin genotype en omgeving expliciet gedefinieerd moeten worden, naar het model met *variabele* effecten, waarin genotype en omgeving 'latente' variabelen zijn. De gemiddelden worden uit de analyse verwijderd, de populaties bestaan uit *paren* individuen die in een bepaalde relatie tot elkaar staan. Het model is $(\underline{P}_X, \underline{P}_Y) = (\underline{G}_X, \underline{G}_Y) + (\underline{E}_X, \underline{E}_Y) + ((\underline{GE})_X, (\underline{GE})_Y) + (\underline{R}_X, \underline{R}_Y)$, en voor iedere populatie maken we een aantal aannamen over de correlatiestructuur van de latente variabelen aan de rechterkant van de vergelijking. Deze aannamen hebben implicaties voor de correlatiestructuur van de manifeste variabelen $(\underline{P}_X, \underline{P}_Y)$, en deze consequenties worden getoetst door willekeurige steekproeven uit de populaties te kiezen.

De verschillen met de benadering via het vaste effecten model zijn duidelijk. Randomisatie is niet meer mogelijk, genotype en omgeving hoeven niet meer expliciet gedefinieerd te worden, en zijn daardoor noodzakelijk vage begrippen geworden, die uitsluitend gedefinieerd zijn in termen van de aannamen over correlaties die we maken. Met name zijn alle biologische en sociale aspecten van het model interpretaties van de vrije parameters geworden, interpretaties die vanzelfsprekend een zekere mate van willekeurigheid hebben en dikwijls uitermate vergezocht lijken. We lichten dit toe aan de hand van een eenvoudig, maar representatief voorbeeld.

Stel, we hebben zes verschillende populaties, bestaande uit paren individuen. De eerste populatie zijn apart opgevoede monozygotische tweelingen, de tweede populatie zijn tezamen opgevoede monozygotische tweelingen, de derde zijn tezamen opgevoede broers, de vierde apart opgevoede broers, de vijfde tezamen opgevoede niet-verwante jongens, en de zesde apart opgevoede niet-verwante jongens. Een verdere verfijning is mogelijk, als we de broers vervangen door mannelijke dizygote tweelingen, omdat we dan in alle populaties beide leden van het paar even oud kunnen maken. We kunnen ons zelfs voorstellen dat we alleen maar individuen bekijken die op dezelfde dag geboren zijn. Voor de twee individuen in een paar nemen we aan dat $\underline{P}_X = g\underline{G}_X + e\underline{E}_X + r\underline{R}_X$,

$$\underline{P}_Y = g\underline{G}_Y + e\underline{E}_Y + r\underline{R}_Y,$$

dat wil zeggen interactie ontbreekt, de parameters (g, e, r) zijn hetzelfde voor alle zes populaties. Door het gebruik van deze parameters kunnen we aannemen dat de

varianties van $\underline{G}_X, \underline{G}_Y, \underline{E}_X, \underline{E}_Y, \underline{R}_X, \underline{R}_Y$ gelijk aan één zijn, we nemen tevens aan dat al de gemiddelden gelijk zijn aan nul. Wat wel tussen populaties verschilt, zijn de correlaties tussen genotype en omgeving. We nemen aan, min of meer per definitie, dat \underline{R}_X en \underline{R}_Y ongecorreleerd zijn met elkaar en met alle andere variabelen. Dit impliceert dat

$$\text{VAR}(\underline{P}_X) = \text{VAR}(\underline{P}_Y) = g^2 + e^2 + 2ge\text{COR}(\underline{G}_X, \underline{E}_X) + r^2,$$

$$\text{COV}(\underline{P}_X, \underline{P}_Y) = g^2\text{COR}(\underline{G}_X, \underline{G}_Y) + e^2\text{COR}(\underline{E}_X, \underline{E}_Y) + ge\{\text{COR}(\underline{G}_X, \underline{E}_Y) + \text{COR}(\underline{E}_X, \underline{G}_Y)\}.$$

De symbolen $\text{VAR}(\quad)$, en $\text{COV}(\quad)$, worden gebruikt voor variantie, covariantie, en correlatie. Door de symmetrie binnen paren mogen we aannemen dat $\text{COR}(\underline{G}_X, \underline{E}_Y) = \text{COR}(\underline{E}_X, \underline{G}_Y)$. Voor iedere populatie blijven er nog steeds vier covariantie parameters over, te weten

$$\mu = \text{COR}(\underline{G}_X, \underline{G}_Y),$$

$$\rho = \text{COR}(\underline{E}_X, \underline{E}_Y),$$

$$\gamma = \text{COR}(\underline{G}_X, \underline{E}_X),$$

$$\delta = \text{COR}(\underline{G}_X, \underline{E}_Y) = \text{COR}(\underline{E}_X, \underline{G}_Y).$$

Dus, in het kort,

$$\text{VAR}(\underline{P}_X) = g^2 + e^2 + 2\gamma ge + r^2,$$

$$\text{COV}(\underline{P}_X, \underline{P}_Y) = \mu g^2 + \rho e^2 + 2\delta ge.$$

In totaal zijn er nu dus $3 + (6 \times 4) = 27$ parameters, en slechts $6 + 6 = 12$ geobserveerde varianties en covarianties. Wat staat ons te doen? We kunnen meer vergelijkingen afleiden, bijvoorbeeld voor hogere orde momenten. Omdat biometrisch genetische analyse over het algemeen gebaseerd is op de normaal-verdeling, wordt dit nooit gedaan. We kunnen meer populaties bekijken, maar dit vergroot alleen het aantal parameters. Er zit niets anders op dan meer aannamen te gaan maken. In de eerste plaats aan de genetische kant. We nemen daar aan, dat voor monozygoten geldt $\mu = 1$, en voor niet verwante personen geldt $\mu = 0$. Voor dizygoten geldt dat μ onafhankelijk is van apart of tezamen opvoeden. Meer aannamen zijn nodig, nu aan de omgevingskant. We nemen aan dat ρ voor apart opgevoede individuen gelijk is aan nul, en voor tezamen opgevoede individuen gelijk is aan één. En tenslotte nemen we aan dat in alle populaties geldt dat $\gamma = \delta = 0$. Voor alle populaties is de variantie dan gelijk aan $g^2 + e^2 + r^2$. Als we $g_A^2 = \mu g^2$ en $g_B^2 = (1 - \mu) g^2$ definiëren, dan geldt voor de covarianties de volgende tabel.

MZT	monozygoten	tezamen	$\text{COV} = g_A^2 + g_B^2 + e^2.$
MZA	monozygoten	apart	$\text{COV} = g_A^2 + g_B^2.$
DZT	dizygoten	tezamen	$\text{COV} = g_A^2 + e^2.$
DZA	dizygoten	apart	$\text{COV} = g_A^2.$
URT	niet-verwant	tezamen	$\text{COV} = e^2.$

URA	niet-verwant	apart	COV = 0.
VAR	variantie	allen	VAR = $g_A^2 + g_B^2 + e^2 + r^2$.

We hebben nu nog maar vier parameters; de twaalf varianties en covarianties zijn lineair in die parameters; we kunnen de vier parameters schatten op basis van onafhankelijke steekproeven uit de zes populaties. Gebruik van efficiënte schatters maakt het mogelijk de aanpassing van het model met behulp van een chi-kwadraat toets te schatten. Het artikel van Jinks en Fulker uit 1970 baarde onder andere opzien, door aan te tonen dat dit zeer eenvoudige model goed paste op gegevens verzameld door Burt. Hetzelfde model, zonder URA en URT en zonder MZA en DZA, wordt nog steeds volop gebruikt in tweelingen-onderzoek (bijvoorbeeld voorbeeld 1 in Eaves, 1977, of Martin, 1975). Hetzelfde model, maar dan uitsluitend bestaand uit MZA, wordt door bijvoorbeeld Jensen (1969) en Munsinger (1977) gebruikt om g^2 , de proportie variantie veroorzaakt door genotype, te schatten. En het model bestaande uit uitsluitend URT wordt gebruikt om de proportie variantie veroorzaakt door gezamenlijke omgeving te schatten in zogenaamde adoptiestudies. Het is niet al te ver gezocht dit bovenstaande model als prototype van biometrisch genetisch onderzoek, en daardoor ook als prototype van psychometrisch genetisch onderzoek, te beschouwen.

Het is overigens geenszins het geval, dat ons voorbeeld het meest gecompliceerde model uit de biometrische genetica is. In 1918 stelde Fisher al een aanzienlijk gecompliceerder model voor, met name aan de genetische kant veel meer uitgewerkt, zodat berekeningen van COV ($\underline{G}_X, \underline{G}_Y$) in veel meer typen populaties mogelijk waren (halfbroers/zusters, overgrootvaders en achterkleinkinderen, enzovoort). De moderne biometrische genetici, zoals Eaves, modelleren ook COV ($\underline{G}_X, \underline{E}_X$) en COV ($\underline{G}_X, \underline{E}_Y$). In meer gecompliceerde modellen neemt het aantal parameters natuurlijk toe, zodat we gedwongen zijn sterkere aannamen te maken op andere plaatsen, of om gegevens in meer populaties te verzamelen. Wat intelligentie betreft, is dit laatste een nogal hopeloze zaak, vanwege het feit dat IQ-tests in de tijd van overgrootvader nog niet bestonden, in de tijd van grootvader er anders uitzagen, en ten minste éénmaal per generatie opnieuw genormeerd moeten worden.

4. Enige kritische kanttekeningen

Het eerste punt van kritiek ligt voor de hand. Om het aantal parameters te reduceren hebben we een groot aantal aannamen moeten maken, die bijvoorbeeld in het geval van IQ niet aantoonbaar juist en in sommige gevallen aantoonbaar onjuist zijn. Hoe is dat te rijmen met het feit dat het model als geheel op een aantal typen gegevens een behoorlijke fit geeft? Het antwoord is zo eenvoudig, dat het verbazing wekt dat er in de geschriften van de Burt-Birmingham school uitermate weinig aandacht aan besteed wordt. Kempthorne zegt het als volgt: 'In connection with the Fisher and

Wright models, it seems worth mentioning that it is a childish error to deduce from the facts (i) theory \rightarrow "expected" results, and (ii) "expected" results \neq observed results that the theory is true. This is so often forgotten that one can only marvel. How often was it suggested in that "queen of sciences" physics that electrons and protons were *the* elementary particles' (1978, p. 18). In de discussie van Eaves (1977) zegt Curnow: 'The fact that more complex models do not improve the fit should not be used on its own to argue that the simple model is correct or even sufficiently accurate for some purpose or other' (1977, p. 350). En Goldberger zegt: 'To explain the persistent use of such assumptions, we need only recognize that without them the models would be indeterminate. If less restrictive, and hence more plausible, specifications were made, the number of unknown parameters would approach and soon exceed the number of observations. Implausible assumptions are needed to identify the parameters and produce the estimates, and thus to keep the model-fitters happy' (1979, p. 337). En Lewontin (1975) stelt: 'The parameters must be estimated from the data to which the model is to be fitted, so that, in part at least, the genetic model is ad hoc, and the number of degrees of freedom left to actually test the model is not great. Superimposed on this problem is the fact that any reasonable environmental hypothesis predicts data with the same general structure as the genetic hypothesis, i.e. that correlation decreases progressively with more distant relationship' (1975, p. 390).

Als illustratie van deze kritiek gebruiken we dezelfde zes populaties als boven, en we vatten een argument van Jaspers en De Leeuw (1980, p. 51-65) kort samen. Als we wat minder assumpties maken, bijvoorbeeld alleen maar over de relatieve grootte van de correlatieparameters in de zes populaties, dan vinden we een model dat de gegevens perfect kan beschrijven, maar waarin de parameters niet geïdentificeerd zijn. We kunnen nu alle mogelijke deelmodellen van dit volledige model bekijken, en parameters schatten die een zo goed mogelijke aanpassing geven. We vinden dan een groot aantal modellen die goed op de gegevens passen, die dikwijls tot zeer verschillende inhoudelijke conclusies kunnen leiden, en die soms tot absurde resultaten en schattingen leiden, al hebben ze nu nog zo'n goede fit. De conclusies van Jaspers en De Leeuw zijn: '(1) All tests and estimates only make sense within the particular model we have adopted. ... (2) Selection of a single submodel from the saturated model can be very misleading. ... (3) The interpretation of the model must be separated from the model itself' (1980, p. 65).

Biometrische genetici antwoorden op dit soort kritiek met een verbluffende naïviteit. Eaves: 'Even so, the level of complexity which need be assumed for a particular trait, human or otherwise, is decided neither by appeals to evolution nor by invocation of "common sense", but by the analysis of data' (1977, p. 354). En nogmaals: 'I would have no one doubt, however, that I see the statement by Professor Curnow that "This approach relies too much on the data" as a compliment rather than a criticism. In a discipline where prejudice and speculation thrive upon ignorance of the facts, I make no apology for preferring data' (1977, p. 353-354). Uit deze citaten blijkt een kloof, en als wij ons niet vergissen, is het dezelfde kloof als er inder-

tijd tussen biometrici en Mendelianen bestond. De filosofie van Pearson was empiristisch en instrumentalistisch; een theorie is een zo kort mogelijke samenvatting van een zo groot mogelijk aantal feiten; causaliteit is een verouderd begrip en moet vervangen worden door correlatie. Het Scheermes van Occam draagt ons op onze modellen zo eenvoudig mogelijk te maken. Het is deze methodologie die van de psychometrie in het algemeen een vormloze massa van miljoenen correlaties gemaakt heeft, en het lijkt of deze methodologie in Birmingham nog steeds vaste voet onder de grond heeft. De lezer van de boeken van Eysenck, en dan met name van de citaten waar zijn hoofdstukken mee beginnen, zal waarschijnlijk ook een schok van herkenning ervaren hebben. Keer op keer blijkt uit de geschiedenis van het IQ-debat, of bijvoorbeeld uit de discussies in de nauw verwante kwestie Buikhuisen, dat er over dit soort uitgangspunten geen vruchtbare discussie mogelijk is. Het moet een zware klap zijn voor toegewijde empiristen dat uit bijvoorbeeld Eysenck en Kamin (1981) nogmaals blijkt dat er zelfs geen overeenstemming te bereiken is over wat de *feiten* nu eigenlijk zijn.

Een tweede punt van kritiek op de modellen van de biometrische genetica wordt ook uitvoerig besproken in Jaspars en De Leeuw (1980). In het model van Fisher is de omgeving uitsluitend *ruis* in het genetische model, in het model van Sewall Wright wordt van het begin af aan plaats ingeruimd voor een actieve omgevingscomponent. Fisher modelleert dominantie in Mendeliaanse zin als niet-additiviteit in een kleinste kwadraten model. Dit heeft tot gevolg dat de correlatie tussen broers/zusters hoger is dan die tussen ouders en kinderen. Wright (1921, en in veel andere publikaties) veronderstelt dat de correlatie tussen de omgeving van broers/zusters groter is dan de correlatie tussen de omgeving van ouders/kinderen, met precies hetzelfde gevolg. De modellen zijn niet te onderscheiden. Opvolgers van Wright, zoals Morton (1974) en Rao e.a. (1976), krijgen dan ook onenigheid met Eaves e.a. (1977, p. 2) omdat zij geen dominantie modelleren. De discussie is natuurlijk zinloos, maar de conclusies van de twee scholen zijn nogal verschillend, in wezen omdat dezelfde parameters in de twee modellen anders geïnterpreteerd worden. Dit is ook de tendens van Jaspars en De Leeuw. Naarmate sociale wetenschappers zich bezig zullen gaan houden met het ontwerpen van modellen waarin de omgeving meer uitvoerig gemodelleerd wordt, zal g^2 , de erfelijkheidscoëfficiënt, dalen. Omdat de omgeving uit een oneindig aantal variabelen met een grote verscheidenheid van correlaties bestaat, is het eenvoudig het model te formuleren als

$$\underline{P}_X = g \underline{G}_X + \sum_{s=1}^{\infty} e_s \underline{E}_X^s + r \underline{R}_X,$$

en als de computer maar groot genoeg is, kan men de bijdrage van het genotype zo klein maken als men wil. Niet dat daarmee iets gewonnen is, het geeft alleen maar een eigenaardigheid weer van lineaire modellen met een toenemend aantal gecorrigeerde predictoren. Het is echter merkwaardig dat social wetenschappers zich toegelegd hebben op het overnemen en verdedigen van genetische modellen, die van het IQ iets permanents en biologisch verankerds maken, in plaats van de omgevings-

component van de modellen uit te breiden. Dit wordt pas de laatste tijd gedaan, nota bene voornamelijk door genetici.

Een laatste punt van kritiek is weer zo voor de hand liggend, dat we het nauwelijks durven te noemen. Correlatie, al dan niet vermomd als variantie-analyse, en al dan niet met een respectabel genetisch-statistisch sausje overgoten, heeft niets te maken met causatie. Statistische terminologie suggereert dit wel, door te praten over *verklaarde* variantie. Biometrische genetici zijn relatief voorzichtig in dit opzicht, maar psychologen als Jensen, Herrnstein, Eysenck, Munsinger menen dat IQ voor 80% genetisch *bepaald* is, en gebruiken dit gegeven om uitspraken te doen over mogelijke gevolgen van interventie. Kempthorne zegt in dit verband: 'I am not deliberately attempting to make a caricature of a portion of the literature, although the reader might suspect so. A referee has asked for actual references here. I suggest that the reader pick a work *at random* from the writers who advocate "heritability" of, say 80%' (1978, p. 12). In deze context verwijzen we ook naar De Leeuw (1978), en De Leeuw en Stoop (1980), die het verband tussen correlatie en causaliteit bespreken bij een verwant soort onderzoek.

Er wordt dikwijls gezegd, dat de genetische modellen in ieder geval precieze predicties maken, en dat de omgevingsmodellen veel vager zijn (bijvoorbeeld Urbach, 1974). Het is juist dat sommige genetische modellen precieze predicties maken, wat niet wil zeggen dat deze ook altijd uitkomen. Andere modellen hebben veel vrije parameters, en zijn niet erg precies. Weer andere modellen hebben een interpretatie die vrijwel geheel gebaseerd is op identificatie-condities. En tenslotte kunnen de precieze predicties dikwijls uit veel algemenere modellen afgeleid worden. Een laatste voorbeeld, ontleend aan De Leeuw (1981). De 'law of ancestral inheritance' van Galton, die rond de eeuwwisseling de genetica terroriseerde, en die nu in de psychometrische genetica een relatief grote rol speelt, voorspelt dat de correlatie tussen het IQ van een individu, en dat van zijn voorvader s generaties terug gelijk is aan θ^s . Dus vader-zoon correleren θ , grootvader-kleinzoon θ^2 , enzovoorts. In navolging van Galton en de meeste van zijn opvolgers volgen we uitsluitend de mannelijke lijn, maar daarlangs zijn de predicties inderdaad vrij precies. Als het al mogelijk zou zijn IQ-gegevens voor een aantal opeenvolgende generaties te verzamelen, en ze zouden overeenkomen met Galton's wet, dan zou daaruit nog niet de juistheid van die wet volgen. De Leeuw (1981) laat zien dat de reeks θ^s volgt uit de aannamen dat overerving een stationair Markov-proces is en dat de regressie van zoon op vader lineair is. Dit soort aannamen zijn echter ook waar of bij benadering waar bij legale overerving, en bij sociale mobiliteit, en bij nog een groot aantal andere processen die met erfelijkheid niets te maken hebben.

5. Conclusie

In het IQ-debat, en in de nauw verwante discussie rond de sociobiologie en de bio-sociale criminologie, worden door zowel de aanhangers van G als door de verdedi-

gers van E een groot aantal elementaire fouten gemaakt. Deze fouten liggen op statistisch, maar vooral op methodologisch terrein. Ze hebben vrijwel allemaal te maken met de *interpretatie* van resultaten uit relatief geavanceerde statistische analyses in wetenschapsgebieden met een geringe theoretische onderbouw. Het is een *sprookje* dat wetenschappelijke theorieën uit de computer rollen, zelfs bij analyse van heel grote data-bestanden en bij gebruik van gecompliceerde statistische technieken. Met name het IQ-debat heeft ertoe geleid dat over de gehele wereld velen zich verbaasd hebben over de *incompetentie* op methodologisch en wetenschapstechnisch gebied van een aantal aan de weg timmerende sociale wetenschappers. Dat is betreurenswaardig, maar het is een vorm van *regressie* om daaruit te concluderen, dat we terug moeten naar gezond verstand, begrip-als-bedoeld, en wijsheid gestold in taal.

Literatuur

- Anastasi, A., Heredity, environment, and the question 'How?'. *Psychological Review*, 65, 1958, 197-208.
- Anastasi, A., & Foley, J.P., A proposed reorientation in the heredity-environment controversy. *Psychological Review*, 55, 1948, 239-249.
- Burt, C., & Howard, M., The multifactorial theory of inheritance and its application to intelligence. *British Journal of Statistical Psychology*, 8, 1956, 95-131.
- Cavalli-Sforza, L.L., & Feldman, M., Cultural versus biological inheritance. Phenotypic transmission from parent to children. *American Journal of human genetics*, 25, 1973, 618-637.
- Eaves, L.J., Inferring the causes of human variation. *Journal of the Royal Statistical Society, A*, 140, 1977, 324-355.
- Eaves, L.J., Last, K.A., Martin, N.G., & Jinks, J.L., A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of mathematical and statistical psychology*, 30, 1977, 1-42.
- Elderton, E.M., A summary of the present position with regard to inheritance of intelligence. *Biometrika*, 14, 1923, 378-408.
- Eysenck, H.J., *Race, intelligence, and education*. London, Temple-Smith, 1971.
- Eysenck, H.J., *The Inequality of Man*. London, Temple-Smith, 1973.
- Eysenck, H.J., *The structure and measurement of intelligence*. Berlin-Heidelberg-New York, Springer, 1979.
- Eysenck, H.J., & Kamin, L., *Intelligence: the battle for the mind*. London, Pan Books, 1981.
- Fisher, R.A., The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 1918, 399-433.
- Galton, F., *Hereditary genius: an enquiry into its laws and consequences*. London, MacMillan, 1869.
- Goldberger, A.S., Heritability. *Economica*, 46, 1979, 327-347.
- Herrnstein, R.J., *IQ in the meritocracy*. Boston, Atlantic Monthly Press, 1973.
- Hogben, L., *Nature and nurture*. London, Allen & Unwin, 1939.
- Jaspars, J.M.F., & Leeuw, J. de, Genetic-environment covariation in human behaviour genetics. In: *Psychometrics for educational debates*. New York, Wiley, 1980.
- Jensen, A.R., How can we boost IQ and scholastic achievement? *Harvard educational review*, 39, 1969, 1-123.
- Jinks, J.L., & Fulker, D.W., Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behaviour. *Psychological Bulletin*, 73, 1970, 311-349.
- Kamin, L.J., *The science and politics of IQ*. Potomac, Earbaum, 1974.
- Kempthorne, O., Logical, epistemological and statistical aspects of nature-nurture data interpretation. *Biometrics*, 34, 1978, 1-23.

- Layzer, D., Heritability of IQ-scores: science or numerology? *Science*, 183, 1974, 1259-1266.
- Leeuw, J. de, De politieke relevantie van korrelaties. *Sociologische gids*, 25, 1978, 31-39.
- Leeuw, J. de, Psychometrische genetica. In: *Codex psychologicus*. Amsterdam, Elsevier Argus, 1981.
- Leeuw, J. de, & Stoop, I., Sekondaire analyse 'Van Jaar tot Jaar' met behulp van niet-lineaire multivariate technieken. In: *Van achteren naar voren*. 's-Gravenhage, Staatsuitgeverij, 1979.
- Lewontin, R.C., Genetic aspects of intelligence. *Annual Review of Genetics*, 9, 1975, 387-405.
- Loevinger, J., On the proportional contributions of differences in nature and nurture to differences in intelligence. *Psychological Bulletin*, 37, 1940, 725-756.
- Martin, N.G., The inheritance of scholastic ability in a sample of twins. *Annals of human genetics*, 39, 1975, 219-229.
- Mather, K., & Jinks, J.L., *Biometrical genetics*. London, Chapman & Hall, 1971.
- Morton, N.E., Analysis of family resemblance. I: Introduction. *American Journal of human genetics*, 26, 1974, 318-330.
- Munsinger, H., The identical twin transfusion syndrome. *Annals of human genetics*, 40, 1977, 307-321.
- Overton, W.F., On the assumptive base of the nature-nurture controversy. *Human development*, 16, 1973, 74-89.
- Pearson, K., On the laws of inheritance in man. *Biometrika*, 3, 1904, 131-190.
- Rao, D.C., Morton, N.E., & Yee, S., Analysis of family resemblance. II: A linear model for familial correlation. *American Journal of human genetics*, 26, 1974, 331-359.
- Urbach, P., Progress and degeneration in the IQ debate. *British journal philosophy of science*, 25, 1974, 99-135 en 235-259.
- Vroon, P., *Intelligentie*. Baarn, Ambo, 1980.