



Universiteit Utrecht

Jan J.F. ter Laak
Afdeling Ontwikkelingspsychologie
Faculteit Sociale Wetenschappen
Universiteit Utrecht

Diagnose van uw intelligentie, kennen en prestaties

Wat biedt het kennisbestand van de psychologie om uw intelligentie, kennen, geschiktheden en prestaties te bepalen? Voorspellen ze uw succes en falen? Hoe ontwikkelen ze zich en wat is de invloed van de sociale context?

Copyright Jan J.F. ter Laak

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier.

“Een van de meest bekende zwakheden van het menselijk denken, kennen en verstand is dat het tegenstrijdige principes met elkaar wil verzoenen en het verstand vrede wil in plaats van logica”.

Vrij naar De Tocqueville (19^{de} eeuw). Hoofdstuk VI ^a Boek II eerste deel.
En in zijn Hoofdstuk X Toepassingen van de wetenschap:

“In onze tijd moet de menselijke geest bij de theorie worden gehouden; zij dwaalt vanzelf af naar de toepassing, en in plaats van haar voortdurend terug te leiden naar het uitvoerig bestuderen van secundaire effecten, is het goed haar daar soms vanaf te houden om haar te verheffen tot de aanschouwing van de eerste oorzaken”.

Voor Marie-Louise, Mose, Ebba en Bor

Inhoud

Voorwoord

Ten geleide 1

Leeswijzer 11

Doelgroep en doel 11

Vooruitblik 12

Diagnose van uw intelligentie, kennen en prestaties

Wat biedt het kennisbestand van de psychologie om uw intelligentie, kennen, geschiktheden en prestaties te bepalen? Voorspellen ze uw succes en falen? Hoe ontwikkelen ze zich en wat is de invloed van de sociale context? 14

Figuur 1: Drie elementen en oriëntaties voor de psychologische diagnose van intelligentie en cognitie 15

I Geschiedenis en betekenis van de concepten intelligentie, cognitie, geschiktheid en prestatie 16

1. Intelligentie, cognitie, geschiktheid en prestatie 17

Intelligentie 17

Cognitie 18

Geschiktheid 18

Creativiteit 18

Prestatie 19

Samenvatting en conclusie 19

2. Geschiedenis van kennis-, intelligentie- en cognitiebegrippen 20

Pre-Socratisch Griekenland 20

Plato 20

Aristoteles 21

Middeleeuwen 22

Filosofie van de nieuwe tijd 22

Brits empirisme 24

Verlichting 25

Wijsbegeerte van de 20^{ste} eeuw 26

Wat biedt dit voor de diagnose van een cliënt? 26

Samenvatting en conclusie 27

3. Reflectie en evaluatie 28

II Impliciete theorieën van intelligentie en prestaties 30

1. Individuele verschillen 31

Westerse opvattingen over intelligentieverschillen en de relatie met creativiteit 31

Niet-westerse opvattingen over intelligentieverschillen 34

Hardnekkige ideeën over intelligentieverschillen 35

Samenvatting en conclusie 38

2. Ontwikkeling 38

Samenvatting en conclusie 40

3. Sociale context 40

Samenvatting en conclusie 40

4. Reflectie en evaluatie 41

III Expliciete theorieën van intelligentie en prestaties 42

1. Individuele verschillen 43

Intelligentie 44

Charles Spearman 44

De G-factor 46

Kritiek 47

Interpretatie van de G-factor 47

Thurstone 48

Guilford 48

Figuur 1: Guilfords Structure of Intellect (SOI) 50

Ordering van de factoren 50

Figuur 2: Het drie lagen concept van Carroll 51

Box: G, S, PMA, SOI: Het ene of het vele? 52

Een nieuwe generatie 53

Sternberg 53

Embretson 54

Pragmatiek: De Wechsler schalen 55

Oorzaken van individuele verschillen 57

Geschiktheden 64

Prestatie 68

Box: Het denken over leren en prestatie verandert voortdurend: niet bij te houden 68

Samenvatting en conclusie 70

2. Ontwikkeling 72

Jean Piagets theorie 72

Kritiek 74

Piagetaanse psychologie nu 75

Ontwikkeling van geschiktheden en prestaties 75

Samenvatting en conclusie 76

3. Sociale context 76

Genen en IQ 77

Sekse en IQ 77

Inkomen en IQ 77

Sterftecijfer en IQ 78

Gezondheid en IQ 78

IQs in verschillende landen 79

Toename van IQ over generaties 80

Geschiktheden 80

Prestaties 80

Samenvatting en conclusie 81

4. Reflectie en evaluatie 81

IV Alternatieve theorieën van intelligentie en prestaties 86

1. Individuele verschillen 87

Naturalistische intelligentie (Gardner) 87

Competenties 88

Leerpotentieel 89

Succesvolle intelligentie 90

Emotionele intelligentie 90

Samenvatting en conclusie 93

2. Ontwikkeling 94

Informatieverwerking 94

Samenvatting en conclusie 95

3. Sociale context 96

Samenvatting en conclusie 97

4. Reflectie en evaluatie 98

V Integratie van expliciete theorieën van intelligentie en prestaties 100

1. Eén theorie? 101

Claims van psychometrie, cognitieve ontwikkelingspsychologie en contexttheorieën 102

Intelligentie en hersenen 102

Veel verklaringen 103

Geschiktheden 104

Cognitieve ontwikkelingspsychologie 104

Contexttheorieën 105

Informatieverwerkingstheorie 105

Effect van de context op intelligentie en cognitie 106

Eén intelligentie- en cognitieve theorie? 106

Samenvatting en conclusie 106

2. Individuele verschillen x Ontwikkeling: stabiliteit tegenover cognitieve ontwikkeling 107

Homotypische stabiliteit van intelligentie en cognitie 108

Steekproef stabiliteit is niet individuele stabiliteit 109

Integratie van psychometrische intelligentie en cognitieve ontwikkeling 109

Samenvatting en conclusie 112

3. Individuele verschillen x Context: aanleg tegenover omgeving 112

Ontwarren van variantie in genen en de omgeving 113

Het Flynn effect 114

SES en prestaties 119

Scholing en prestaties 121

Klassengrootte en prestaties 121

Hoge verwachtingen van interventies 122

Samenvatting en conclusie 123

4. Ontwikkeling x Context: groeien versus leren 125

Versnellen van cognitieve ontwikkeling 125

Integratie op abstract niveau 125

Samenvatting en conclusie 126

5. Reflectie en evaluatie 126

VI Operationalisatie, meting en instrumentatie van intelligentie, cognitie, geschiktheid en prestatie 131

1. Individuele verschillen 132

Psychometrische intelligentie en geschiktheden 132

Interpretaties van correlaties 134

Een alternatieve representatie van de correlaties tussen tests 134

Elementaire processen 135

Figuur 1: Cognitieve tests afgebeeld als afstanden in een plat vlak 137

Alternatieve theorieën 139

Metten van prestaties 139

Figuur 2: De tests in het centrum representeren de G 140

Instrumentatie 141

Intelligentietests: Binet-Simon, WAIS, WISC, Rakit, Intelligentietest eindtoets

Basisonderwijs, NIO, Raven's Matrices, GIT 142

Geschiktheidstests: DAT en GATB 147

Prestatietests 148

Periodiek peilingsonderzoek 149

Criteria: betrouwbaarheid 150

Samenvatting en conclusie 152

2. Ontwikkeling 153

Ontwikkeling van cognitie vastgelegd met behulp van tests 153

Stadiumovergangen 155

Ontwikkeling van prestatie 157

Samenvatting en conclusie 158

3. Sociale context 158

Samenvatting en conclusie 159

4. Reflectie en evaluatie 159

Onderwerpen en namen Hoofdstuk I, II, III, IV, V, VI 162

Slotbeschouwing 164

Referenties en geraadpleegde literatuur 170-185

Auteur 186

Voorwoord

In neoliberale samenlevingen, kastensels in India en zelfs egalitaire ideologieën, zoals communisme, overal dus, tellen individuele prestaties en intelligentie. De testpsychologie is met het meten van intelligentie (IQ) groot geworden. We hebben allemaal een idee over wat intelligentie is en wat goede prestaties zijn. De individuele verschillen (test)psychologie heeft de domeinen van intelligentie en (school/werk)prestatie beheerst. Er is volop theorievorming over hoe het IQ is opgebouwd en waar het op berust. Het westers begrip heeft - niet verwonderlijk gegeven de filosofische achtergrond - zowel pragmatische als rationalistische achtergronden. Niet-westerse intelligentie- en prestatiebegrippen en alternatieve theorieën benadrukken de functie van intelligentie voor een soepel verloop van onze omgang met elkaar. De ene ware theorie van intelligentie is gewenst, maar het komt er niet van. Er zijn voor een klein gebruikersgebied als Nederland voldoende intelligentie- en prestatietests. Door de oprichting van het Cito is de instrumentatie een monopolie geworden van dat goed uitgeruste instituut. Prestatiedomeinen worden daar omschreven door inhoudelijk deskundigen. Zij schrijven de items en proeven. De moderne testtheorie is uitgewerkt door een aantal psychometrici, dat een eigen relaxed Rasch model ontworpen heeft. Rondom intelligentie en prestaties wordt het nooit stil. Het zijn geen uitgemaakte zaken. Dat blijkt hopelijk uit dit boekje.

Utrecht, najaar 2015

Jan ter Laak

Ten geleide

Dit geschrift bevat vooringenomenheden, willekeurige onderscheidingen en ongearticuleerde epistemologische en theoretisch inhoudelijke standpunten. Het maakt gebruik van ander werk en van vakgenoten als inspiratiebronnen. Een auteur ontkomt niet aan vooroordelen en preoccupaties over wetenschap, haar toepassing en beoefenaren. Hij is niet objectief en niet neutraal. Als een onderwerp hem niet interesseert of niets oplevert, leest en schrijft hij er niet over. Er wordt geen poging gedaan om deze persoonlijke bagage of ballast te verbergen. Als U deze prolegomena of reflectie vooraf liever mijdt, kunt U beginnen bij **Leeswijzer**, p. 11.

Voringenomenheden over wetenschap(pers), psychologie en diagnostiek
Wetenschappelijke arbeid is bedoeld om de werkelijkheid binnen en buiten de mens - het Ik en het Niet-Ik - te beschrijven, verklaren, controleren en voorspellen. Wat tot het Ik gerekend wordt wisselt in de psychologie: leer-, ontwikkelings- en informatieverwerkende mechanismen worden soms vervangen door biologische en neurologische. Hier wordt gekozen om gedrag psychologisch te beschrijven en te verklaren. Van gedrag zijn we voor een deel de eigenaar en dat zijn we niet direct van neurologische en biologische mechanismen.

Sociale wetenschap heeft met concurrenten van doen (Westbroek, 2013) in de vorm van systemen die beloven vragen te beantwoorden en behoeftes te vervullen: ideologieën, religies, heersende opvattingen, reclame, lobbyisten, en ga zo maar door. De psychologie is vooralsnog geen geslaagd systeem. De resultaten van studies en de successen van trainers, leraren en behandelaars geven daar onvoldoende reden toe.

De opbrengst van de diagnostiek wordt geridiculiseerd door universitaire psychologen (werk van doctorandi), hulpverleners (wat is uw meerwaarde; het is een mager halffabricaat) en leken (IQ is maar een cijfer). Deze houding belemmert het zicht op wat diagnostiek kan en is.

Diagnostiek gaat over het gedrag van de cliënt; in dit boek over kennen, informatie verwerken en intelligentie van de cliënt. Het is geen biologisch, medisch of natuurwetenschappelijk onderzoek. Daar heeft een diagnosticus geen verstand van en het zet hem op het verkeerde been. Dat denken, die definitie van het object en het onderzoek erover kan slechts metaforen bieden om gedrag te beschrijven.

Psychologie is zelfbetrokken. Het betekent dat *psychologische diagnostiek tegelijkertijd diagnostiek van de psychologie* is. Het kennisbestand van de psychologie en de activiteiten van diagnosticus en theoreticus/onderzoeker blijven niet buiten schot.

Het willen kennen en onderzoeken van het intelligent gedrag van de persoon vindt zijn

oorsprong in de rationele en emotionele uitrusting van de mens. Rationaliteit leidt tot theoretische constructen en empirische toetsing. Emotionaliteit leidt tot verlangen naar inzicht in wat de cliënt doorziet, begrijpt en kan. Het komt soms voort uit een hang naar een veilig geloof. Datzelfde kennen kan ook aanleiding geven tot ongerustheid en onzekerheid over wat je kunt bereiken. Beide zijn nodig voor dynamiek en balans.

Wetenschappers en practici hebben niet volledig toegang tot de fysisch en sociale werkelijkheid. Kennis verschuift de horizon, verlegt de grens maar volledige kennis van objecten, verschijnselen en subjecten bestaat niet. Ze komt ook niet tot stand door integratiepogingen binnen en tussen wetenschappen.

Elke activiteit - dus ook kennen, leren en denken - is ingebed in het alledaagse leven. Onze geleefde ervaring is de basis van onze kennis met inbegrip van onszelf en mensen om ons heen: Merleau-Ponty (1945/1957) en zijn opvolger Piaget.

Wetenschap wordt door elites soms gebruikt om hun positie te behouden en privileges te beschermen. Het publiek wordt op afstand gezet terwijl het ziet dat wetenschappers het zelden eens zijn en er zich onder hen fraudeurs bevinden (Van Kolfshoten, 2012). Er is een geloof dat het beoefenen van wetenschap is weggelegd voor de begaafden. Als het publiek en de politiek zien dat die begaafde deskundigen geen inzichten geven en problemen oplossen wordt de weldenkende burger gevraagd thema's te noemen en vragen te stellen waar onderzoekers zich over gaan buigen en antwoorden op geven.

Discussie op grond van argumenten is een onvermijdelijk onderdeel van wetenschappelijk werk. Als de psychologie een wetenschap is - daar ben ik niet zo zeker van - hoort ook daar een dergelijke discussie thuis. Als ze geen wetenschap is maar informatieverzameling over gedrag dan blijft die discussie ook nodig.

Wetenschappers kunnen ontsporen door zich in te stellen op het overwinnen van anderen en het behalen van succes door bijzaken-barok, dat wil zeggen door zich 'met overmaat aan geleerdheid en sluwheid te richten op kleine gebieden of maatschappelijke thema's, daarbij een vaardigheid ten toon spreidend die weinigen hen zullen verbeteren en waar geldig en vals en goed en *tactische* begrippen zijn' (vrij naar Menno ter Braak, cursief, jtl).

Psychologiebeoefening heeft evenveel te maken met politiek als met het belangeloos verwerven van inzichten. De onderzoeksagenda en subsidies worden niet (alleen) door intrigerende belangeloze vragen bepaald maar door belangen en lobby's van gevestigde groepen. Op een vraag aan een naoorlogse Britse premier wat nu het politiek handelen stuurde zei hij: *Events*. Psychologisch onderzoek wordt wellicht meer dan we denken gestuurd door *events*. Historici van de psychologie noemen haar '*eventful*'.

Het 19^{de} eeuwse romantisch beeld van de wetenschapper is dat van de uitzonderlijk begaafde of van de briljante psycholoog die door je heen kijkt. Het cliché van transpiratie

(95%) en inspiratie (5%) komt dichterbij wat er gebeurt in onderzoek en praktijk: respectievelijk het routinematig uitvoeren van een gefinancierd project (*output*) plus het opschrijven volgens richtlijnen van het *Publication Manual* van de APA en het rapporteren over de cliënt volgens een protocol.

Wetenschappers zijn conservatief. Wat ze weten is kostbaar en moeizaam verworven. Een succesparadigma wordt niet gauw opgegeven. De VU wiskundige Meester (2014) voegt er aan toe dat sommige arrogant zijn en hun kennis en weten overschatten. Ze zien verschijnselen vanuit een tunnel duiden en duiden afwijkingen van hun visie als onwetenschappelijk. Het ging daarbij over onzekerheid in de evolutieleer. De onzekerheid werd opgevat als afwijking van die leer met een opening naar een schepping en een schepper.

Psychologen vertellen wat het publiek en tijdschriftredacties willen horen en vermijden te zeggen wat deze niet wil horen. Gemakkelijk bij het publiek en redacties liggende thema's worden uitgewerkt, bijvoorbeeld man-vrouw verschillen, agressie, positieve psychologie, persoonlijke groei en ontwikkeling, genetische basis van gedrag, het lijden dat zich overal voordoet en de rol van hersenen voor begaafdheid en pathologie. Als onderzoeksgegevens afwijken van wat *Dass Man* - van het publiek tot de bestuurlijke elite vindt - valt er iets uit te leggen. Succes is niet verzekerd.

Psychologen kunnen het publiek en de bestuurlijke elite ook een spiegel voorhouden en aan de hand van gegevens tonen dat ze niet het centrum van de wereld zijn, geen *tabula rasa*, niet rationeel of irrationeel, niet vrij of onvrij, niet gelijk of ongelijk, niet transparant voor elkaar, maar ook geen vreemden en niet onafhankelijk of afhankelijk.

Psychologists go, where the money goes Dit uit zich in aansluiting bij modieuze thematieken, goed in het gehoor liggende theorievorming en methodologie. Een *kitchen and sink* typologie van psychologen met mijn voorkeur voor driedelingen is: de Opportunist: steeds een nieuw thema kiezen, dat het goed doet: van geleerde hulpeloosheid naar positieve psychologie, van Piaget naar *theory of mind*, naar agressie, naar..., de Monopolist: één eenvoudig populair thema kiezen en dat monopoliseren: bijvoorbeeld de *Big Five*, *Attachment*, Adoptie en Echtscheiding en de Bekeerling: van psychometricus naar persoonlijkheidspsycholoog of vice versa. Een bekering heeft iets moedigs vooral als je er niet mee op de TV komt.

Het is het moeilijk uit te maken of meta-analyses en meta-analyses van meta-analyses iets nieuws brengen of een eind maken aan eenzelfde type onderzoek, c.q. een vraag afdoende beantwoorden of een probleem oplossen. Wellicht is het slechts een superieure vorm van boekhouden. Is dit het westers ultieme geloof in de Zelfopenbaring van het Zijn? Meta-analyses, *big data* en *data mining* zijn niettemin zinvol, al steunen ze op de empiristische veronderstelling: hoe meer onderzoek, hoe beter. Empirisch onderzoek ontuchtigt en matigt hoge verwachtingen weliswaar, maar geen enkel onderzoek kan zonder een

gedisciplineerd geloof. Dat wil zeggen het kan niet buiten de empirie om bedachte theorieën, structuren en modellen, gecreëerde betekenissen in de hermeneutiek en analyse van vooronderstellingen over wat de wetenschap en de mens zijn en bewerkstelligen.

Er is een pikorde in de wetenschap. Deze zet de psychologie op afstand van de fysica, scheikunde, neurologie en biologie. Sociologie, antropologie, taalkunde en wijsbegeerte worden er onder geplaatst (Simonton, 2004, 2009). Binnen de psychologie is er eveneens een pikorde met functieleer aan de top en ontwikkelings- en klinische & persoonlijkheidspsychologie onderaan. Deze pikordes verhinderen om van een andere (sub)discipline te leren en te profiteren.

Onderscheidingen Om de diagnostiek van de cliënt te structuren worden drie niet willekeurige bronnen van theorie onderscheiden: impliciete alledaagse, expliciete theorievorming uit onze tekstboeken (de heersende paradigma's van Kuhn, 1962) en alternatieven voor expliciete theorievorming. Deze worden geordend in drie oriëntaties: individuele verschillen in intelligentie, cognitieve ontwikkeling en invloed van de (gemanipuleerde of natuurlijk gegeven) fysische en sociale context op kennen, denken en intelligentie. De drie dekken bijna alle theorieën in de tekstboeken.

In dit boek staat het denken, kennen, cognitie, intelligentie, geschiktheid en prestatie van de cliënt centraal. Dat is een hoog aggregatieniveau in de studie van gedrag. Voor atomistisch denkende onderzoekers is dat onbegonnen werk.

Beschrijven, voorspellen, verklaren/controleren en beslissen worden als doelen van diagnostiek geoormerkt. Hun succes hangt af van toegankelijkheid en helderheid van conceptuele analyses en empirische bevindingen, onder meer resultaten van meta-analyses en van het gezonde verstand, want sommige analyses vragen naar de bekende weg en andere zijn pseudo-empirisch (Smedslund, 2009).

Beschrijven is naast categoriseren - bijvoorbeeld zoals Linnaeus dat met planten deed en de DSM – ook het ontwerpen van een wiskundig geformuleerd model dat uit zijn aard intrinsiek klopt. Dat kan een doel op zich zijn. Deze voorkeur is terug te vinden bij psychometrici met hun modellen, functies en formules, maar ook bij structuralisten waar Piaget een voorbeeld van is. Van vertegenwoordigers van beide werk- en denkwijzen wordt verwacht dat ze een brug naar observeerbaar gedrag slaan. Als dat niet zichtbaar gemaakt wordt is het elegant, maar zonder betekenis voor diagnostiek.

Verklaren gaat terug op de Griekse Pre-Socratici en Plato en Aristoteles. Het ging hen om een verklarend principe (oorsprong, *archè*), zodat alles opgenomen kon worden in een overkoepelend geheel of uit één bron oplichtte. Zo worden uiteenlopende zaken verklaard. Aristoteles wilde bijvoorbeeld met zijn principe van de gulden middenweg zowel evenwichtstoestanden bij natuurverschijnselen (harmonie der sferen) als politiek (zijn vorm van stadsstaat democratie) en ethiek (hoe het goede leven: *eudamonia* bereiken) verklaren. Bij diagnostiek als een doelgerichte activiteit komen beide als een *hybride* aan de orde:

diagnostische kennis bestaat voor een deel uit categoriesystemen, structuren en formules én berust op empirische feiten en ervaringskennis (inter- en supervisie) én op diagnostische praktijken als observatie, experimentjes om gedrag te ontlokken aan de cliënt en het gebruik van een valide instrumentarium.

Standpunten De grenzen van voorspellen, verklaren/controleren en beslissen van en over gedragingen tekenen zich af als onderzoek voortgaat op de wijze waarop het nu vormgegeven is en gepubliceerd wordt. Dit standpunt wordt gemotiveerd en geïllustreerd met resultaten van meta-studies.

Gedragingen en sociale contexten zijn dynamisch en variabel. Bij het verkennen van grenzen van de impact van oorzaken en de hechtheid van samenhangen wordt nadruk gelegd op effectgroottes, betrouwbaarheidsintervallen, *power* van toetsen, *eye-balling* en gezond verstand.

Constateringen en aarzelingen De psychologie is rijk aan tweedelingen. Ik kies niet voor de een of de ander. Ze zijn gefabriceerd en uitgevonden, zoals lichaam versus geest (*mind-body*), theorie versus meten, persoon versus situatie, stabiliteit versus verandering, natuur versus cultuur, links versus rechts en genen versus omgeving.

Er is aarzeling om complexe protocollen, regels en statistische modellen te aanvaarden als veilige en vooral afdoende hulpen bij het diagnosticeren van de persoon van de cliënt. Dit verwijst niet zozeer naar Piet Vroons slagzin: 'Hoe meer regels hoe meer vlegels' als wel naar het feit dat ze hun doel voorbij kunnen schieten, gegeven de beperkte informatieverwerkingscapaciteit en de natuurlijke werkwijze van de diagnosticus.

Er is terughoudendheid om correlaties tussen stimuli en responsen uit experimentele en predictor-criterium studies te benutten als basis voor individuele voorspellingen, verklaringen, beslissingen en voor diagnose-behandeling-combinaties.

Aard van de tekst Deze bevat geen recepten voor afnemen van interviews en tests, het uitvoeren van interventies en het schrijven van psychologische rapporten. Ze is bedoeld als hulp bij een *rijke en voldoende beschrijving van het probleem/de vraag van de cliënt*, zoals Simon dat voorstond en De Groot aanbevolen heeft. Daarnaast is de ecologische oriëntatie van Brunswik inspiratie en leidraad. Het gaat hier om de oude Simon. Hij verdedigde later dat intelligentie louter het manipuleren van abstracte symbolen is en zet de relatie met de biologische *wetware* en psychologische processen, de *software* tussen haakjes of denkt die niet nodig te hebben. De symbolen en structuren zijn bij hem geen metafoor. Ze zijn letterlijk bedoeld. En elke symbolische en formele structuur kan niet zonder semantische interpretatie in de psychologie (Searle, 1984).

Methode en inhoud Methodologie gaat over het tegengaan van vertekeningen en het trekken van causale conclusies. Ze heeft daarnaast een epistemische dimensie nodig om het object van onderzoek, het gedrag van de cliënt, recht te doen. Er wordt naar gestreefd om

een relatie tussen model, formule en iets buiten het model of de formule te verduidelijken. Waarheid, geldigheid en objectiviteit gaan over een relatie en dat is niet alleen de consensus tussen experts of die van een forum van weldenkende collega's. Zij sluit een contrafactisch streven in want die relatie, uitspraak, *bewering* over gedrag en het *waargenomen* gedrag is nooit volledig en sluitend gelegd. Als dat het geval was dan zouden wetenschap en onderzoek al lang voltooid zijn.

Er wordt naar gestreefd om naast aandacht voor individuele verschillen, ontwikkeling en sociale context te bespreken. In diagnostisch intelligentie onderzoek overheerst de individuele verschillen oriëntatie. Dat betekent dat daar het meeste onderzoek verricht is en gepubliceerd is.

Het onderwerp van elke diagnose is een gedrag, d.w.z. vraag, probleem van de cliënt. Menselijk gedrag is onderwerp van elke wetenschap. Zonder de mens is er immers geen wetenschap. De mens en zijn gedragingen zijn bijgevolg intrinsiek een *go-between* tussen alle wetenschappen. Het risico om daarbij buiten je boekje te gaan is groot.

Aanleunen tegen andere vakken Fysica, scheikunde en biologie hebben een hogere status dan de psychologie. De identiteit van psychologen ligt minder vast dan die van fysici en medici. Meedeinend met de tijdgeest en zo nu en dan geconfronteerd met een identiteitscrisis, richt de psycholoog zich soms op andere takken van wetenschap, bijvoorbeeld neuropsychologie en hersenonderzoek: het geloof in het DNA-maals als substituut voor het hierna-maals. En, evenals bij religie gaat het daarbij niet zelden om belangen. Verder doet zich in een poging gerespecteerd te worden een toewending voor naar beschrijving met behulp van formele en statistische modellen. Wiskunde is immers de exactheid en geldigheid zelf en staat hoog in de wetenschappelijke pikorde. Men zet zich af tegen losjes geformuleerde verklarende semantische theorieën en constructen. De modellen zijn waar. Ze kloppen op de wijze van regels van het schaakspel. En, bij een confrontatie met psychologische onderzoeksvragen lopen ze door hun complexiteit soms voor op de realiteit van de beschrijving van concrete gedragingen.

Sociologie, culturele antropologie en taalkunde gaan over groepsgedrag en culturele praktijken. Deze worden in de psychologie enigszins op afstand gezet. Ten onrechte want ze liggen niet zover van de psychologie af en formuleren mechanismen die er toe doen om het gedrag van de cliënt te begrijpen, beschrijven en verklaren.

Een *horizontal world view* wordt aanvaard (Van Dijk & Wilthagen, 2014). Deze steunt op Wittgenstein en James' werk waarin het hogerop zoeken door reductie, zoeken naar de kleinste eenheid en naar verborgen onder- en bovenliggende lagen secundair is aan de concrete beschrijving met behulp van voorbeelden. Wittgensteins opmerking: (1953, sec. 2 xiv): *The existence of experimental method makes us think we have the means of solving*

problems which trouble us; though problem and method pass one another by heeft niets van zijn kracht verloren.

Psychologische mechanismen In deze tekst ligt de nadruk op de vraag/het probleem van de cliënt en op psychologische mechanismen en op constructen die zijn gedrag beschrijven, verklaren en voorspellen. Omdat je niet alles tegelijk kunt beperk ik me in dit boek tot intelligentie (het IQ), kennen, denken, informatieverwerking en cognitie van de cliënt. Dit is *maximum performance* ter onderscheiding van *typical performance*.

Om de intelligentie van de cliënt in concrete situaties voor ogen te houden wordt aangesloten bij Brunswiks representatieve ontwerpen voor onderzoek, bij Rorty's afwijzen van *foundationalism* en bij Smedslunds nadruk op culturele praktijken en de invloed van onmiddellijke situaties.

Vrijheid zolang dat helpt kennis te verkrijgen Onderzoekers en diagnostici mogen vrij analyseren, denken en theoretiseren. De vrijheid is er ook aan de methodologische zijde. Er rust niet bij voorbaat een taboe op intuïtieve, subjectieve of kwalitatieve benaderingen naast de experimentele en correlatieve methoden. Dit heeft zin zolang de practicus niet verdwaalt in vertalingen in de vorm van reducties, verborgen mechanismen, producten van neurale activiteiten, herenscans, kleinste eenheden en veronderstelde hogere mentale bewustzijnstoestanden. Dit leidt tot een *Lost in translation*, zoals in de film waar de twee congresgangers, de acteurs Bill Murray en Scarlett Johansson, niets verstaan van de lezingen in het Japans en maar met elkaar gaan praten, zoals studenten bij hoorcolleges. De verticale metafoor van het dieper en hoger zoeken, bevat als iedere metafoor - ook de horizontale - het risico *of holding us captive* (Wittgenstein, 1953, Sectie 115). Een open, eclectische houding past bij diagnostisch werk en omdat de diagnosticus zélf een rol in speelt in de diagnose is inter- en supervisie nodig.

Wat betekent psychologie voor intelligentiediagnostiek Ze is hofleverancier van theorieën, methoden en geldigheidscriteria voor diagnostische uitspraken over intelligentie. Vooruitgang of stagnatie in diagnostiek hangen samen met ontwikkelingen in de psychologie. Is er na Watson, William James, Skinner, Freud, Piaget, Spearman, Newell, Simon, Tversky en Kahneman en in Nederland na Strasser, Linschoten, Kouwer, De Groot en vele anderen iets nieuws te verwachten? Komt er een doorbraak, een andere kijk die ons dichterbij brengt bij begrijpen/verklaren van het gedrag van de cliënt of een nieuw en belangrijk facet van zijn gedrag aan het licht brengt? Kan er originaliteit verwacht worden op een gebied dat zo breed is en waar alle mogelijke standpunten, voorkeuren en temperamenten al lang en vaak beschreven en vertoond zijn en iedere zogenoemde nieuwe gedachte meteen in een vertrouwde categorie ondergebracht kan worden? Of doen er zich subtiele, bijna onopgemerkte verschuivingen voor, zoals met de PC en het internet? En, als het zich voordoet, wordt het dan opgemerkt? Of gaat het ermee als met de klokkenluider:

hij wordt niet gehoord en als er wel naar hem geluisterd wordt, belandt hij op dood spoor of wordt buitengezet.

Diagnostiek heeft robuuste bevindingen en resultaten nodig om het gedrag van de cliënt te kunnen voorspellen & controleren en recht te doen. Meta-studies beloven zulke bevindingen te bieden maar laten bescheiden tot gemiddelde waarden en grote variabiliteit zien: r capaciteiten - werkprestaties tussen .15 en .65, predictieve validiteit van tests: r tussen de .02 en .75 en sekseverschillen in schoolprestaties: d -waarden tussen -0,30 tot 0,17. Dat duidt op steekproef- en lokaal gebonden zijn van resultaten. De recente replicatiecrisis is er altijd geweest, zie de varianties van d - en r - waarden in iedere meta-studie.

Literatuur Geraadpleegde literatuur en vermelde studies zijn noodzakelijkerwijs een selectie. Ze zijn zo gekozen dat kennis nemen daarvan in staat stelt specifieke onderwerpen uit te diepen. Omdat tijdschriften elektronisch toegankelijk zijn is dat te verwezenlijken. Als begrippen en namen niet bekend klinken zijn ze te vinden via zoekmachines. Opsomming van empirische onderzoeken uit de literatuur is droge kost. Ze is bedoeld als voorbeeld, illustratie van bevindingen en is genummerd en apart gezet. Naast klassieke studies wordt recent onderzoek vermeld met voorkeur voor meta-studies. Recent onderzoek toont gesofisticeerde analyses maar ze zijn bijna nooit aanleiding om oude conclusies te herzien. Na opsomming van empirisch onderzoek volgt een conclusie. Er is naar tegenvoorbeelden gezocht. Als ze gevonden zijn, worden ze vermeld.

Lezen en schrijven Elk boek is, zoals Plato (4^{de} eeuw *Before Common Era*: BCE) zei een *farmakon*, dat wil zeggen zowel een medicijn tegen het vergeten als een gif voor het geheugen en het kennen. Elke lezer loopt het risico ontgift te moeten worden na lezing van welke tekst dan ook. Je kunt een tekst op verschillende manieren lezen. Een is zoeken naar eigen thema's en vaststellen of die er niet of onjuist in staan. 'Ik word niet/verkeerd geciteerd'. Een grondslagen zoekende tekst wordt nauwelijks gelezen. Onderzoekers zijn immers op precies te omschrijven onderwerpen uit werken ze uit in brokjes: *least publishable units*. Een collega zei eens onderzoek doen is als een blind paard op je doel afgaan. Verder wordt originaliteit wordt betwist: 'Dat heb ik al eerder verteld in mijn artikel van 10 jaar geleden'. 'Waar heb ik dat eerder gehoord'. Een tweede is nagaan of het waar of juist wat er staat. Dat is moeilijk in de psychologie. Daar geldt eerder hermeneutiek dan waar/juist versus onwaar/vals. Het betreft ontdekken en verbreden van betekenis en zin. Wat is de boodschap? Een derde is retorisch lezen: de argumentatiestructuur ontrafelen. Het schrijven van een boek is pretentief en dit bevat ook nog eens geen nieuwe inzichten en is niet oorspronkelijk. Het brengt iets bijeen van wat de psychologie te bieden heeft voor de diagnostiek van het denken de intelligentie van de cliënt. De pretentie is dat bestaande kennis en methoden bijeengezet, uitgelegd en geëvalueerd worden. Theorie en onderzoek worden geordend in drie elementen, drie bronnen en drie oriëntaties.

Wetenschappelijk psychologisch onderzoek Er wordt geen poging gedaan om de overmaat aan onderzoek en theorie te verzoenen of te integreren. Ze krijgen afzonderlijk plaats en aandacht. Combinaties leiden tot onvruchtbare hybriden, *muddy dichotomies*. Je mag en kunt betwijfelen of psychologie een wetenschap is. Ze bevat misschien slechts informatie over samenhangen tussen gedragingen onderling en over oorzaken uit sociale en fysieke context op gedrag. Wetenschappelijke tijdschriften zijn in dat geval dagbladen, *journals* en onderzoekers *journalists*. Ze bevatten informatie, geen kennis of inzicht, laat staan dat ze wetten opleveren in de Newtoniaanse zin. Deze twijfel wordt toegelaten want misschien geldt dat voor de psychologie. Er wordt alleen geprobeerd een stap verder te komen. Er is een kennisbestand en dat zinvolle informatie, feiten over de kennis, denken en intelligentie van de cliënt op een rij zet.

Zonder als diagnosticus bescheiden hoeven te zijn over het vak - het is wat het is - leiden theorie en onderzoeksresultaten tot een bescheiden houding over wat beschreven, verklaard en voorspeld wordt over het gedrag van de cliënt.

Persona van de diagnosticus Een archetypisch beeld van wat een diagnosticus is - gegeven zijn karakter en drijfveren - ziet er in deze tekst zo uit: Hij biedt geen wetten, waarheden, dogma's en geen nieuwe apparaten of technieken. Hij toont relativiseringsvermogen en een door het kennisbestand van de psychologie getemde verbeelding over het kennen van de persoon die de cliënt is. Hij is niet in eerste instantie degene die problemen oplost, onderzoeksgeld binnenhaalt of een onderzoeksbedrijf opzet en runt. Hij is probleem-, niet oplossingsgericht. In dit boek wordt in problemen en niet in oplossingen gedacht. Het is niet gericht op hulpverleners. Het wil de vorming van een *rich and sufficient description of the problem field* bevorderen met inbegrip van de ecologische structuur waarin een individuele vraag of probleem is ingebed. Het doel is het midden te houden tussen het Hegeliaanse *Dass Wahre ist das Ganze* en het reductionisme van het behaviorisme, neurologisme, empirisch-analytisch methodologisme, staticisme, monotheorisme en beperking tot één psychologische thema of één methodische invalshoek.

Het midden houden, maat houden is één van de vier Aristotelische deugden. Ze kwamen in de Middeleeuwen meestal aan de orde bij begrafenissen, nu bij afstudeerpraatjes, feestjes en begrafenissen en daar hebben we de *Big Five* voor. De vier passen op diagnostiek van de cliënt: matigheid, verstandigheid/voorzichtigheid, dapperheid/sterkte en rechtvaardigheid. Met het wijzen op deugdzaamheid is het uitkijken, want deugden wordt door niets zozeer onderuitgehaald als door de saaiheid van haar pleitbezorgers.

Psychologische diagnostiek is een bezigheid *sui generis*. Daar bedoel ik mee dat de diagnosticus zelf zijn weg moet zoeken. Dat wil niet zeggen dat hij autonoom is, want diagnostiek staat in relatie tot het kennisbestand van de psychologie, de cultuur en de samenleving. Haar rol en bijdrage worden steeds opnieuw uitgevonden. Recent staat *valorisatie* in de aandacht. Wat draagt ze bij aan analyse van vragen en problemen in de samenleving?

Dit boekje is niet om te volgen of te repliceren, maar om eigen te maken, kritisch te evalueren en te valoriseren om terminologisch bij de tijd te zijn. Elke tekst, protocol, model, empirisch resultaat wordt rammelende ballast, dood gewicht als het niet wakker gelezen en getoetst wordt door een geïnteresseerde, gemotiveerde en onafhankelijke lezer.

Leeswijzer

Ordering van de tekst Elk onderwerp is gemarkeerd door twee nummers: het eerste duidt het hoofdstuk aan en het tweede de sectie. De secties bevatten subthema's die cursief gedrukt zijn. Elke sectie wordt afgesloten met een Samenvatting en conclusie. Voorbeelden van empirische studies zijn genummerd. Als er tegenvoorbeelden gevonden zijn worden ze vermeld. Bovendien wordt de variantie van de uitkomsten van vergelijkbare studies vermeld. Elk hoofdstuk eindigt met Reflectie en evaluatie. Reflectiviteit houdt in dat de ambigue, gefragmenteerde en betwistbare aard van psychologische kennis wordt erkend met inachtneming van feiten en aannemelijke interpretaties. Het houdt ook in dat er multiple perspectieven op gedrag zijn en er verschillende constructen geldig zijn om die te beschrijven. Er wordt gelet op de positie die een onderzoeker inneemt. Is hij lid van een groep met bepaalde praktijken? Dit beïnvloedt zijn waarneming en begrip van gedragsverschijnselen. Evaluatie verwijst naar een oordeel over de kwaliteit van de constructen en kennis: wat is bereikt, zijn pretenties waargemaakt, worden bevindingen realistisch weergegeven? Aan het eind van de hoofdstukken worden *Onderwerpen* vermeld die als subjectindex gelezen kunnen worden.

Geen glossary Psychologische constructen liggen niet vast. De Wittgenstein van *Philosophical Investigations* (1953) beweert dat definities in de psychologie niet veel verhelderen omdat ze betekenis krijgen in een context. De sinoloog Schipper (1988, p. 13) noemde het ontbreken van definities kenmerkend voor de open leer van het Chinese Taoïsme. Heideggers *Sein und Zeit* (1927/1962) kan gelezen worden als een kruistocht tegen de solide definities van het Zijn door de christelijke filosofische traditie. Het ernstig nemen van deze standpunten houdt in dat ik er niet naar streef de diagnosticus definitieve constructen aan te praten.

Boxen Er zijn boxen tussen de tekst gevoegd in een kleinere letter. Deze vermelden achtergronden, debatten en controversen en bevatten ook idiosyncratische gezichtspunten op een psychologisch of maatschappelijk verschijnsel.

Referenties De vele literatuur kan nooit recht gedaan worden. De referenties zijn bedoeld als vindplaats. Er wordt vanuit gegaan dat de lezer toegang heeft tot elektronische tijdschriften. Voor niet onmiddellijk inzichtelijke begrippen zijn er de zoekmachines.

Voorkennis Er wordt bij de lezer een actieve kennis van methodologie, psychologische theorievorming, psychometrie en statistiek op BA niveau verondersteld.

Doelgroep en Doel

Dit boek is bestemd voor studenten BA en MA psychologie en pedagogiek en HBO toegepaste psychologie. Het is ook bedoeld voor die practici en GZ psychologen in opleiding, die enerzijds het aan de leiband een model of protocol lopen waarderen als het verstrekken van duidelijkheid hoe het moet, maar anderzijds willen reflecteren op voorschriften die ze soms ontoereikend achten voor hun beroepsuitoefening. De tekst is ook bestemd voor leken die psychologische diagnostiek van intelligentie kritisch willen volgen. Er wordt een

hoofdstuk gewijd aan het 'leken'oordeel en -inzicht in intelligent functioneren en handelen. Niet gearticuleerde beweringen en kennis kunnen opgezocht worden op het internet. Het doel is het bevorderen van een zoektocht naar de grondslagen voor diagnostiek van ons gedrag en de activiteit van het diagnosticeren. Er wordt een reflectieve houding aangemoedigd: denk na en kijk uit, voor je springt. Er worden geen praktische handvatten aangereikt die voorschrijven wat een diagnosticus moet doen in de vorm van diagnose-behandel-combinaties (DBC's). Daar zijn andere en betere teksten voor van bijvoorbeeld Witteman et al. (2014) en De Bruyn et al. (2015).

Vooruitblik

Eerder heb ik aannemelijk gemaakt (ter Laak 2011, 2015) dat psychologische diagnostiek heeft *geen* eigen en/of specifiek gedragsdomein (materieel object) en geen eigen gezichtspunt op gedrag (formeel object) heeft. Daarin wijkt ze af van de subdisciplines in de psychologie. Ze gaat in beginsel over *elk* gedrag en ontleent *ad hoc* aan het kennisbestand van de psychologie al mogelijke theorieën, methoden en empirische bevindingen. Ze is uit zichzelf onbepaald. Een organisatorisch gevolg daarvan is dat ze niet ingebed is in de structuur van een faculteit, afdeling of vakgroep. Niettemin wordt diagnostiek binnen vrijwel alle afdelingen van de psychologie en pedagogiek faculteiten bedreven en in de praktijk van de gezondheids-, onderwijs- en organisatiepsychologie en de pedagogiek. In Nederlandse handboeken over diagnostiek wordt gesuggereerd dat de diagnose en het diagnostisch proces zo vormgegeven zijn dat zij de toets van de wetenschappelijke kritiek kunnen doorstaan. Witteman et al. (2014) beschrijven de diagnostische stappen die

'...een leidraad voor het uitvoeren van psychodiagnostisch onderzoek in de praktijk bieden; ...inzicht geven in hoe verschillende stappen in het psychodiagnostisch onderzoek op een goede manier uitgevoerd kunnen worden' (onderlijning jtl).

De Bruyn et al. (2015) maken een handleiding, een praktijkleer die laat zien...

'dat diagnostische besluitvorming in de praktijk op een wetenschappelijk-professioneel verantwoorde wijze kan worden doorlopen' (p. 18) (onderlijning jtl).

Als *caveat* wijzen laatstgenoemde auteurs weliswaar de beperkingen van het kennisbestand van de psychologie. Niettemin leidt het volgen van het prescriptieve model van de diagnostische cyclus bij hen tot wetenschappelijk en professioneel verantwoord diagnosticeren. In deze tekst doe ik *een stap terug* en werk uit wat de psychologische diagnose van de intelligentie van de cliënt aan het kennisbestand van de psychologie kan ontlenen om zijn kennen en presteren te beschrijven en die kennis te benutten voor de diagnose van zijn vraag/probleem. Ik vraag me ook af wat de cliënt erop kan zeggen. Hij is mondig en praat terug.

Hoofdstuk I gaat over geschiedenis en inhoud van onze opvattingen over denken en intelligentie. Het domein wordt voor in drieën verdeeld: intelligentie/cognitie, geschiktheid

en prestatie. De hoofdstukken II en III betreffen impliciete *beliefs* en expliciete theorievorming over intelligentie/cognitie, geschiktheid en prestatie. In hoofdstuk IV komen alternatieve theorieën aan bod. Deze zijn te begrijpen als verzet tegen de reducties en eenzijdigheden van expliciete theorievorming. Hebben de 'alternatievelingen' een punt of is het gerommel in de marge? In hoofdstuk V vraag ik me af wat pogingen tot integratie van de drie oriëntaties (individuele verschillen, ontwikkeling en sociale context) opgeleverd hebben. Verrijken ze ons inzicht of zijn het onwerkbaar hybriden? Tocqueville (18^{de} eeuw) zag al in dat verzoenen heilloos was. Elke expliciete theorie bevat reducties. Intelligentie/cognitie, geschiktheid en prestatie zijn gearticuleerd als begrippen. Wetenschappelijke begrippen of constructen zijn abstracties die inhoud krijgen door operationalisaties in taken en opdrachten. Deze operationalisaties leiden tot metingen. Dan wordt in realistisch systeem afgebeeld in een formeel systeem. Het meten van intelligentie (IQ) geschiktheid (bijvoorbeeld differentiële aanleg tests) en prestaties op school en in het beroep is een kernactiviteit van psychometricus en diagnosticus.

Diagnose van uw intelligentie en prestaties

Wat biedt het kennisbestand van de psychologie om uw intelligentie, kennen geschiktheden en prestaties te bepalen? Voorspellen ze uw succes, hoe ontwikkelen ze zich en wat is de invloed van de sociale context?

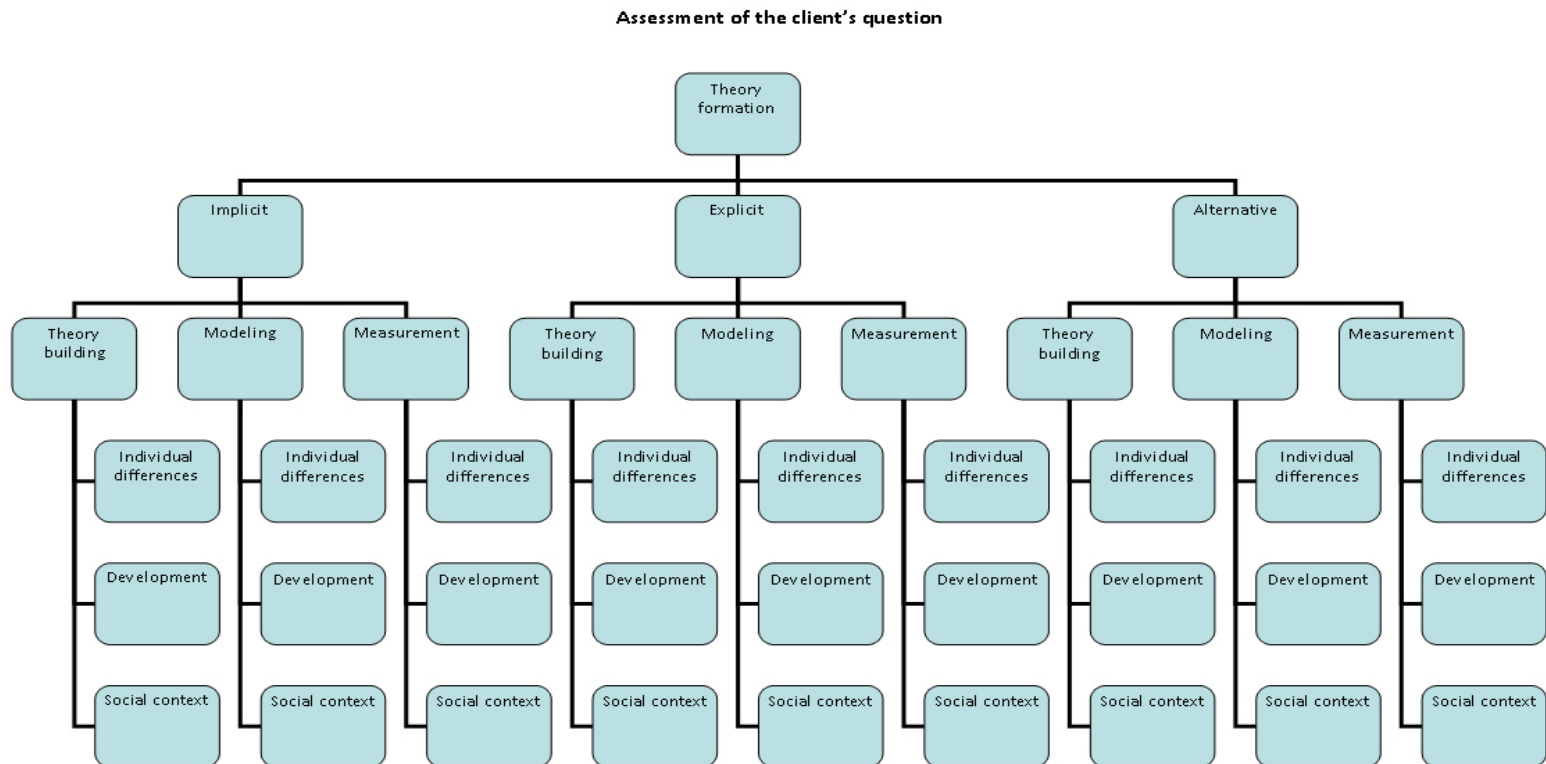
Intelligentie, eerst de mentale leeftijd daarna het deviatie IQ met een gemiddelde van 100 en een standaarddeviatie van 15 per leeftijdsgroep zijn bedacht met als doel selectie van leerlingen en personeel en voorspelling van school- en werkprestaties. Cognitie is nauw verbonden met ontwikkeling van intelligentie en met informatieverwerking. Beide verwijzen naar de manier waarop we kennis verwerven. Dat wil zeggen hoe we ons verhouden tot en een cognitieve, rationele, bewuste uitwisselingsrelatie aangaan met fysische objecten en verschijnselen, sociale gebeurtenissen en (gedrag van) mensen. Aan intelligentie worden kenmerken verbonden als kennen, gericht denken, doelgericht handelen, rationeel beslissen en effectief interacteren. Daarnaast verwijzen begrippen als emotionele intelligentie en competenties naar niet onmiddellijk rationele kenmerken van intelligent functioneren.

Informatie wordt omschreven als een reeks vormen van uitwisseling van een organisme - hier de cliënt - met de interne en/of externe omgeving. Ze brengen ons nader tot relevante kenmerken van mensen en objecten. Bij de eersten ligt de nadruk op bewustzijn waar de psychologie zich aan vertild heeft maar ze kan er niet omheen.

Informatie is in ook in andere wetenschappen een kernbegrip. Dat gaat van fysica, scheikunde, levenswetenschappen, neuro- en sociale wetenschappen tot sociologie, antropologie en taalkunde. Materie, levende organismen en ons bewustzijn kunnen als informatieverwerkende systemen opgevat worden. Een dergelijke abstracte kwalificatie moet voor elk vak en elk onderzoek concreet uitgewerkt worden. Het is ermee gesteld als met de concepten uit de evolutie (bijvoorbeeld natuurlijke selectie en overleven) en uit de natuurkunde (bijvoorbeeld zwaartekracht en energie). Ze passen zo ongeveer op alles. Informatie is te beschouwen als een structuur die in open verbinding staat met andere structuren (grammatica's, een aantal verbonden machines, computers), betekenissen (semantiek: ook verschillende perspectieven op hetzelfde verschijnsel) en consequenties voor handelen (pragmatiek met regels, de maxims van Grice). Informatie gaat in de natuurkunde bijvoorbeeld gepaard met een statistische benadering in de thermodynamica en kwantummechanica. Het DNA met zijn code is ook op te vatten als een informatieverwerkend systeem.

In de psychologie verwijst informatieverwerking naar onze uitwisseling met objecten en personen en naar mechanismen die daarbij betrokken zijn zoals waarnemingsgewoonten (of -wetten, geheugen en allerlei verwerkingsregels van 'heuristics' en 'biases' tot eenvoudige vuistregels, logica en wetenschappelijke methodologie. Drie bronnen (impliciete, expliciete en alternatieve theorieën) drie oriëntaties (individuele verschillen, ontwikkeling en sociale context), operationalisatie en meten en instrumentatie vormen het kader voor de bespreking

van cognitie en intelligentie inclusief geschiktheden: verworven vaardigheden op een domein en prestaties. Dat is wat we geleerd hebben (Figuur 1).



Figuur 1: Drie elementen en oriëntaties voor de psychologische diagnose van intelligentie en cognitie. Modeleren omvat hier ook operationalisatie: empirisch specificeren van modellen en theoretisch constructen en de activiteit van het meten. Het laatste resulteert in categorieën, ordinale- en intervalschalen of een aantal vectoren, factoren of dimensies. Bij meten horen ook de gevolgen in de vorm van instrumenten, zoals tests en vragenlijsten en andere methoden van diagnosticeren.

In de twee millennia omvattende westerse wijsbegeerte zijn begrippen die op intelligentie (geschiktheid, prestatie) en cognitie lijken al geanalyseerd. Dit is de sub-discipline van de epistemologie en die gaat over wat we (kunnen) kennen van de wereld van objecten en subjecten. Wat is de inhoud van deze begrippen in de psychologie? Hoe vatten leken kennen en intelligentie op? Welke expliciete theorievorming vult onze tekstboeken? Welke alternatieve opvattingen zijn er? Wat ontbreekt volgens sommige auteurs in die tekstboeken? Intelligentietheorie en operationaliseren en meten zijn oorspronkelijk nauw verbonden want de eerste theoretici waren ook psychometrici. Hoe hebben operationaliseren en meten van intelligentie en cognitie in de oriëntaties vorm gekregen? Welke intelligentie-, geschiktheids- en prestatietests worden gebruikt en wat hun is waarde en functie?

I **Geschiedenis en betekenis van kennen, intelligentie, cognitie, geschiktheid en prestatie**

We hebben de onweerstaanbare neiging om objecten, gebeurtenissen en mensen te begrijpen, te verklaren, te beheersen en voor hun doelen en belangen aan te wenden. Intuïtief weten we dat voor moeilijke taken slimme mensen nodig zijn. In de oudheid werden kinderen geadopteerd om zich te verzekeren van een geschikte, intelligente opvolger. Caesar (100-44 BCE) was een adoptiekind. In de Middeleeuwen werden de kloosters gevuld met monniken die óf rijk óf intelligent waren. Dat gaat niet altijd samen. Bij de opkomst van de universiteiten vanaf de 13^{de} eeuw en in de Renaissance vanaf de 15^{de} eeuw werden studenten geselecteerd. We willen de besten, high potentials wat daar ook onder verstaan wordt.

Intelligentie, inzicht, kennen, denken, begrijpen en evalueren zijn cognitieve handelingen die niet voor het eerst door psychologen zijn beschreven. Vanaf de homo sapiens is er gedacht, kennis verzameld en benut. We beginnen bij de Grieken door onze bekendheid met hun geschriften en eindigen bij de moderne wijsbegeerte. De psychologie over denken en kennen sprong in vanaf het eind van de 19^{de} eeuw. De industriële revolutie van de 19^{de} eeuw veroorzaakte een trek naar de steden. Migranten kregen nieuw en ongewoon werk. Degenen die zich gemakkelijk konden aanpassen en snel leerden waren aantrekkelijk voor werkgevers. In het begin van de 20^{ste} eeuw werd er in sommige landen leer- en dienstplicht ingevoerd. Hoe al die leerlingen en jonge soldaten te plaatsen? Intelligentie- en persoonsdiagnostiek werden op grote schaal benut voor selectie en plaatsing.

Hoe is intelligentie en cognitie opgevat in de geschiedenis vanaf de Grieken, via de Middeleeuwen, Descartes, Hume en Locke, Kant tot de moderne wijsbegeerte? Waarom heeft de eind 19^{de} -eeuwse psychologie ons op het spoor gezet van mentale leeftijd en IQ? Waarom zijn we intelligentie en cognitie gaan zien als een vermogen dat ons helpt in te zien hoe de wereld van objecten, gebeurtenissen, (sociale) verschijnselen en wijzelf in elkaar zitten?

1. Intelligentie, cognitie, geschiktheid en prestatie

Intelligentie en cognitie verwijzen naar de prestatie van een persoon. Het gaat niet om zijn mening of opvatting over objecten, gebeurtenissen, verschijnselen en gedrag. Het gaat om 'objectieve' kennis van fysieke en sociale werkelijkheid. Eerder zijn ken- of denkhoudingen vermeld, onder meer het (naïef) realisme, idealisme/rationalisme, empirisme en de pragmatische denkhouding. Elk biedt criteria om te bepalen of een bewering over de werkelijkheid - hier het intelligente gedrag en handelen van een cliënt - waar/vals/nuttig is. Daarbij noemde ik drie standpunten c.q. benaderingen:

- (1) de empirisch-analytische in combinatie met het logisch positivisme. Dit bevat rationalistische en empiristische kenmerken.
- (2) De interpretatieve waaronder constructivistische, interpretatieve, metaforische en hermeneutische werkwijzen vallen.
- (3) De kritische die betrekkelijk weinig genoemd wordt. Ze kan niet uitgevlakt kan worden omdat daar een emancipatorisch kennisbelang geldt. Dat kan uitwerken als het afwijzen van dominante opvattingen over kennen, denken, cognitie en intelligentie. Het kan ook resulteren in het benutten van deze kennis om de cliënt te bevrijden van in de samenleving heersende vooroordelen over zijn vermogens en kansen.

Epistemologisch/epistemisch is een uitdrukking voor kenmerken van een leer die theorie, (geordende begrippen, constructen, formules) verbindt met iets buiten die theorie. In het geval van diagnostiek is de verbinding tussen het kennisbestand van de psychologie en de vraag/het probleem van de cliënt over zijn gedrag. We weten spontaan en uit onderzoek dat we fouten kunnen maken bij het leggen van de verbinding. Dat wijten we aan onze bedrieglijke klinische blik, gebrekkige informatieverwerking, vertekeningen en vooringenomenheden. De methodologie wapent ons tegen die gebreken door fouten te onderkennen, op te sporen en er voor te corrigeren.

Ons kennen en denken is feilbaar maar tegelijkertijd nemen we aan dat ons verstand, bewustzijn, geest, kennen betrokken is bij en nodig is voor het verwerven van geldige kennis over objecten, verschijnselen en mensen. In de psychologie hebben intelligentie en cognitie de connotatie van alert en snel. We nemen waar, herkennen en verzinnen betekenissen bij gebeurtenissen en gedragingen. We spreken, overleggen en communiceren daarover met anderen om een gemeenschappelijke wereld een gedeeld referentiekader te scheppen.

Intelligentie wordt wisselend gedefinieerd. Je kunt de nadruk leggen op uiteenlopende kenmerken van kennen, leren, informatieverwerken en het benutten van ons denken en bewustzijn. Neisser (1996) heeft in overleg met collega's een middenweg gekozen. Ze deden dat naar aanleiding van de zoveelste discussie over het al of niet aangeboren zijn van intelligentie. Dit leidde tot deze omschrijving:

Intelligentie is een algemeen vermogen en bevat de capaciteit tot redeneren, plannen, problemen oplossen, abstract denken, complexe ideeën begrijpen en snel leren van ervaring.

Ze sluit niet in: uit het hoofd leren of specifieke schoolse vaardigheden, zoals leren lezen en schrijven. Het gaat niet om handig zijn met tests ('test wise'). Het is een breed en diep vermogen dat verwijst naar inzicht in en begrip van het dagelijks leven. Ze helpt beslissen wat gedaan moet worden. Ze verschilt van creativiteit, karakter, persoonlijkheid en andere trekken'.

Deze bevat het IQ en sluit ook cognitie, ontwikkeling en gevoeligheid voor training en scholing in. Intelligentie is een kenmerk van een levend, dynamisch en open organisme. Vooral bij de *homo sapiens*. Ze staat ze niet los van leeftijd want intelligentie/cognitie ontwikkelt zich en is gevoelig voor de sociale context. Gottfredson (1977, p. 13) gaf onafhankelijk van Neisser et al. een *mainstream* omschrijving en komt bijna op hetzelfde uit:

Intelligence involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience. It is not merely learning from a book; a narrow academic skill or test-taking smarts. Rather it reflects a broader and deeper capacity for comprehending our surroundings, 'catching on', 'making sense of things', or 'figuring out what to do.

Cognitie wordt vaak genoemd in combinatie met ontwikkeling. Het begrip verwijst naar het gedurende de ontwikkeling leren kennen van de fysische en sociale wereld. Naast cognitieve ontwikkeling gaat cognitie over processen van informatieverwerking. Dit betreft onder meer de experimentele studie naar het effect van manipuleren van kenmerken van stimuli en situaties op ons waarnemen, onthouden, bewerken en verwerken (*updating, shifting, inhibiting*) van informatie over de fysische wereld en over het gedrag van mensen.

Geschiktheid (aptitude) verwijst naar de capaciteit, het vermogen om op een bepaald domein te presteren na doelgerichte training. Het bepalen van geschiktheid is begonnen bij selectie voor het leger en de universiteit en daarna voor het beslissen over het aannemen van sollicitanten. Het onderscheid tussen vloeibare (*fluid*) Gf en uitgekristalliseerde (*crystallized*) intelligentie Gc van Horn en Cattell is gebruikt in het kader van geschiktheidstests. Dit herinnert aan het alfa-bèta onderscheid dat bij persoonskenmerken gemaakt wordt. Ze zijn elementen van Thurstones *Primary Mental Abilities*. Deze worden zowel intelligentie- als geschiktheidsfactoren genoemd. Intelligentie en geschiktheid liggen met andere woorden niet ver uit elkaar. Ze kunnen conceptueel echter onderscheiden worden. IQ is een algemene vaardigheid (intelligentie, de G-factor van Spearman). Geschiktheden gaan over het vermogen om kennis te verwerven op een gebied, bijvoorbeeld op basis van ruimtelijk, technisch en verbaal inzicht.

Creativiteit Gf speelt een rol in het onderscheid tussen intelligentie en *creativiteit*. Er wordt een beperkte overlap verondersteld. Voorbeelden:

(1) Een meta-analyse door Kim (2005) levert een gemiddelde correlatie over studies van $r = .17$ tussen intelligentie en creativiteit op.

(2) Silvia en Beaty (2012) laten zien dat het formuleren van creatieve metaforen, een item uit creativiteitstests, gecorreleerd is met de Gf: de gestandaardiseerde bèta = $.49$. Een latent variabele model toonde dat 24% van de variantie in metafoor-kwaliteit 'verklaard' wordt door de Gf. Dat is een substantiële samenhang van $r = .49$.

(3) Benedek et al. (2014) onderscheiden de executieve functies *updating*, *shifting* en *inhibition* en gebruiken ze om de relatie tussen intelligentie (Gf) en creativiteit (divergent denken) te duiden. Gf werd gemiddeld tot goed voorspeld door *updating* en niet door *shifting* en *inhibition*. Creativiteit werd door *updating* en *inhibition* voorspeld maar niet door *shifting*. De gedeelde variantie van intelligentie en creativiteit werd toegeschreven aan *updating* en aan de BF factor Openheid.

Er is kortom empirisch enige overlap tussen IQ en creativiteit. De samenhang (β) varieert tussen $.17$ en $.49$. Dat is volgens Cohens vuistregel het bereik tussen bescheiden, via gemiddeld tot groot. Verderop volgt een uitwerking met een drempelhypothese over de IQ-creativiteit relatie. Hoeveel IQ heb je nodig om creatief te kunnen zijn?

Het onderscheid tussen intelligentie en geschiktheid hangt historisch samen met verschil tussen een Britse en een Amerikaanse opvatting. De Brit Spearman koos voor de G-factor die dicht bij het idee van het IQ komt. De Amerikaan Thurstone prefereerde gedifferentieerde geschiktheidsfactoren in zijn land met een gevarieerde bevolking als gevolg van immigratie.

Prestatie (achievement) verwijst naar wat aan kennis en vaardigheden verworven is na het volgen van een curriculum. *School Achievement Tests (SATs)* en de Cito-toetsen zijn voorbeelden van instrumenten die gemaakt zijn om school- en beroepsprestaties te meten. De drie stellen een persoon in staat om te laten zien wat hij doorheeft, begrijpt en aankan. Ik beschouw intelligentie en geschiktheid als een twee-eiige tweeling. Ze zijn niet inwisselbaar maar familie en lijken op elkaar. Ze willen echter af en toe graag van elkaar verschillen.

Samenvatting en conclusie

Al lang voor de psychologen intelligentie en prestatie gingen onderzoeken waren er opvattingen over de relatie tussen wat we zeggen over objecten en gedrag en wat die objecten en die gedragingen zijn. Dat is een epistemische relatie die in de psychologie empirisch-analytisch, dat wil zeggen met nadruk op voorkomen van fouten, uitgelegd wordt. Daarnaast is een hermeneutische en kritische epistemologie mogelijk. We weten dat we feilbaar zijn. Methodologie is een stelsel werkwijzen en voorschriften om fouten op te sporen en te controleren. In de psychologie verwijzen intelligentie, geschiktheid en prestatie achtereenvolgens naar een algemeen vermogen om problemen op te lossen, een aantal specifieke capaciteiten om te leren van training/oefening op een bepaald gebied en naar verworven kennis bijvoorbeeld na het volgen van een curriculum. Intelligentie en geschiktheid zijn verwante maar te onderscheiden concepten. Meestal sluit intelligentie

aliertheid en snelheid in en het in staat zijn om te leren van de context en zich te ontwikkelen. Cognitie wordt gebruikt in combinatie met cognitieve ontwikkeling en informatieverwerking. Het gaat over het verwerven van kennis over de fysische en sociale werkelijkheid gedurende de ontwikkeling en van kennis over kenmerken van stimuli van buiten en de verwerking en opslag daarvan. Intelligentie en creativiteit zijn verbonden, gecorreleerd maar vallen niet samen.

2. Geschiedenis van kennis, intelligentie- en cognitiebegrissen

De Duitse geheugenonderzoeker Hermann Ebbinghaus (1835-1909) zei eens dat de psychologie een lang verleden en korte geschiedenis heeft. Dat geldt ook voor kennen en intelligentie. Het concept is uitgevonden als meetbare individuele verschillen variabele in de 19^{de} eeuw. Als onderdeel van de westerse wijsbegeerte en wetenschap heeft ze een lang verleden. Ze heeft betrekking op mogelijkheden en beperkingen van ons kennen: de epistemologie (Kenny, 2001). De eenheid, de bouwsteen van kennen in de filosofie is vaak het begrip, concept. Ik ga hieronder zonder scrupule met zevenmijlslaarzen door de westerse geschiedenis van kennen heen.

Het pre-Socratisch Griekenland De oorsprong van reflectie op ons kennen ligt voor ons in het antieke Griekenland. De eerste wijsgeren van wie geschriften ons bekend zijn, of waar we ons toe beperkt hebben, hielden zich bezig met verklaren opgevat als de *oorsprong* van al wat is. Is het water, lucht, zijn het getallen, atomen, of een combinatie? De pre-Socratische tijd was een tijd met enige rust. Pythagoras had bijvoorbeeld een harmonieuze wereld voor ogen, waarin het getal, het decimale stelsel, de uitdrukking voor de aanwezige ordening was. Daarna kwam een periode van twijfel. Het idee won terrein dat er geen zekere kennis bestond en 'de mens de maat van alle dingen' was. Sofisten reisden de Griekse steden en eilanden af en leerden mensen tegen betaling zo te argumenteren dat ze gesprekken en discussies op de markt, in de handel en voor de rechtbank konden winnen. Het lijkt op de moderne communicatietraining, en dat is business. In Nederland zijn er ongeveer 18.000 titels over communicatie beschikbaar: 'Vandaag besteld, morgen thuisbezorgd!'.

Plato (427-347 BCE) kritiseerde de sofisten en zette tegenover de twijfel *absolute ideeën* over Eenheid, Waarheid, Goedheid en Schoonheid. We kunnen daar echter alleen de schaduwen van waarnemen. Al zijn we gebrekkige kenners, we hebben geen twijfel over het bestaan van de Ideeën. Kennis is herinnering: anamnesis. Het geheugen is altijd selectief en daardoor schieten we tekort in het verinnerlijken van die absolute Ideeën. In het dagelijks leven was Plato niet alleen de filosoof die mooie ideeën koesterde en voorschriften formuleerde over hoe goed te leven en te besturen. Hij streefde ook naar politieke macht, werd slaaf, maar werd vrijgekocht en verbrandde geschriften van zijn tegenstanders, toen hij weer invloed had. Hij verzette zich tegen het schrift, zoals sommigen zich nu tegen Google, Wikipedia, Twitter, enzovoort verzetten: (Plato, vertaling, 1980, IV, p. 476)

‘...het schrift zal in de ziel van hen die het leren, vergetelheid doen ontstaan, doordat ze zullen vergeten hun geheugen te oefenen. Door hun vertrouwen in het schrift zullen ze immers het middel om zich iets te herinneren buiten zichzelf gaan zoeken, in vreemde lettertekens, in plaats van in zichzelf, door hun geheugen in werking te stellen’.

Geschriften tot zich nemen hield volgens Plato in dat je een veelweter en manipulator wordt zonder zich iets van de Ideeën te herinneren en daardoor de wereld te begrijpen. Zijn denken werd omarmd door het Christendom. Het denken paste voor een deel. Het lichaam is de kerker van de ziel en er is een eeuwig bron van al wat is was en zal zijn: God: het Ene, ware, goede en schone. Het menselijk kennen schiet tekort. Dat gelovigen niet konden lezen, was geen probleem voor de clerus. Dat werd pas in de Reformatie doorbroken: Luther vertaalde de Bijbel in de volkstaal. Heidegger (1927/1962; Nederlandse vertaling 1980, p. 220) zou in de 20^{ste} eeuw zijn variant van dat tekortschieten als Zijnsvergetenheid opvoeren. Hij had ook nog een aversie tegen gepraat (*Das Gerede*). Van Plato mocht je al niet schrijven of schilderen en tekenen. Er mag veel niet, en dat is tot nu gebleven.

‘...en omdat het gepraat de primaire zijnsbetrekking tot het gesproken zijnde heeft verloren, respectievelijk nooit heeft gehad, deelt het zich niet mee op de wijze van een oorspronkelijke toe-eigening van dit zijnde, maar via het doorvertellen en napraten’ (onderlijning: Heidegger).

*Aristoteles (384-322 BCE) vatte menselijk kennen en intelligentie op als *spontane actieve krachten*. Het verstand vormt concepten, constructen door het algemene uit het individuele en bijzondere te halen. Materie zonder vorm is chaos, maar heeft de potentie om in de greep te komen van de vorm. Ziel, geest is vorm en maakt materie tot iets intelligibels, dat wil zeggen iets dat toegankelijk wordt en is door en voor ons verstand. Hij was een nuchter man en er zich van bewust dat kennis ook afhangt van interesses en belangen van de kenner, de gelegenheid en het streven naar macht. Constructen die we van de wereld vormen zijn een compromis tussen elementen uit de objectieve wereld en wat bedacht, uitgevonden of gegeven is door of met ons verstand. De ziel is het levensprincipe van wat beweegt: planten, dieren en mensen. De laatste tonen naast kennen emoties als verlangen, wraak, overheersing, vreugde. De basis van zijn ethiek is het bewaren van het midden: ‘niets té’, ‘de gulden middenweg’ en zelfcontrole.*

Dit is in het Christendom overgenomen en er is een hang naar extremen aan toegevoegd, denk aan het verheerlijken van ascese. Die belangstelling zien we in de psychologie terug in de aandacht en subsidie die extreem gedrag (ernstig lijden, agressie en cognitief presteren) trekt. Het Christendom deelt deze belangstelling met andere doctrines en ideologieën.

Middeleeuwen Thomas van Aquino (1224-1274) is waarschijnlijk de grootste christelijke denker ooit. Hij combineerde het werk Plato, Aristoteles, Augustinus, de Arabieren Averroes

en Avicenna en vele anderen - van de laatste drie vooral hun commentaren op de christelijke Bijbel - tot een theologie die nu nog onderwezen en bestudeerd wordt. De zintuigen (Aristoteles) spelen een belangrijke rol in het kennen van de wereld en God. Eén van zijn vijf 'godsbewijzen' (vijf wegen: *Quinque Viae*) berust op de oneindige eenheid, waarheid, goedheid en schoonheid van de wereld (Plato). De overige ontleende hij aan Aristoteles: dat van de Onbewogen Beweger, de Onveroorzaakte oorzaak, het uit zich volstrekt Noodzakelijke, de vierde is de Platoonse en de vijfde weg is nu nog populair: de intellectuele Ordenaar der naturen (zie Peters, 1957, pp. 451-461). Zo'n idee of ideaal van schoonheid en orde is kennelijk nooit ver weg. Als er sprake is van een 'Ontwerper' - de vijfde weg - bedoelen ze daar niet iemand mee die in Delft bouwkunde gestudeerd heeft of het heil heeft gezocht in de Eindhovense *Design Academy*.

Volgens Thomas werken zintuigen en het actieve verstand samen om algemene constructen te vormen over objecten en verschijnselen in de wereld. Een *concept* is zijnde van het verstand (de rede) met een fundament in de werkelijkheid die los van ons verstand bestaat. Het is een abstractie van de zichtbare wereld die bestaat uit de massa aan objecten en mensen, omschreven door Wittgenstein als 'dat wat het geval is'. De abstractie bevat de kern of het wezen van een object. Intelligentie is 'in de objecten lezen': inzicht. Het verstand 'ziet' geen vervormde schaduwen (Plato) maar heeft toegang tot de kern, de essentie (Aristoteles). Intelligentie is bij Thomas het vermogen om te onderscheiden op basis van reële verschillen en te combineren op basis van werkelijke gelijkenissen. Het concept 'invariantie': hetzelfde zien in objecten die verschillend lijken, het abstraheren van de kern, is later door Piaget bij baby's en kinderen onderzocht. Zij 'leren te weten' dat een object nog bestaat ook al is het niet meer zichtbaar (objectpermanentie) en dat een object niet verandert ook al wijzigt de uiterlijke verschijningsvorm (conservatie).

Als christen moest Thomas de aardse mens verzoenen met eeuwige goedheid, schoonheid en waarheid. De mens had volgens hem individueel verschillend en beperkt deel aan die goedheid, waarheid en schoonheid. Hij is immers een schepsel uit die eeuwige waarheid: God. Men kan Hem (m/v/o) opvatten als een licht dat alles verlicht maar dat zelf niet gezien kan worden.

Thomas leefde in een onrustige tijd en moest vanuit Italië naar Parijs vluchten. Dat verhinderde hem niet gedisciplineerd en systematisch te schrijven. Door de in die tijd voorgeschreven bureaucratische stijl en de verplichte opbouw van geschriften is het taaie leesstof. Je redeneert altijd naar een doel, vooral het kloppend maken van een zinsnede uit de Bijbel. Dat spreekt ons niet meer aan in zijn hoofdwerk: een samenvatting van de theologie. De inhoud blijft de moeite waard en wordt nog uitgelegd en becommentarieerd (zie bijvoorbeeld Peters': *Metaphysica, een systematische inleiding*, 1957).

Filosofie van de nieuwe tijd Descartes (1569-1650) wordt de filosoof van de nieuwe tijd genoemd. Hij kwam op het probleem van het kennen terug. Hij moest niet veel hebben van de oude scholastieke filosofen na Thomas. Je kon het immers niet zo gek bedenken of een of andere filosoof had daar wel een uitspraak over gedaan, bijvoorbeeld hoeveel engelen er

op de punt van een naald konden zitten. Doordat in de 16^{de} eeuw schepen uitvoeren en er nieuwe continenten ontdekt werden brak bovendien het inzicht door dat kennen cultureel bepaald en dus relatief is. Descartes probeerde zich te onttrekken aan het stellen van zinloze vragen en aan een hulpeloos en hopeloos relativisme en ging op zoek naar *zekerheid*. We kunnen aan alles twijfelen zelfs aan het bestaan van de wereld want dat kan een illusie zijn, maar we kunnen de twijfel zélf niet betwijfelen. Het 'Ik denk dus ik ben': *Cogito ergo sum* is een onontkoombare intuïtie. Kennen en waarheidsvinding worden daardoor een aangelegenheid van het bewuste denken, de ratio en bergt de mogelijkheid tot zekerheid in zich.

Filosofen moeten heldere en duidelijke ideeën formuleren. Hij vond die helderheid van de wereld van objecten en verschijnselen in de kenmerken lengte, breedte, hoogte en beweging. Het bewustzijn beschikt over ingeboren, met de geboorte gegeven ideeën. De belangrijkste zijn:

- (a) Een oneindige substantie van het meest volkomen zijn dat uit zichzelf bestaat en God heet. Daarmee redde hij zich uit een materialistisch atheïsme dat hem de kop gekost zou hebben.
- (b) Uitgebreidheid dat wil zeggen van de eerste drie kenmerken van lichamen en voorwerpen.
- (c) Niet-uitgebreidheid van de denkende substantie, de ziel of het bewustzijn. Hij noemt daarnaast nog andere ingeboren ideeën, zoals getal, tijd, plaats en vorm of gestalte.

Enkele daarvan keren in de 19^{de} eeuw terug in het werk van Kant als synthetische oordelen a priori: tijd, ruimte en oorzaak.

Met Descartes is een rationalistische traditie in het Westerse denken gevormd die tot vandaag voortduurt. De psychologie vertoont er sporen van, vooral in haar empirisch-analytische, logisch positivistische kenhouding. Studies over het getalbegrip laten bijvoorbeeld zien dat baby's een notie van 'tweeheid' hebben. Ze reageren (hartslag, kijkrespons) anders op twee willekeurige voorwerpen als op drie willekeurige andere voorwerpen. Bij de interpretatie van dit verschijnsel wordt soms aan een aangeboren idee van hoeveelheid gedacht (Gelman & Williams, 1998). Chomsky (1978) heeft laten zien dat het onmogelijk is om alle zinnen uit de moedertaal op behavioristische wijze als stimulus-respons verbindingen te leren. Er moet een *Language Acquisition Device* (LAD) verondersteld worden dat ons met de geboorte gegeven is. Een evolutionair psycholoog zal zeggen dat het een voorlopig punt in een evolutie is waarin inderdaad taal geleidelijk een ommekeer teweeg heeft gebracht. Zo is er iets ontstaan dat we bewustzijn noemen.

De Franse en wat later Duitse wijsgerige tradities zijn respectievelijk rationalistisch en idealistisch. Bij beide heeft de geest, het verstand, het idee voorrang bij het beschrijven en verklaren van de wereld en de mens. In de Franse rationalistische traditie verdwijnt de twijfel door de zekerheid van de ingeboren ideeën. Dat houdt in dat je moet steunen op je eigen bewuste verstand. In de Duitse idealistische traditie heeft het idee voorrang op de werkelijkheid. In de opvatting van Hegel (19^{de} eeuw, Nederlandse vertaling: J. Hahn, 1980)

wordt de hele geschiedenis voorgesteld als het door dialectiek geleidelijk verschijnen van de Geest: de *Phaenomenologie des Geistes*.

Brits empirisme John Locke (1623-1704) gaf de nieuwe filosofie een radicale wending. Kenleer is onder meer door zijn werk een centraal onderdeel van de wijsbegeerte geworden. Hij was een prototype van een klassieke Engelse wijsgeer: gericht op ervaring, met werkelijkheidszin en weerszin tegen speculatie. Hij was gezegend met een afgewogen oordeel en toonde evenwicht tussen conservatief en progressief denken en handelen. Hij stelde vast dat we bij kinderen niets merken van ingeboren ideeën. Kennis komt volgens hem tot stand door *zintuiglijke ervaring: sensations* en door waarneming van onszelf: *reflection*. Die twee samen leiden tot voorstellingen: *ideas* die met elkaar verbonden raken door *associations*. Hiermee is afgerekend met de inbreng en voorrang van het werkzame verstand en ingeboren ideeën. Hoe die verbindingen tot stand komen en hoe we tot gemeenschappelijke voorstellingen, *ideas* komen is een probleem. We hebben immers unieke ervaringen en kunnen we nooit de ervaring en het bewustzijn van een ander doorgronden.

David Hume (1711-1776) beweerde dat de associaties niet willekeurig zijn. Ze zijn wetmatig, raken verbonden door gelijkenis, eenzelfde oorzaak en dezelfde plaats en tijd. We kunnen tot een gemeenschappelijk beeld van de fysische werkelijkheid komen omdat we in een wereld leven met dezelfde natuurwetten: Newtoniaanse universele wetten. Dit past ook op de sociale werkelijkheid. Het doet denken aan de behavioristische leerwet van contiguiteit, dat wil zeggen dat een stimulus en een respons verbonden raken omdat ze in tijd en plaats nabij zijn. Het herinnert ook aan de Gestaltwetten die van nature orde in de waarneming van particulariteiten aanbrengen.

Hobbes (1651/1992), Locke en Hume (Honderich, 2005) formuleerden de *empiristische traditie* waarin psychologie en diagnostiek voor een deel naast rationalisme ook stevig geworteld zijn. In die traditie lopen we achter de feiten aan en proberen we weinig of niets aan de feiten op te leggen. Er is de pretentie van het verkrijgen van objectieve kennis en deze kennis heeft autoriteit, dat wil zeggen je kunt er niet omheen en moet ze accepteren. Hume was mild over mensen die daar moeite mee hadden. Hij was gematigd in wetenschappelijke en religieuze debatten waaraan volgens hem veel gelijk hebbers deelnamen. Hobbes (1992, p. 120) dacht dat wetenschap een einde zou moeten maken aan de bloedige religieuze oorlogen en stelde een 'social contract' voor om het moorden te stoppen. Zijn beroemde, controversiële en absolutistische regel luidt:

'I authorize and give up my Right of Governing myself, to this Man, or to this Assembly, on this condition, that thou give up thy Right to him, and authorize all his Actions in like Manner. This done, the Multitude so united in one Person, is called COMMON WEALTH'.

Hij was er van overtuigd dat om aan oorlog – de oorlog van iedereen tegen iedereen - en angst te ontsnappen de macht aan een sterke staat gegeven moest worden.

Het *Esse est percipi* (Zijn is waargenomen worden) van de Engelse empiristen kent een moderne variant in de opvatting dat de mens niet alleen product of consument is. Hij bestaat ook in wat *anderen* van hem vinden en in hem waarderen. Ik ben wat ik ben in de ogen van anderen, denk aan de socioloog Goffman en de filosoof Sartre.

Verlichting De continentale en Britse tradities van rationalisme/idealisme en empirisme zijn door *Immanuel Kant* (1724-1804) verbonden. Hij vroeg zich af wat we kunnen weten en wat niet. Wat en hoeveel kan het verstand (*Vernunft*) los van de ervaring kennen? Is er geen kennis die al in een construct of in ons zit opgesloten, los van enige ervaring? Dat de cirkel rond en lichamen en objecten uitgebreid zijn, is een gegeven vooraf: een *apriori*. Het is niet een resultaat van onze waarneming van veel cirkels of het zien van veel objecten. Dat oordeel over de rondheid en uitgebreidheid is geen empirisch maar een analytisch oordeel. Andere oordelen over kenmerken vellen we achteraf na waarneming en studie: *aposteriori*. Bijvoorbeeld: deze objecten zijn zwaar, hebben een kleur. Om van de toevallige associaties van Hume en de ingeboren ideeën van Descartes af te komen ging Kant op zoek naar samengestelde (synthetische) oordelen vooraf (apriori). Naar zijn overtuiging komt er iets uit onszelf, uit ons verstand, ons intellect dat er al is vóór alle ervaring. Voorbeelden van deze oordelen zijn *ruimte, tijd en causaliteit*. Een causaal verband wordt volgens hem dus niet alleen uit empirisch onderzoek afgeleid. Ons verstand kan immers niet onder dat synthetisch oordeel apriori uit. Het 'moet' objecten, gebeurtenissen en gedragingen als een oorzaak-gevolg keten opvatten. Daarnaast kunnen we informatie over de werkelijkheid verkrijgen door empirisch onderzoek, bijvoorbeeld door kenmerken van objecten en mensen vast te stellen.

Volgens Kant is dit een copernicaanse revolutie. Het kennen richt zich immers niet naar de voorwerpen. Deze 'richten' zich naar ons kenners. We kennen de dingen op zichzelf (*an sich*) niet. We kennen ze in het licht van onze synthetische oordelen apriori en kunnen daarnaast onderzoek doen en geldige synthetische oordelen aposteriori over verschijnselen en mensen vormen en uitspreken. We hebben geen kennis van het wezen van de objecten, los van onze 'voorkennis'. Kennen en weten zijn bij Kant relationeel dat wil zeggen met een eigen inbreng waaraan je je niet kunt onttrekken en met een inbreng van de objecten met hun kenmerken.

Het bestaan van oordelen vooraf: voor-oordelen is uitgezocht in de psychologie. In 1955 vatten Cronbach en Taft al studies samen met de boodschap dat ons oordeel over en waarneming van het gedrag van anderen (*person perception*) bepaald is door onze opvattingen, stereotypen en belangen. En minder door feitelijke gedragingen van anderen. Dit bevestigt wel niet op epistemologisch maar op empirisch niveau dat ons oordeel niet aan de objecten en personen gewijd is maar het voor een deel onze 'voor-oordelen' steunt ('voor' kan gelezen worden als apriori) en onze belangen weerspiegelt.

De Verlichting was een optimistische periode. Kant dacht zelfs dat het geen doel diende om nog oorlog te voeren en heeft een korte tekst geschreven waarin hij wilde rationeel wilde laten zien dat oorlogen tot het verleden zouden behoren. De titel van zijn boekje was

ironisch bedoeld: *Zum ewigen Frieden*: de richtingwijzer naar het kerkhof. Hij schreef *voorwaarden* op waaronder er altijd vrede zou kunnen zijn. Het kleine werk heeft een rol gespeeld in het schrijven de Verklaring van de Rechten van Mens. Hobbes dacht van *niet* en wilde daarom iedereen onder een dwingende regel brengen. Wat eens 'verlicht' was zouden sommigen nu - en Hobbes al veel eerder - *naïef* noemen. Moderne schrijvers en denkers verklaren dan ook vaak de Verlichting dood, bijvoorbeeld de Franse auteur Houellebecq.

Wijsbegeerte van de 20^{ste} eeuw Deze is niet meer in te delen volgens het schema Rationalisme, Empirisme, Idealisme, Verlichting, enzovoort. Ze bevat nu thema's die aansluiten bij maatschappelijke vragen. Evenals in de psychologie is in de filosofie de tijd van de grote systemen, ideologieën, theorieën voorbij. De discussie over wat we kunnen kennen en weten lijkt afgeschaft of is irrelevant geworden door technologische successen. Sommige moderne filosofen verzetten zich daartegen want technische kennis levert biedt geen inzicht in het menselijk bestaan. Deze filosofen beïnvloeden nauwelijks de hoofdstroom in de psychologie. Zij benutten filosofie i.c. kenleer of wijsgerige antropologie niet voor theorievorming en onderzoek. Het zijn voor hen losse compartimenten.

Een enkeling heeft een stempel gedrukt op de 20^{ste} eeuwse filosofie, bijvoorbeeld Heidegger. Hij spreekt in zijn hoofdwerk *Sein und Zeit* (1927, 1986) over *Existentialen*: onontkoombare zijnswijzen, zoals het 'in-de-wereld-zijn'; 'het met-anderen zijn'; 'het zelf-zijn' en 'het ten-dode-zijn'. Deze drukken de tijdelijke, ruimtelijke en sociale aard van het menselijk bestaan uit en doen herinneren aan Kants synthetische oordelen apriori. Heidegger zegt immers dat gebeurtenissen, voorwerpen, de wereld, de medemens geen bestaan op zichzelf hebben maar slechts in relatie tot elkaar. Bovendien komt in deze *Existentialen* tot uitdrukking dat de mens een individu, een zelf is. Hij is historisch dat wil zeggen dat hij een ontwikkeling door de tijd heen doormaakt. De gebeurtenissen, verschijnselen, de wereld en medemens vormen het brute, ruwe materiaal, de ruwe stof die door de kennende geest gevormd wordt.

De filosofie maakte aannemelijk dat we vastzitten aan onze geest, ons verstand en lichamelijke en daarmee objecten en verschijnselen tegemoet treden. Descartes maakte er een dualisme van en Kant probeerde ze te verbinden. Een in de psychologie vergeten poging tot synthese/verzoening is van Merleau-Ponty (1945; Engelse uitgave, 1958). Het lichaam is geen puur object en geen puur subject maar een lichaam-subject: *corps sujet*.

Wat biedt dit voor de diagnose van een cliënt? Het wijsgerig verleden van de diagnostiek gaat *niet* over theorievorming van individuele verschillen, ontwikkeling en sociale context. Het heeft echter wel sporen in de psychologie- en diagnostiekbeoefening nagelaten. Galton maakte tests vanuit zijn empiristische opvatting dat de zintuigen de poorten zijn waardoor kennis binnenkomt. Dat werd meten, selectie bij de poort, bijvoorbeeld visuele scherppte. Het Britse empirisme van Galton aangevuld met Amerikaans pragmatisme (onder meer James, 1907) domineren de psychologie. Wat minder prominent maar herkenbaar is het denken over de verhouding geest-lichaam in het werk van Brunswik en Gigerenzer. Het

multi-attribuut-nut-model vooronderstelt een Cartesiaans rationalistische beslisser gecombineerd met een behavioristische, economische hedonist, die de goede meest bevredigende kenmerken selecteert en de optie met de grootste winst kiest. In de ontwikkelingspsychologie van Piaget worden Kantiaanse vraagstukken over ruimte, tijd, oorzaak voorgelegd aan kinderen en adolescenten met behulp van ingenieuze proefjes.

Valt er iets te leren van deze onvolledige geschiedenis van het kennen? De Kantiaanse inzichten hebben onder meer als consequentie dat we erop letten wat ons verstand (ons lichaam-subject) meebrengt bij begrijpen en verklaren van individuele gedragsverschillen, cognities en emoties, hun ontwikkeling en de effecten van de context. Aandacht voor werkzaamheid van ons verstand en de daaruit voortvloeiende analyses treffen we een enkele keer aan bij het bespreken van de filosofische herkomst van het behaviorisme of van de informatieverwerking. We leren ook dat in de geschiedenis rustige perioden van denken over het menselijk kennen en bestaan worden afgewisseld met twijfels. Kuhn (1962) noemde dit de rust van het heersen van een *paradigma*.

De grenzen van het kennen worden van tijd tot tijd opnieuw uitgevonden en geherdefinieerd: een nieuw paradigma. Meestal gaan er socio-culturele veranderingen aan vooraf of er mee gepaard. Dit geldt ook voor de psychologie: Galton was geïnteresseerd in het menselijk genie. Het is mogelijk dat hij de positie van de Britse hogere klassen wilde onderzoeken en bevestigen in de tijd van de opkomst van industriële ondernemers. Dat was een nieuwe klasse burgers. De steden stoomden vol met arme migranten. Binet beantwoordde een vraag uit de samenleving met een intelligentietest die een mengsel is van rationalistische en pragmatische uitgangspunten. Deze interesse en vraag zorgden voor het belang van het vaststellen van individuele verschillen in intelligentie. Dat was geen thema van de filosofie van het kennen.

De VS werden rond 1870 het rijkste land ter wereld. Ze moesten een weg vinden om veel immigranten op te nemen. Er was van meet af aan belangstelling voor individuele verschillen tussen mensen: aan wie heb ik het meest voor werk; van wie krijg ik het minst last als hij hier woont. De rationalistische traditie is bij Binet te herkennen. Hij nam syllogismen op in zijn test en die stammen uit de logica. Hij was ook praktisch: wat moeten kinderen weten om het verplichte onderwijs te volgen en te overleven?

Samenvatting en conclusie

Intelligentie, cognitie, geschiktheden en prestaties hebben te maken met het kennen van de fysische en sociale wereld. Hoe komen we aan die kennis en hoe weten we of die geldig is? Dit zijn vragen die we al in de oudste, ons bekende geschriften aantreffen. Bij Plato is de bron van kennis de *herinnering* aan absolute Ideeën. Volgens Aristoteles en Thomas van Aquino stamt algemene kennis weliswaar uit de waarneming van het concrete maar is er een werkzaam verstand dat de algemene kern, het wezen eruit haalt. In de nieuwe filosofie vanaf Descartes ontstaat het rationalisme als poging tot zekere kennis. Deze berust volgens hem op ingeboren ideeën. Engelse filosofen hebben traditioneel meer vertrouwen in de waarneming en Locke zet de veronderstelling van ingeboren ideeën opzij. Voorstellingen

worden gevormd uit waarnemingen en de reflectie daarop. Kant vraagt zich af wat de kenner onvermijdelijk meebrengt bij het kennen. Zijn synthese is dat de waarneming het eerst is maar die wordt gevormd door de kenner met oordelen die hij al heeft voorafgaand aan iedere waarneming. De daaruit voortvloeiende houding is nuttig voor de beoefening van diagnostiek want elke diagnosticus moet proberen te onderscheiden wat hij aan hypothesen, theorie, gedachten en (voor)oordelen meebrengt in het diagnostisch proces en wat hij aan waarneming van de gedragingen van zijn cliënt ontleent. In de hedendaagse filosofie is er geen aandacht voor grote theorieën. De filosofie van de 20^{ste} eeuw laat niettemin sporen zien van deze vragen. Heidegger beschrijft kenmerken van het menselijk bestaan die even intuïtief onontkoombaar lijken als de *apriori*'s van Kant. Merleau-Ponty neemt Descartes' dualisme van lichaam/object met uitgebreidheid en bewustzijn met ingeboren ideeën als uitgangspunt voor een synthese: op het niveau van het lichaam zijn er immers al intenties, bedoelingen. De socio-culturele condities aan het eind van de 19^{de} en begin 20^{ste} eeuw waren aanleiding om individuele verschillen in cognitie en intelligentie vast te stellen. Daarmee is de wijsgerige oorsprong verlaten want deze verschillen zijn geen onderwerp van de filosofie van het kennen.

3. Reflectie en evaluatie

Het filosofisch verleden van de psychologie toont de dominantie van westerse rationalistische, idealistische, empiristische en pragmatische tradities. Dit blijkt uit de dominante rol van het logisch positivisme en het methodologisch analytisch-empirisme in de psychologie. Er zijn in de psychologie sporen van alle 'ismen' uit de filosofie maar empirisme en rationalisme domineren. Dit is aanleiding voor filosofen om psychologen te verwijten dat ze een zwakke en inconsistente epistemologische basis hebben en een beperkte opvatting van objectiviteit door hun poging/wens om op natuurwetenschap te lijken. Andere epistemologische houdingen, interpretatieve en kritische komen nauwelijks aan bod. Het wijsgerig verleden van diagnostiek van intelligentie is indirect te herkennen bij het vaststellen van *individuele verschillen* in intelligentie, geschiktheid en prestatie. Het gaat in de wijsbegeerte om het kennend subject en dat is een abstractie. Er is wel een relatie met de tijdgeest van de 19^{de} - 20^{ste} eeuw in de tests die Galton en Binet hebben ontworpen. Het doel was niet het kennen los van kennende subjecten te bestuderen maar uit te zoeken of genialiteit overgeërfd werd en schoolsucces voorspeld kon worden.

De ontwikkelingspsychologie van Piaget is niet empiristisch en pragmatisch maar Europees rationalistisch geïnspireerd. Ontwikkeling, tijd, verandering en geschiedenis speelden een rol in zijn denken. Hij wilde in eerste instantie de geschiedenis van de wijsbegeerte bestuderen om de cognitieve ontwikkeling van eenvoudige reflexen tot logische operaties te beschrijven. De gedachte dat de ontogenetische ontwikkeling (ontwikkeling in een mensenleven) de fylogenetische ontwikkeling (ontwikkeling van de menselijke soort) herhaalt, bracht hem tot bestudering van kinderen van 0 tot 12/18 jaar. Zijn onderzoeksthema's sluiten aan bij wijsgerige vraagstukken over inzicht in natuurwetten, morele regels en socio-cognitief handelen. Deze vragen stelt hij aan kinderen aan de hand

van proefjes en situatieschetsen. De kinderen proberen deze op te lossen en geven antwoorden op de vraag *waarom* ze antwoorden als ze antwoorden. Deze kennis is er niet gericht op voorspelling, controle en beslissen. Ze spelen dan ook in de praktijk van de diagnosticus nauwelijks een rol.

De gemanipuleerde context van het klassiek behaviorisme in de gedragsmodificatie speelt in de diagnostiek een beperkte rol. De sociale context in scholing en training staat voor een deel los van hoe kennen in zijn werk gaat. Er wordt vooral een stimulus-respons verbinding voorondersteld bij een organisme dat op bevrediging van behoeften uit is.

De invloed van het wijsgerig verleden op de psychologie is er wel maar selectief en in beperkte mate. Dat geldt ook voor de diagnostiek. Dit kan de suggestie wekken dat intelligentie onderzoek geëmancipeerd is en onafhankelijk van vragen over de verhouding tussen kenner en gekende. In de studie van de cognitieve ontwikkeling (Piaget) speelt de wijsgerige achtergrond wel een rol. Bij de studie van informatieverwerking is dat al minder het geval maar de verhouding tussen kenmerken van stimuli en situaties en die van de waarnemer/kenner stuit uiteraard op de vraag wat de waarnemer meebrengt en wat van objectieve stimuli in de buitenwereld afkomstig is.

De diagnosticus kan gegeven de verschillende kenhoudingen met recht eclecticisch te werk gaan. Diagnostiek van intelligentie is een empirisch vraagstuk geworden. Dit neemt niet weg dat de *kritiek* van filosofische zijde is dat het empirisme beperkt is in het beschrijven en verklaren van kennen en denken en de actieve rol van de kenner onderschat. Er is ook kritiek op de opvatting van objectiviteit, want die valt niet samen met het objectief scoren van antwoorden op testvragen. Het is ook niet het met behulp van formules integreren van gegevens en beslissen tussen opties. Ten slotte valt het niet samen met het toetsen of verschillen tussen condities significant zijn. Het object van onderzoek is een subject dat recht moet worden gedaan. Met andere woorden zoals het is in zijn rijkdom aan verschijningsvormen. De subjectieve belevingswereld van de cliënt hoort daarbij. Het blijft een twistpunt of kennis daarvan betrouwbaar en valide is gegeven de methodologisch empirisch-analytische criteria. En, als dat al het geval is, blijft het voor de diagnosticus open, of deze kennis bijdraagt aan de beantwoording van de vraag of het oplossen van het probleem van zijn cliënt. Onderzoek schat en toetst immers populatieparameters.

II **Impliciete theorieën van intelligentie en prestaties**

Intelligentie en cognitie laten leken niet onberoerd. Ouders hebben graag intelligente kinderen. Datende personen letten meer en meer op gelijkwaardigheid van opleiding en intelligentie. Scholing, opleiding en intelligentie worden beschouwd als voorwaarden om te stijgen op de sociaal-economische ladder.

Welke opvattingen hebben leken over aard en oorzaken van individuele verschillen in intelligentie en cognitie, over de ontwikkeling van cognitie en de invloed van de sociale context? Zijn leken ervan overtuigd dat intelligentie een gave is of geloven ze in de mogelijkheid van het verbeteren van hun intellectuele capaciteiten? Waarom zijn lekenopvattingen relevant voor de diagnosticus? Hoe formuleren cliënten hun vraag of probleem? Is het een verhaal, dat hun 'beliefs', hun opvattingen weerspiegelt? Maakt de officiële theorievorming verschil met wat de man in de straat en wat men achter redactie bureaus van bladen en magazines over intelligentie denkt?

1. Individuele verschillen

Intelligentie interesseert ons. We willen graag intelligent overkomen en doen óf gemakkelijk uitspraken over het IQ (Eysenck reed een auto met als nummerbord: IQ 180; we kijken naar de 'nationale IQ test', of 'de slimste mens') óf we doen er geheimzinnig over. Het concept heeft een functie in het leven van alledag. Als we te maken hebben met een ingewikkeld vraagstuk waarbij we hulp kunnen gebruiken kijken we niet alleen of iemand aardig is maar vooral of hij slim of handig is.

Eerst wordt de structuur en inhoud van opvattingen van leken en experts over intelligentie beschreven. Wat hoort er wel en wat hoort er niet bij en hoeveel dimensies zijn erin te onderscheiden? Ten tweede worden enkele niet-westerse opvattingen over intelligentie besproken. Ten derde komen enkele hardnekkige gedachten over intelligentie, bijvoorbeeld over de relatie van IQ met schedelomvang en fysieke aantrekkelijkheid aan de orde.

Westerse opvattingen over intelligentieverschillen en hun relatie met creativiteit
Opvattingen over individuele verschillen zijn ten eerste te vinden in ons taalgebruik en ten tweede in opinies van ouders en leraren en ten derde zijn er empirische studies over inhoud en structuur van wat leken en experts zeggen en geloven dat intelligentie is.

Ons taalgebruik bevat woorden die uitdrukken in welke mate wij anderen en onszelf intelligent, verstandig, slim, handig, gehaaid, uitgekookt of geletterd vinden. Neisser (1976) zegt dat we een prototypisch concept van een intelligent persoon hebben. Dat houdt in dat sommige kenmerken meer tot het concept 'intelligent persoon' behoren dan andere. Beschouwt u bijvoorbeeld belezen zijn (denk aan Plato, die het schrift afwees) als een directer kenmerk van intelligentie dan een goed geheugen, communicatief vaardig zijn of met een iPad kunnen omgaan? Is snel een probleem doorzien een belangrijker kenmerk dan assertief zijn?

Ouders en leraren hebben opvattingen over wat een intelligent kind of leerling is. Ze denken na over de vraag of intelligentie kan veranderen. Als ze dat geloven zoeken ze naar omstandigheden die veranderingen in de gewenste richting bevorderen. Leraren denken dat een intelligent kind zijn woordje goed kan doen, veel weet en open en gewetensvol is. Dat kan verschillen per (sub)cultuur. Er is empirisch onderzoek naar impliciete deze opvattingen. Voorbeelden:

(1) Nevo (1993) heeft opvattingen van studenten over intelligent gedrag verzameld door hen te vragen waarom, bij welke kenmerken en onder welke voorwaarden antwoorden op testvragen en prestatietoetsen correct waren. Volgens hen is een antwoord correct als een aantal regels wordt toegepast. Ze zeggen bijvoorbeeld: 'Er is aansluiting bij algemeen geldige opvattingen of bij bestaande kennis; het staat in de boeken; ik weet het antwoord gewoon; je kunt er achter komen door goed te kijken of te luisteren en, er is een logische of rekenkundige procedure of regel gevolgd'. De argumenten zijn voor een deel conventioneel (wat we allemaal vinden; het staat in boeken) en voor een deel wetenschappelijk want gebaseerd op methodologisch aanvaarde regels.

(2) Er is empirisch onderzoek naar de inhoud en structuur van opvattingen van leken en deskundigen over intelligentie. Sternberg et al. (1981) vroegen 500 studenten en 150 cognitieve en

persoonlijkheidspsychologen om kenmerken en handelingen te noemen die bij een 'intelligent persoon' horen. Ze noemden onder meer logisch redeneren, veel lezen, een open geest en gezond verstand. Vervolgens gaven ze op een 7-puntsschaal aan in welke mate de kenmerken pasten bij een 'ideaal intelligent persoon'. De variantie in de antwoorden gaven ze met drie factoren weer:

I *Practical problem solving*: redeneert logisch, legt en ziet verbanden tussen ideeën, overziet de belangrijke aspecten van een probleem.

II *Verbal ability*: spreekt duidelijk en gearticuleerd, spreekt vloeiend, weet veel van een bepaald onderwerp af.

III *Social competence*: accepteert anderen zoals ze zijn, geeft eigen fouten toe, toont interesse in de wereld om hem heen.

Uit de antwoorden van de experts trokken ze eveneens drie factoren:

I *Verbal intelligence*: heeft een grote woordenschat, leest en begrijpt, is nieuwsgierig.

II *Problem solving ability*: past kennis toe op concrete problemen, neemt goede beslissingen, formuleert problemen helder.

III *Practical intelligence*: doorziet situaties, bepaalt hoe op effectieve wijze gestelde doelen te bereiken, is gevoelig voor de concrete wereld om hem heen.

(3) In 1985(a) verrichtte Sternberg een studie met de concepten intelligentie, creativiteit en wijsheid. Hij vroeg collega's van de Yale universiteit van afdelingen als kunstgeschiedenis, filosofie, natuurkunde en economie om kenmerken van deze concepten op te schrijven. Uit hun antwoorden leidde hij handelingen en kenmerken af. Deze werden voorgelegd aan docenten en studenten. Zij gaven met behulp van 9-puntsschalen aan of de kenmerken en activiteiten pasten bij het prototype van een 'intelligent, wijs' en 'creatief persoon'. Een niet-metrische multidimensionele schaalanalyse - in die tijd een alternatief voor factoranalyse - leverde drie tweepolige dimensies op. Ze 'verklaarden' 82% van de variantie in de schaalcores:

Dimensie 1: vaardigheid om praktische problemen op te lossen, met aan de ene pool: ziet of doelen haalbaar zijn, gebruikt procedures flexibel, past kennis gemakkelijk toe en aan de andere pool: verbale vaardigheid, kan over veel thema's zinvol meepraten, beschikt over een goede woordenschat, speelt met taal.

Dimensie 2: intellectuele balans en integratie, met aan de ene pool: ziet verschillen en overeenkomsten, ziet verschillende kanten aan een vraagstuk, doorziet abstracte ideeën en doelgerichtheid en aan de andere pool: zoekt en benut doelgericht informatie, presteert veel, is gemotiveerd om doelen te bereiken.

Dimensie 3: contextuele intelligentie, met aan de ene pool: leert van successen en fouten, begrijpt en interpreteert de omgeving, weet wat er aan de hand is in de wereld en *fluid thought* (snel nieuwe dingen leren) en de andere pool: het vermogen wiskundige en ruimtelijke relaties te begrijpen, zeer intelligent, denkt snel.

De vragen over creativiteit en wijsheid leidden tot vier en drie tweepolige dimensies. Intelligentie en wijsheid hingen met gemiddelde r-waarde (rond de .30 .40) samen. De

relatie tussen creativiteit en intelligentie was wat lager. Creativiteit vraagt behalve schoolse, abstracte, formele of *academic* intelligentie ook ruime inhoudelijke kennis van een vakgebied, een zelfsturende denkstijl, tolerantie voor onzekerheid en een hoge motivatie (Sternberg & Hubart, 1992). We zagen hiervoor dat een facet van creativiteit: 'metaforen kunnen formuleren' met een bèta van .49 met IQ samenhang. Er is ook een waarde van $r = .17$ gevonden. De samenhang varieert van bescheiden tot gemiddeld.

(4) Jauk et al. (2013) verdedigen voor de relatie tussen intelligentie en creativiteit een *drempelhypothese*. Dat wil zeggen dat een bepaald niveau van intelligentie is nodig voor creatieve prestaties, bijvoorbeeld een IQ van 120 of meer. Ze gebruikten gesegmenteerde regressies en konden met iteratieve berekeningen de drempel empirisch bepalen. De deelnemers waren 297 Duitse studenten en andere leken (101 mannen) tussen de 18 en 55 jaar. Ze werden gerekruteerd via een plaatselijke krant en de universiteitsmailinglijst. Het bleek dat voor het genereren van originele ideeën: 'noem twee originele ideeën' een IQ van 100 voldoende was. Meer dan twee originele ideeën werden gevonden bij de IQs van ongeveer 120 en hoger. Om deze bevindingen in verband te brengen met Sternberg en Hubarts' vinding kan je denken aan het volgende: Bij een bepaalde IQ-drempel gaan persoonlijkheidskenmerken meedoen bij het voorspellen van creatieve prestaties, bijvoorbeeld, Openheid en Gewetensvolheid. Het laatste kan ook weer geconditioneerd zijn: artiesten 'bloeien' wellicht op bij een laag niveau van Gewetensvolheid gedurende een fase in het creatieve proces, maar bij empirisch-analytisch georiënteerde wetenschappers zou je een 'Streng Methodologisch Geweten' verwachten, maar dit is speculatie. Dit geweten leek in de methodologie opleiding van psychologen in de roerige jaren 60 tot 80 bijna op een vervanger het zondigheidsbesef. Het verschil tussen intelligentie en creativiteit is ook beschreven met behulp van informatieverwerkingsprocessen als *updating*, *shifting* en *inhibition* (zie verderop).

Leken hanteren kortom vrijelijk, maar niet zomaar allerlei criteria om iemand intelligent te noemen. De intelligentiefactoren van studenten en experts verschilden weinig volgens Sternberg. Hij benadrukt daarom de overeenkomst expliciete en alledaagse theorievorming. Er zijn ook verschillen. Een lekentheorie wordt niet getoetst aan empirische gegevens. Er worden geen objectieve tests gebruikt. Ze leggen bovendien meer nadruk op sociale vaardigheden en communicatie dan experts (Furnham, 1990). Intelligentie wordt gauw vergeleken met creativiteit en men laat graag zien dat ze onafhankelijk zijn (Sternberg). Dat is voor een beperkt deel juist. Diagnostici benadrukken ten opzichte van leken eerder intrapersonlijke kwaliteiten (trekken en intelligentie) dan interpersoonlijke (sociabiliteit).

Praktische intelligentie: Sternberg et al. leggen verder bij het benoemen van de factoren nadruk op praktische, contextuele en creatieve aspecten van intelligent en cognitief handelen. Schoolse, abstracte intelligentie: *academic intelligence*' is niet genoemd. Hij vertelde in voordrachten (Nijmegen, 1992; Washington, 1999) dat zijn kracht daar niet lag. Hij behaalde niet de hoogste scores op IQ en *Scholastic Achievement Tests*. Zijn kracht lag in zijn creativiteit en praktische intelligentie. In 2014 vertelt hij dat hij steeds dat bestudeerde

waar hij niet goed in was: *I study what I stink at*. Hij noemde zijn prestaties op IQ tests en het volhouden van zijn intieme relaties. Zijn triarchische theorie over liefde schreef hij na het mislukken van een relatie. In twee monografieën bepleit Sternberg (2000a, 2000b) het belang van praktische, contextuele intelligentie.

Niet-westerse opvattingen over intelligentieverschillen Intelligentie is een westers concept. Historisch ligt daar de nadruk op ideeën, geest, ratio, het denken. Sternberg voegt daar het slim handelen en praktische intelligentie aan toe. Vooral sinds de Descartes en de Verlichting staan het bewustzijn, de geest, de ziel aan het roer en schepen ze de emotie en motivatie bij zich in. In de 19^{de} en de 20^{ste} eeuw worden verschillen in intelligentie belangrijk. Nemen andere culturen die het westerse economische succes zien of zagen de inhoud en structuur van dat concept over? Voorbeelden:

(1) Berry en Bennett (1992) vroegen aan zestig Eskimo's van de Cree stam in Noord-Canada termen te noemen die met intelligentie te maken hadden. Ze verzamelden ongeveer twintig termen. De deelnemers maakten stapeltjes met gelijke termen. Multidimensionele schaling op deze gelijkenissen leverde twee dimensies op:

I Aan de ene pool termen als wijs, denkt grondig, schenkt aandacht, heeft richtingsgevoel en toont respect en aan de andere pool termen als dom, gek en achterlijk denken.

II Aan de ene pool stevig en moedig en aan de andere pool begrijpt nieuwe dingen en is religieus. Intelligentie heeft voor de Eskimo's niet de Westerse gevoelswaarden van 'hersens hebben en snel en scherp analyseren'.

(2) Soortgelijke bevindingen zijn gedaan bij de volwassenen van de Chewa stam in Zambia. Zij legden de nadruk op sociale verantwoordelijkheid, samenwerking en gehoorzaamheid. Intelligentie werd in Zimbabwe vooral gezien als wijs en voorzichtig handelen. Sternberg et al. (1997) ontmoetten in dorpen in Kenia kinderen die intelligent genoemd werden maar het slecht deden op school. Ze wisten wel veel van kruiden. Voor een deel is het westerse intelligentiebegrip daar wel bekend maar ze leggen de nadruk op zaken die met dagelijks overleven te maken hebben.

(3) Er zijn culturen die met de Amerikaanse in aanraking zijn gekomen en waarvan de leden Amerikaans spreken. Welke opvatting over intelligentie treffen we daar aan? Yang en Sternberg (in: Sternberg, 1997) vonden vijf factoren bij Taiwanese: I Een algemene cognitieve factor II: Interpersoonlijke intelligentie III: Intrapersoonlijke intelligentie IV: Intellectuele zelfbewustheid en V: Intellectuele bescheidenheid.

Er ligt bij sommige niet-westerse gemeenschappen naast rationele aspecten nadruk op verantwoordelijkheid en onafhankelijkheid. Dit hangt samen met de sociale organisatie van die groepen. Een groep liet kinderen vrij en gaven hen al jong verantwoordelijkheid. Zij mochten al vanaf hun negende een geweer dragen en gebruiken. Er is nadruk op behoud van de sociale verhoudingen en bescheidenheid.

Hardnekkige ideeën over intelligentieverschillen Er zijn opvattingen van leken en sommige professionals over hersenomvang en intelligentie en over hoe leken hun m/v verschillen in intelligentie en eigen potentieel inschatten. We vergelijken ook graag landen. Zijn Amerikanen inderdaad de intelligentste mensen ter wereld? Voorbeelden:

(1) Een oude en hardnekkige opvatting is dat intelligentie met de omvang van het hoofd (hersenvolume) samenhangt. De schrijver Willem Elsschot (Verzameld Werk, 1992) schreef een wraage en hilarische novelle waarin de zoon van de sigarenmaker een groot hoofd had en dat bestemde hem voor tot een grootse intellectuele loopbaan. Deze mislukte onder meer omdat hij geen Frans verstond en sprak. In die tijd was Frans de taal van de midden- en hogere klasse en hogere onderwijsinstellingen in België. De opvatting is niet tot de belletrise beperkt. Lombroso (1936-1909) dacht dat hij criminelen kon identificeren door de hersenomvang te meten. Hij was een antropoloog en sociaalbewogen criminoloog. Hij beschouwde criminaliteit als een genetisch defect dat aan het uiterlijk af te lezen was. Galton speculeerde al ook over het feit dat schedelomvang en zijn algemene G-factor gecorreleerd waren. Dat is ook uitgezocht. Deary et al. (2008) zijn psychometrische intelligentie-onderzoekers en rapporteren een bijzondere of eigenaardige $n = 1$ studie. Ze berekenden bij 48 gezonde mannen tussen 71 en 79 jaar de correlatie tussen schedelinhoud en IQ. De waarde was $r = .50$. Nu maten ze het volume van de schedel van Koning Robert I van Schotland (1274-1329). Ze gingen er kennelijk van uit dat die niet gekrompen was na zolang in Moeder Aarde gelegen te hebben. Ze keken daarvoor naar volume, schedelbreedte en -lengte. Ze gebruikten de regressiecoëfficiënt (de b in de lineaire relatie van twee variabelen X en Y : $Y = a + bX$) bepaald bij die 48 mannen en schatten daarmee het IQ van Koning Robert. Ze kwamen uit op een IQ van 128 (95% betrouwbaarheidsinterval: tussen 106 en 130). Dit mag controversieel klinken maar het is geen uitzondering: McDaniel (2005) rapporteerde in een meta-studie een correlatie van $r = .33$ tussen hersenvolume en IQ.

(2) Burgaleta et al. (2012) observeerden dat mannen gemiddeld grotere hersenen hebben dan vrouwen en dat hersenvolume correleert met IQ. Ze lieten ook zien dat de significante verschillen tussen de hersenvolumes *niet* gecorreleerd waren met verschillen in Spearman's G . Sekseverschillen waren er wel: mannen scoorden hoger op ruimtelijke vaardigheden dan vrouwen maar dit was niet gecorreleerd met G . Hiervoor is beweerd dat de verschillen tussen mannen en vrouwen gering zijn, maar in grote steekproeven significant. De herseninhoud van de homo sapiens is groter dan die van apen. Ze waren door rechtop te lopen meer op waarneming (kijken) dan op reuk aangewezen.

(3) Kanazawa (2011) vond een lage correlatie (voor de VS) van $r = .126$ en een gemiddelde correlatie (voor het VK) van $r = .381$ tussen intelligentie, gemeten met multiple tests die .991 (VK) en .999 (VS) laden op de G-factor met *aantrekkelijkheid*. Ze maten dit laatste door middel van het oordeel door leraren (VK) en *gerated* door *Add Health* ondervragers (VS) op een 5-puntsschaal. De VK steekproef was voornamelijk blank en bestond in het begin uit 17.419 baby's die gevolgd werden tot hun 47^{ste} levensjaar. Er waren 9543 over. De VS steekproef bleek achteraf redelijk representatief en bestond uit 18- tot 28-jarigen ($n = 20.745$). Sekse was een moderator variabele: de correlatie was hoger voor mannen dan voor vrouwen in beide steekproeven. De correlatie bleef significant na controle voor SES, lichaamslengte en gezondheid. De bescheiden en gemiddelde correlaties zijn volgens de auteur het gevolg van een algemene fitness factor.

(4) Een terugkerend thema is het vergelijken van de IQs van landen of staten. Dit soort onderzoek bergt het risico in zich het eigen land superieur te achten ten opzichte van andere. Sommige

onderzoekers schromen niet om dat te doen, bijvoorbeeld Jensen en Rushton. Ze moeten erkennen dat Amerikanen niet de intelligentste zijn.

(5) Leken hebben een opvatting over hun IQ ten opzichte van dat van anderen. Er zijn schattingen gemaakt voor en door leden uit verschillende culturen. Dit is accuraatheidsonderzoek, naarmate je dichter bij je met een IQ-test gemeten score zit ben je accurater. Hoe zien leken zichzelf en hoe leden van andere culturen? Furnham et al. (1999) lieten proefpersonen schatten hoe hoog hun algemene intelligentie was en hoe hoog ze op de zeven factoren van Gardner scoorden. Mannen gaven zichzelf hogere scores dan vrouwen op ruimtelijke en numerieke intelligentie. De verschillen waren klein: slechts twee punten op een schaal met een gemiddelde van ongeveer 106 en een spreiding tussen de 4.5 en 9.0. Er was een cultuurverschil: de Hawaïanen waren het bescheidenst (104.7), de inwoners van Singapore, gemiddeld (106.7) en de Britten het minst bescheiden (109.1). Zo bestaan er naast verschillen in inhoud ook verschillen in gemiddelde intelligentie die mensen uit verschillende culturen zichzelf toekennen. Allen schatten zich als groep overigens hoger in dan het gemiddelde van $M = 100$ en $SD = 15$. Ze overschatten zich met een halve standaarddeviatie ($d = .50$, dat is 7.5 IQ punten). Dit is mogelijk een gevolg van de steekproef. Allen moesten immers naast de moedertaal het Engels machtig zijn.

(6) Ratan et al. (2012) beweren een nieuwe dimensie aan impliciete opvattingen over intelligentie toegevoegd te hebben: het *potential*. Negen en zeventig Californische studenten (52 vrouwen) leeftijd gemiddeld 19.1 jaar, $SD 1.18$, en 69 studenten uit Bangalore, de IT-stad van India, (41 vrouwen) gemiddelde leeftijd 22.4, $SD 1.59$) namen deel. De Indiase studenten geloofden meer dan de Amerikaanse in de kans het ver te brengen (hun potentieel): $t(91) = 2.03$ ($p < .05$). De Indiërs geloofden eerder dan de Amerikanen dat ze een Nobelprijs konden winnen. *Beliefs* hebben gevolgen voor gedrag. Het geloof dat iedereen zeer intelligent kan worden, voorspelt steun bij het gelijk verdelen van de middelen over allen, over rijken en armen. Dit is verwonderlijk voor India omdat het een kastensamenleving is die stilzwijgend maar niet officieel aanvaardt dat middelen en welvaart ongelijk verdeeld zijn. Er wordt vooral lippendienst bewezen aan gelijkheid. Mogelijk vergelijken de studenten zich alleen met hun eigen groep of kaste. Amerikanen geloven naar verhouding meer in het gelijk verdelen van middelen. Dat is ook aan hun politiek te zien. Denk aan massale onderwijsprogramma's zoals *Head Start*. Daar geldt dat als men er na zoveel steun nog niet komt het aan de groep en het individu zelf moet liggen.

(7) In 2001 vatte Furnham de studies over sekseverschillen over een periode van vijftig jaar samen. Mannen schatten zichzelf hoger dan vrouwen, allen denken dat ze hoger scoren dan hun grootouders en iets lager dan hun kinderen. De geschatte scores correleren laag met feitelijke IQ-scores. Ze zijn niet accuraat. Mannen denken dat vrouwen lager scoren op wiskundige en ruimtelijke taken. Een realistische schatting voor de tweede decade van de 21^{ste} eeuw is dat ongeveer 20% van de kinderen hoger op de maatschappelijke ladder komt dan hun ouders. De meesten blijven gelijk en ook ongeveer 20% eindigt lager op die ladder.

(8) Zell en Krizan (2014) vatten de resultaten van 22 meta-analyses samen over de correspondentie tussen zelf-evaluaties van intelligentie, taalvaardigheid, medische kennis en vaardigheden, sportprestaties en beroepsvaardigheden en objectieve metingen van die prestaties. De correlaties varieerden tussen de $r = .09$ en $r = .63$ met een gemiddelde van $r = .29$ ($SD = .11$). Dit wijst erop dat mensen gemiddeld genomen een bescheiden inzicht hebben in hun vaardigheden en context factoren die een accurate schatting bevorderen, onderschatten.

Er blijven kortom hardnekkige gedachten over het bestaan van een relatie van IQ met hersenomvang en aantrekkelijkheid. Hersenvolume en IQ hangen bescheiden samen. De relatie kan niet afgedaan worden als pure onzin. Het kan te maken hebben met voeding en armoede. Man-vrouw verschillen zijn significant in grote steekproeven. De waarden zijn laag en zouden weg *kunnen* vallen na DIF analyses. Het eigen IQ, iemands kennis van zaken en potentieel worden overschat en dat verschilt tussen culturen. De meta-studie over zelf-evaluaties van Zell en Kristan laat als vaak grote variatie in r-waarden zien: van verwaarloosbaar tot hoog. Over het geheel genomen zijn de samenhangen gemiddeld; ongeveer 10% van de variantie is gemeenschappelijk.

Er is mij geen onderzoek bekend naar alledaagse opvattingen over *geschiktheid*. Deze zijn echter wel te verzinnen: som geschiktheden op, bijvoorbeeld ruimtelijk inzicht, technisch inzicht, nauwkeurigheid, rekenvaardigheid, verbale vaardigheid en woordenschat. Plaats daarnaast functies en beroepen, bijvoorbeeld machinebankwerker, advocaat, psycholoog, arts, commercieel manager, leraar en bankemployé. Vraag nu een steekproef verbindingen te maken. Als elke geschiktheid bij elke functie even vaak genoemd wordt dan is er geen verband. Het is echter te verwachten dat bepaalde vaardigheden vaker bij bepaalde functies genoemd worden dan bij andere. We zouden ook een steekproef kunnen vragen om allerlei geschiktheden te noemen die bij concrete functies of beroepen passen.

Nog een manier is het raadplegen van advertenties in vakbladen voor allerlei beroepen. We kunnen dan lezen welke geschiktheden bij welke functies vereist en aanbevolen worden. Dan leren we meteen of personeelsmanagers en commerciële selectiebureaus in staat zijn onderscheid te maken tussen de taken in een bedrijf of dat ze van iedere sollicitant hetzelfde, het maximum eisen. Er is geen belangstelling voor dit onderwerp.

Ook is mij ook geen onderzoek bekend naar alledaagse opvattingen over *prestaties*. We kunnen de betrekkelijk willekeurig samengestelde examens, proefwerken en opstellen uit het onderwijs als zodanig opvatten. Deze lagen onder vuur omdat ze niet betrouwbaar en valide zouden zijn, denk aan De Groot (1967) in zijn boek *Vijven en Zessen*. Tijdens zijn sabbatical in de VS (Stanford) heeft De Groot het *Educational Testing Service* (ETS) in Princeton bezocht en dat als voorbeeld genomen voor het Cito. Er stond hem ook een selectievrij onderwijs voor ogen. Dat is er nooit gekomen is. Het Nederlandse onderwijs kent horden en afbreekrisico's. Het Cito domineert nu het meten van onderwijsprestaties. Of De Groot het eens zou zijn geweest met de ontwikkeling dat formele Rasch modellen uitmaken wat wel of niet zinvolle toets inhoud is, kunnen we niet meer nagaan. Gezien zijn belangstelling voor inhoud, kan men vermoeden dat hij terughoudend geweest zou zijn. Het leraarsoordeel is echter mede door kennis en inzicht in de Cito-toetsen nauwkeuriger geworden. Beishuizen et al. (2003) vonden dat het advies van de schoolkeuze door de schoolleiding bij 361 leerlingen in 87% van de gevallen overeenkwam met een advies op basis van een objectieve intelligentietest met een goede predictieve validiteit. Slechts bij 1% was het advies op grond van de testuitslag ongunstiger dan dat van de schoolleiding. De

leerkrachten beschikten over de testresultaten van de leerlingen. Het verschil Cito-toets en leraar oordeel is klein. Het kan de ene of de andere kant opgaan. Soms is hun oordeel *lager of hoger* dan het advies op grond van de Cito-uitslag.

Prestatiemetingen door leraren zelf hoeven niet zonder meer opzij gezet te worden. Een niet zo betrouwbare en valide mondelinge of schriftelijke overhoring kan een *formatieve* functie vervullen. Dat wil zeggen dat het resultaat wordt benut om het onderwijs bij te stellen. De leraar moet zich er van meet af aan van bewust zijn dat zijn formatieve 'toets' de prestatie kan over- of onderschatten. Hij merkt dat meteen en dan hoeven er geen batterij tests en een leger psychometrici voor in stelling gebracht te worden.

Samenvatting en conclusie

Intelligentie opvattingen van leken kunnen begrepen worden als een prototypisch construct: bepaalde termen behoren er meer toe dan andere. Studenten gebruiken conventionele, bijvoorbeeld 'het staat in boeken', logische en empirische regels als argumenten dat antwoorden op testvragen goed zijn en dus van intelligentie getuigen. De inhoud en structuur van termen die naar intelligentie verwijzen zijn bij leken en experts onderzocht. De antwoorden tonen drie vergelijkbare factoren: I probleem oplossen, II verbale vaardigheid en III sociale en praktische intelligentie. Hoewel de uitkomst van leken en experts vergelijkbaar is zijn de methoden verschillend: experts leggen nadruk op intrapersonlijke kenmerken (trekken en intelligentie) en leken op interpersoonlijke vaardigheden (sociale intelligentie). Niet-westerse, niet-geïndustrialiseerde culturen leggen nadruk op sociale functies, zoals verantwoordelijkheid, samenwerking en bescheidenheid. Er zijn geen studies over de opvattingen van leken over geschiktheden. Leken maken waarschijnlijk onderscheid tussen geschiktheden die nodig zijn voor verschillende beroepen. Opvattingen van leken over prestaties zijn niet rechtstreeks uitgezocht.

Er zijn en blijven bijzondere opvattingen over de relatie tussen lichamelijke kenmerken en intelligentie, bijvoorbeeld tussen schedelinhoud, aantrekkelijkheid en IQ. Mensen overschatten gemiddeld genomen hun eigen IQ met bijna een halve SD. Er is samenhang tussen IQ en schedelomvang of je dat nu wil of niet. Onderzoek laat een bescheiden correlaties zien. De interpretaties zijn wisselend, bijvoorbeeld gezonde voeding, armoede, verwaarlozing, enzovoort. Kanazawa zoekt het in een algemene fitheidsfactor. De variatie correlaties is groot. De uitslagen zijn niet stabiel over steekproeven.

2. Ontwikkeling

Er is beperkt onderzoek naar *beliefs* over ontwikkeling van intelligentie. Er is niet veel nagedacht over de ontwikkeling van geschiktheden en prestaties. Voorbeelden:

(1) Ross (1989) vroeg studenten grafieken te tekenen die uitdrukten hoe gedragskenmerken zich ontwikkelden tussen 5 en 85 jaar. Allerlei vormen kwamen voor: toename, afname, stabiliteit en pieken gedurende de levensloop (zie ook Adolph et al., 2008).

(2) (Gergen et al. (1990) vonden dat jonge moeders en vrouwen zonder kinderen een organismisch groeimodel veronderstellen voor de ontwikkeling van baby's. Ze beschrijven stadia tussen de 0 en 50 maanden. In de eerste maanden verschijnen sympathie en trots. Rond 3 jaar zien ze sporen van een eigen wil en het vermogen tot planning. De rol van de moeder loopt parallel: eerst tonen ze affectie (sussen, affectie, comfort bieden) maar aan het eind van het tweede jaar leggen ze de nadruk op sociale uitwisseling en interactie: spelletjes doen, samen spelen, prijzen, maar ook afwijzen en grenzen stellen.

(3) Cruys (1992) constateerde dat mensen spontaan stadia (babytijd, peutertijd, lagere schooljaren, puberteit) en domeinen van ontwikkeling (lichamelijk, verstandelijk, emotioneel, sociaal) onderscheidde.

(4) Furnham (2001) liet zien dat ouders veronderstellen dat hun kinderen iets *intelligenter* zijn dan zichzelf. Zij beschouwen zichzelf op hun beurt intelligenter dan hun ouders. Ze verwachten een geleidelijke toename van intelligentie over generaties. Het Flynn effect bevestigt dit. Er bestaan ook individuele verschillen in *beliefs* over de mate waarin intelligentie beïnvloedbaar is.

Leken hebben dus opvattingen over ontwikkeling van intelligentie. Die passen in hun beeld over bijvoorbeeld hoe een kind zich ontwikkelt en welke rol ouders daarbij spelen. Impliciete opvattingen over de ontwikkeling van *geschiktheden* zijn niet onderzocht. Er worden effecten van training verwacht. Bovendien is het *belief* waarschijnlijk dat er domeinspecificiteit is. De één leert bijvoorbeeld snel wiskunde, een ander gemakkelijk een taal.

Leken hebben impliciete opvattingen over de ontwikkeling van *prestaties*. Ze gebruiken tijdtafels voor wat, wanneer geleerd kan of moet worden. Ieder wordt geconfronteerd met eigen en andermans prestaties door de tijd heen. Er zijn 'groeitheoretici' die aannemen dat het volgens een organismisch plan van binnenuit verloopt en er zijn 'mechanistische theoretici' die veronderstellen dat ongeveer alles op ieder tijdstip geleerd kan worden, gegeven goede training en voldoende tijd. Ouders proberen een verklaring te vinden voor het verloop van schoolprestaties van hun kinderen. Als de resultaten stabiel en op een aanvaardbaar niveau zijn dan wordt meestal een normale groei verondersteld in een doorsnee sociale context. Als het niveau afwijkt en lager is dan verwacht worden verklaringen gezocht. Oorzaken worden 'gesitueerd' in het kind, de sociale context (gezin, school) of in het curriculum. Bijvoorbeeld, het kind is nerveus, speels, lui, kan zich niet concentreren; de leraren zijn niet bekwaam, hun onderwijsstijl past niet bij het kind, of de inhoud van het curriculum sluit niet aan bij de belevingswereld van het kind. Ouders kunnen ook zichzelf beschuldigen dat ze vroeger ook niet de hoogste prestaties behaalden of te weinig gemotiveerd waren. Zijn de prestaties hoger dan verwacht dan worden ze meestal toegeschreven aan kenmerken van het kind en de gezinsomstandigheden. Dit is een voorbeeld van de attributiefout. Dat houdt in dat we oorzaken van het gedrag van anderen toeschrijven aan hun disposities en die van onszelf ook aan de situatie.

Opvattingen van leken over de ontwikkeling van *prestaties* zijn niet systematisch onderzocht. De diagnosticus komt ze wel tegen in zijn praktijk. Er zijn socio-cultureel

bepaalde tijdtafels voor prestaties die hem helpen te bepalen of prestaties op school, in werk en in het sociale leven een 'natuurlijke' of een cultureel opgelegde ordening volgen.

Samenvatting en conclusie

De vroegste ontwikkeling van cognitie wordt door ouders meestal organismisch opgevat, dat wil zeggen als een geleidelijke of stadiumgewijze groei van binnenuit. Leken hebben opvattingen over de oorzaken van de ontwikkeling van intelligentie. Ze veronderstellen dat hun kinderen een iets hogere intelligentie hebben dan zichzelf. Geschiktheid wordt opgevat als de aanleg, het vermogen zich te specialiseren op een domein. Leken verschillen in hun opvatting over geschiktheid: is het een vast of een onbestemd, flexibel beginkapitaal? De opvattingen over de ontwikkeling van prestaties zijn voor een deel bepaald door socio-culturele tijdtafels. Als prestaties achteruitgaan of niet aan de verwachting voldoen, worden de oorzaken in de persoon en zijn sociale context gezocht.

3. Sociale context

Volgens de *beliefs* van leken moet intelligentie tot uitdrukking komen in de sociale context. Sternberg (1985) omschreef de derde leken-dimensie als contextuele intelligentie'. Antwoorden op de vraag welke kenmerken bij een intelligent persoon horen leverden als derde factor praktische intelligentie op. De invloed van de sociale context op IQ-scores is een oud vraagstuk: ouders noemen genetische en omgevingsfactoren als oorzaken voor IQ-verschillen bij hun kinderen. Meestal veronderstellen ze dat de verschillen *tussen* gezinnen groter zijn dan *binnen* hun gezin. In gedragsgenetische studies wordt dat niet bevestigd. De IQ-verschillen van kinderen binnen een gezin variëren en zijn gemiddeld bijna 1 SD ($M = 100$; $SD = 15$). Het gemiddeld verschil bedraagt ongeveer 12 IQ-punten. Deze waarde is van de jaren 70 van de vorige eeuw. Nieuw onderzoek zou deze waarde mogelijk kunnen verlagen omdat er minder kinderen per gezin zijn en er meer aandacht voor individuele kinderen is door ouders en kinderopvang. Hoe de invloed van de omgeving op intelligentie en cognitie precies wordt opgevat is niet empirisch onderzocht. Het is aannemelijk dat vooruitgang verondersteld wordt, bijvoorbeeld van naïef naar ontwikkeld, van eenvoudig naar gesofisticeerd, van eenvoudig naar complex en van aandacht voor bijzaken naar hoofdzaken. Als oorzaak wordt geen specifiek kenmerk in de context aangewezen want het doet zich voor afgezien van kenmerken van de context.

Leken veronderstellen een invloed van de vormgegeven context, bijvoorbeeld school- en trainingsprogramma's en van de natuurlijke context, bijvoorbeeld gezin, leeftijdgenoten, buurt op geschiktheid en prestatie. Ze investeren immers veel in het vinden van de juiste context om hun geschiktheden te ontplooiën: scholing en training worden gezien als kansen om vooruit te komen in de maatschappij.

Samenvatting en conclusie

Leken verwachten dat intelligenties binnen het gezin ongeveer gelijk zijn en toenemen over generaties. Er wordt ook progressie verwacht bijvoorbeeld van eenvoudig naar complex,

afgezien van specifieke contextinvloeden. Het belang van contexten van scholing en training voor geschiktheden en prestaties blijkt uit investeringen die men zich getroost om zich te scholen en te bekwamen.

4. Reflectie en evaluatie

De leek legt als amateurtheoreticus meer nadruk op praktische en contextuele intelligentie dan de diagnosticus. Deze stelt *academic* kenmerken vast zoals algemene intelligentie en ruimtelijke en numerieke vaardigheden. Als de diagnosticus met andere culturen te maken heeft is hij er zich van bewust dat er naast assimilatie en aanpassing aan het westerse intelligentiebegrip ook andere aspecten van het verstand zoals verantwoordelijkheid voor de groep en onafhankelijkheid centraal gesteld worden. Westerlingen overschatten hun IQ en waarschijnlijk ook hun geschiktheid en prestaties. Ze doen dat bovendien meer dan leden van andere culturen. Ieder schat de eigen groep hoger in dan het berekende gemiddelde. Dat geldt niet alleen voor het IQ. Tachtig procent van de ondervraagden noemt zich humoristischer dan het gemiddelde (50% dus). Sommigen noemen dit verschijnsel humor. Impliciete opvattingen over ontwikkeling van cognitie tonen sporen van een vooruitgangsgedachte: van eenvoudig naar complex, van onrijp naar rijp en van naïef naar gesofisticeerd. Het IQ wordt als betrekkelijk stabiel beschouwd en er wordt vooruitgang over generaties verwacht. Leken zoeken oorzaken van IQ-verschillen in genetische aanleg en omgeving. De omgeving is vooral het gezin en de school. Ze nemen ten onrechte aan dat verschillen binnen gezinnen klein en tussen gezinnen (de burens) groot zijn. De context komt bij leken vooral ter sprake als het om geschiktheid en prestatie gaat.

De diagnosticus kan verwachten dat bij het niet tot ontplooiing komen van geschiktheden en het achterblijven van prestaties allerlei contextfactoren als oorzaak genoemd worden. Diagnostici komen onvermijdelijk in aanraking met deze opvattingen. Meer onderzoek over die opvattingen zou hen van pas komen.

III **Expliciete theorieën van intelligentie en prestaties**

Intelligentie is het meest onderzochte psychologische construct. Het begrip speelde aanvankelijk een rol bij selectie van personeel en toelating van leerlingen tot het onderwijs. Galton was geobsedeerd door meten en zette zijn kennis in om na te gaan of genialiteit overerfbaar was. Evenals bij persoonlijkheid overheerst de individuele verschillen oriëntatie die van de ontwikkeling en de sociale context. Het belang van het vaststellen van intelligentieverschillen nam toe na de industrialisatie en immigratie aan het eind van de 19^{de} en het begin van de 20^{ste} eeuw in Europa en de VS.

Cognitie volgde een ander spoor. Dat was gericht op het beschrijven en begrijpen van het ontstaan van (wetenschappelijk) kennen. Daarnaast is cognitie opgevat als informatieverwerking en dat past bij de Britse en Amerikaanse empiristische houding over kennisverwerving.

Na het instellen van de leerplicht werd het zaak om de geschiktheid om kennis te verwerven op specifieke domeinen en op school- en beroepsprestaties te meten. Interesse voor intelligentie, cognitie en prestatie staat niet los van de maatschappelijke context. Het vooruitgangsgeloof en het idee van de vorming van kinderen en volwassenen leidden er toe omgevingen zo in te richten dat intelligentie, cognitie en prestaties toenamen of verbeterden. Intelligentiemeting diende praktische doelen. De epistemische vraag hoe de verhouding is tussen kenner en objecten, verschijnselen en gedrag wordt daarbij tussen haakjes gezet. Dat zijn wijsgerige vragen en daar heeft de psychologie zich immers van ontdaan.

De informatieverwerkingsbenadering ontkomt niet aan een (impliciet) epistemologisch standpunt. Je vraagt je af wat de kenner meebrengt en wat komt van objecten, verschijnselen en gedrag. De Gibsons karakteriseerden de relatie met het begrip 'affordances'. Objecten, verschijnselen en gedrag dagen de kenner uit ze te beschrijven onder een bepaald opzicht, bijvoorbeeld wat biedt het mij voor mijn doelen, wat kan ik voorspellen of wat is het object eigenlijk.

Piagets theorie gaat wel over de relatie kenner en de wereld, dat wat het geval is. Deze is nu minder populair. Ze niet past in de methodologische empirisch-analytische stroming en de nadruk ligt volgens critici te zeer liggen op rijping. Dat is een betrekkelijk autonoom proces en niet alles is dus maakbaar aan ons kennen.

Dit hoofdstuk gaat over vragen als: Welke individuele verschil(len) dimensie(s) worden onderscheiden in intelligentie, geschiktheid en prestatie? Hoe wordt er geordend als er meer intelligentiefactoren verondersteld en bepaald worden? Welke oorzaken worden genoemd voor de verschillen in intelligentie, geschiktheid en prestatie? Hoe wordt de ontwikkeling van intelligentie/cognitie, geschiktheden en prestaties beschreven en verklaard. Wat is de rol van de natuurlijke (bijvoorbeeld gezin) en geconstrueerde sociale context (bijvoorbeeld scholing) op intelligentie, cognitie, geschiktheid en prestatie?

1. Individuele verschillen

De discussie over verschillen in intelligentie werd, ten eerste, gedomineerd door Spearman (VK, in navolging van Galton) en door Thurstone en Guilford (VS). Zij waren statistisch geschoold en legden de grondslag van wat 'psychometrische intelligentie' genoemd wordt. Ten tweede schied Vernon (VS) orde in de vele intelligenties die deze auteurs voorstelden. Carroll (1993) heeft die poging herhaald om de definitieve structuur van intelligentie te beschrijven door de gegevens tot dan toe samen te voegen. Dit werk en de inzichten uit informatieverwerking en psychometrie inspireerden, ten derde, sommige auteurs om nieuwe, theoriegeleide opvattingen voor intelligentie te bedenken en te toetsen. Sternberg baseerde zich bijvoorbeeld op informatieverwerking en Embretson haalde haar inspiratie uit de informatieverwerking en de moderne testtheorie. Naast de theorievorming over de bouwstenen van intelligentie is er, ten vierde, de pragmaticus Wechsler die een bruikbaar instrument ontwikkelde als medisch psycholoog van het New Yorkse Bellevue ziekenhuis: de Wechsler schalen. Ten vijfde is over de oorzaken van de individuele verschillen in intelligentie voortdurend gespeculeerd. Niet-psychologische vakken als genetica, neurologie, hersenonderzoek en biologie worden erbij gehaald om intelligentieverschillen te verklaren. Deze uitstapjes naar andere disciplines worden vanzelfsprekend als verdiepend opgevat: hoe biologischer, fysiologischer, neurologischer of fysischer, hoe beter, hoe dieper. Er wordt daarbij niet naar Oppenheimer en Putnam of Fodor (zie hiervoor) verwezen om de verbinding tussen intelligentiepsychologie en andere disciplines te funderen. Het is volgens mij onbegonnen werk. Het psychometrische idee dat de latente trek of het theoretisch attribuut de nuttige eenheid van studie is, ten zesde, aangevuld met studie van elementaire cognitieve taken (ECTs). De prestaties op deze taken zijn gecorreleerd met het IQ. Het oude idee dat onze hersens, ons zenuwstelsel met intelligentie te maken hebben, is uitgewerkt door reactie- en inspectietijden met IQ te correleren.

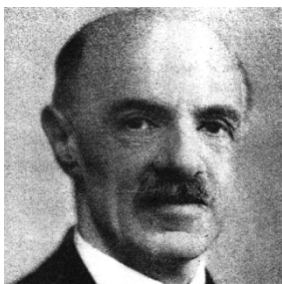
Geschiktheid ligt dicht bij intelligentie. Psychologen die werkten voor specifieke groepen, bijvoorbeeld piloten en operators van ingewikkelde apparatuur, hebben zich gericht op het vaststellen van specifieke geschiktheden. Prestaties op inhoudelijke domeinen werden eerst door leraren en ervaren beroepsbeoefenaars bepaald. Daarna hebben de testpsychologen en psychometrici het vaststellen en meten van prestaties voor een groot deel overgenomen. De Groot (*Vijven en zessen*) van 1967 argumenteerde dat oordelen van leraren onbetrouwbaar en niet valide waren. Binet vertrouwde dat oordeel evenmin en dat bracht hem tot de eerste bruikbare intelligentietest. Er moest een solide basis komen voor het toewijzen van leerlingen aan opleidingen. Daarbij stond De Groot een selectievrij onderwijs voor de geest. Hij wilde niet steeds opnieuw horden opwerpen voor leerlingen. Dat is er niet van gekomen. Hij stond aan de basis van het Cito na een bezoek aan het *Educational Testing Service* in de VS tijdens zijn sabbatical aan de Stanford University, waar Cronbach werkte. De Cito-toetsen zijn misschien horden geworden en De Groot zou zich daar waarschijnlijk tegen verzet hebben. Je kunt zeggen dat de moderne psychometrie het meten van onderwijsprestaties bepaalt. Alles wordt een latente trek volgens een Rasch model van

kennis van Engels, natuurkunde, het theorie examen voor het rijbewijs tot de bekende groep 8 toets. De inhoud van de vakken en vaardigheden lijken op het tweede plan te staan.

Intelligentie is uitvoeriger onderzocht dan enig ander gedragskenmerk. Lynn en Vanhanen (2012) stellen in hun boek over intelligentieverschillen in 160 landen voor om intelligentie te beschouwen als het unificerende construct van de sociale wetenschappen, analoog aan natuurkundige concepten als massa, energie en snelheid. Het IQ brengt echter tot heden *geen* eenheid in de psychologie. Er is voortdurend debat over wat het is, het aantal factoren, of er stadiumgewijze ontwikkeling is en wat de invloed van omgeving en van de aanleg is.

De testbeweging is vanaf het eind van de 19^{de} en begin 20^{ste} eeuw gericht op het meten van intelligentie. *Galton* heeft als een van de eersten proefpersonen taken, vragen en problemen voorgelegd. Hij karakteriseerde ze als uitingen van een aangeboren algemene cognitieve vaardigheid. Deze bestaat weer uit specifieke vaardigheden zoals taal, wiskunde, muziek, geheugen en kunstzinnigheid. Zijn operationalisaties en tests hebben het niet gehaald.

Charles Spearman (1863-1945) zette het werk van Galton voort. Zijn werkkring was het University College in Londen. Hij gebruikte taken en vragen uit bestaande tests en besloot tot een algemene cognitieve vaardigheid: de G-factor. Het bestaan van deze factor wordt bijna altijd bevestigd want er is empirisch steeds een sterke eerste factor. De G betekent ook iets inhoudelijks. Verschillen tussen subgroepen op de subtests van batterij intelligentietests zijn een functie van de cognitieve complexiteit van die subtests, dat wil zeggen de tests leiden tot grote verschillen tussen groepen op subtests met een hoge complexiteit en tot kleine tussen groepen op subtests met een lage complexiteit.



Charles Spearman (1863-1945) was zowel theoreticus als psychometricus. Aan hem danken we de G- en S- factor van de intelligentie.

De G-factor verwijst naar elke activiteit die mentale inspanning vergt. Spearman was ook een van de eersten die factoranalyse gebruikte. De G is daardoor een hoeveelheid van 'iets', dat afgeleid wordt met behulp van een statistische operatie. De G-factor is een ontdekking van het kaliber van de wetten van *reinforcement* door Thorndike. Jensen (1998) overdrijft, als hij de G-factor vergelijkt met de ontdekking van het DNA door Watson en Crick.

Door Galton, Binet en Amerikaanse legerpsychologen zijn tests ontworpen. De correlaties tussen die tests hoefden/konden niet alleen met de G-factor beschreven of verklaard worden. Daarom veronderstelde Spearman het bestaan van een S-factor die gemeenschappelijk is voor groepjes tests met hetzelfde soort items, bijvoorbeeld verbale, ruimtelijke, numerieke of mechanische taken. Dit is de bekende twee-factortheorie: G en S. Spearman schuwde empirisch werk niet. Hij liet ongeveer honderd tests en taken afnemen bij schoolkinderen en bepaalde welke taken en tests een hoge G-lading hadden en welke een lage. De laatste waren gevoelig voor de specifieke vorm van de taken:

Hoge G-ladingen waren er voor Matrixrelaties (.94), Generalisaties (.89), Reeksen aanvullen (.87), Verbale Analogieën (.83), Gelijkenissen (.77), Rekenvraagstukken (.77) en Begrip van een stukje tekst (.73). Naar verhouding lage G-ladingen vond hij bij Snelheid van doorlopen van een doolhof (.04), Getallen doorstrepen (.12), Stippen tellen (.14), Simpele optellingen (.23), Snelheid van tikken, bijvoorbeeld met een potlood op tafel (.24), Snelheid bij het zetten van stippen (.27), Geheugen voor *paired-associates* en herkenning: *recognition memory* .31.

Spearman's veronderstelling is dat de inhoud van items niet van belang is. Hij drukt dat uit met de term *Indifference of the indicator*. Bij de WISC laadt de subtest Woordenschat naar verhouding het hoogst op G. Coderen laadt naar verhouding het laagst (voor Nederlandse gegevens, zie Kan, 2011). Helemaal *indifferent* zijn de subtests dus niet. Hij was een ruwe debater en verweet Binet bijvoorbeeld dat hij maar wat taken bijeen geraapt had die weliswaar het oordeel van de leraar over schoolsucces van leerlingen voorspelde maar waarvoor een conceptueel kader ontbrak (Spearman, 1927, p. 71):

'When Binet borrowed the idea of such promiscuous pooling, he carried it into execution with a brilliancy that perhaps no other living man could have matched. But on the theoretical side, he tried to get away too cheaply. And this is the main cause of all the present trouble'.

Bij de bekende leerpsycholoog Thorndike trof Spearman onvoldoende inzicht in eenheden van kennen aan. Hij zei:

'... for him, the mind – like the brain as he conceived it, was composed of infinitely numerous minute elements connected together by association, now presented under the name of 'bonds' (geciteerd in Deary et al., 2008, p. 124).

Spearman aarzelde om de G definitief te interpreteren (1927, p. 75): *Eventually, we may or may not find reason to conclude that G measures something that we can appropriately call intelligence.* Als redenen, oorzaken voor de hoge G-ladingen van tests noemde hij de *Apprehension of experience*. Dit verwijst naar het feit dat we ons bewust zijn van de waarneming van kenmerken van objecten en gebeurtenissen. Verder *Eduction of relations* de activiteit waardoor we zinvolle relaties tussen twee kenmerken aanbrengen. We zien

meteen dat er een relatie is tussen 'tak-boom-stam'. We zien dat echter niet onmiddellijk tussen 'Prinses Irene von (of 'zu', daar wil ik afwezen) Lippe Biesterfeld' en 'Relatiebemiddelingsbureau voor Alleenstaande Bomen'. Als laatste oorzaak/reden noemt hij *Eduction of correlates*. Dit drukt uit dat we bij een kenmerk en een relatie onmiddellijk correlerende kenmerken of relaties kunnen opnoemen, bijvoorbeeld 'hoog - laag'.

De G-factor komt het meest tot uitdrukking in tests die *Eduction of relations and correlates* vereisen. Deze vragen van de persoon om inductief (van het bijzondere, concrete naar het algemene) en deductief (van het algemene naar het bijzondere) te redeneren. Dit verschijnsel wordt nog steeds onderzocht, bijvoorbeeld in het werk van Johnson-Laird (1999). Hij onderzocht langdurig deductief redeneren analyseerde en gebruikte daarvoor items uit de Wechsler-schalen, bijvoorbeeld: 'wat is het verschil tussen liefde en affectie' en 'wat is de overeenkomst tussen beloning en straf'?

De G-factor wordt omschreven als een mentale kracht, zoiets als paardenkracht en is '*... determined innately; a person can no more be trained to have it in higher degree than he can be trained to be taller*' (Spearman geciteerd in Deary et al., 2008, p. 126). Een meta-analyse (k = 4; N = 3018) van Te Nijenhuis et al. (2014) leidt tot eenzelfde lezing van de G: IQ winst bij adoptiekinderen die in een stimulerende omgeving terecht kwamen was niet zichtbaar op de G: '*...adopted children do not demonstrate any gain in G*' (p. 60). Testscores met een hoog G-gehalte gaan achteruit volgens Te Nijenhuis et al. door *inbreeding*. Het verschijnsel doet zich voor bij verschillende etnische groepen en correleert met reactietijd en hersenvolume. De auteurs volgen hiermee Rushtons (1999) en Rushton en Jensens (2010) lezing van de koppeling van de G aan onze biologische uitrusting. Het betreft echter steeds een interpretatie van de gemeenschappelijke variantie van een aantal tests. Deze laat - als gezegd - steeds een sterke eerste factor zien. De (sub)tests die een hoge lading hebben op die eerste factor noemt men G-geladen. De vooronderstelling bij deze interpretatie is dat biologische en omgevingscomponenten niet een statistisch artefact zijn als werkelijk onderscheiden kunnen worden.

Spearman's twee factor theorie is de basis geworden voor het omschrijven van intelligentie sinds de introductie in 1904. Mee door zijn toedoen werd factoranalyse een belangrijke techniek om gemeenschappelijk variantie in items te ontdekken. Het werd ook een methode om kenmerken van theoretische constructen te toetsen. Steun voor de G-factor lezing wordt gezien in het feit dat verschillende tests positief correleren - de *positive manifold* (Carroll, 1993). Dat suggereert het bestaan van een 'onderliggende' gemeenschappelijke 'kracht' of 'macht'. De G is of wordt verondersteld stabiel (te zijn) over testbatterijen en niet gevoelig voor verschillende soorten factoroplossingen, bijvoorbeeld centroïd of Kaisers varimax-criterium en toegestane correlaties tussen factoren, bijvoorbeeld orthogonaal of oblique (scheef). De G-factor is een voorspeller van uiteenlopend criteriumgedrag en onderdeel van verschillende cognitieve taken. Hij hangt samen met gedrag dat losstaat van wat we onmiddellijk met intelligentie of mentale macht, kracht verbinden, zoals beslissen, reactie- en inspectietijd.

Kritiek De *positive manifold* leidt niet noodzakelijk tot één factor en verwijst niet naar een functionele eenheid. De ingrediënten voor de G zijn betrekkelijk willekeurig gekozen en zijn nooit compleet. Er is ook geen definitieve statistische test of er één factor is. In feite wordt meestal rond de 50% van de variantie in criteria verklaard met behulp van de tests met een hoog G-gehalte. Messicks *consequentiality* is eveneens als punt van kritiek opgevoerd want als ieder individueel verschil construct kan het tegen groepen in de samenleving worden gebruikt.

Interpretatie van de G-factor Als te verwachten is er voortdurend debat over de *interpretatie*, denk aan de studies van Te Nijenhuis et al., Rushton en Rushton & Jensen hierboven. Voorbeelden:

(1) Kanazawa (2004, 2008, 2010) interpreteert de G-factor als een specifiek informatieverwerkingsmechanisme, een probleem-oplos-module die in de evolutie tot stand is gekomen. Dit mechanisme creëert gelijkheid tussen mensen en vormt een contrast met dieren. Wij hebben het, zij niet. Tegelijkertijd zijn er binnen de soort individuele verschillen in het beschikken over dat gelijke mechanisme. Hij beweert dat het domein waar de algemene intelligentie op past *evolutionary novelty* is. Een dergelijke hypothese kan alleen ondersteund worden door *circumstantial evidence*. Penke et al. (2011) beweren dat de hypothese te ver gaat: De G-factor is een psychometrisch construct dat een matrix met positieve correlaties redelijk weergeeft. Zo'n variabele is niet noodzakelijk verbonden met een specifieke modulaire adaptatie. Ze noemen het voorbeeld van hardlopen: het is geassocieerd met het hebben van twee benen maar ook met het cardiovasculair systeem, met het motief om meer en harder te trainen, enzovoort. Dat zijn alle adaptaties die op zich staan. Peters (2013) wijst het idee af dat mentale structuren ingeboren en domein-specifieke mechanismen van informatieverwerking zijn, gevormd uit de noodzaak specifieke evolutionaire problemen op te lossen. Het centraal zenuwstelsel is een lange weg gegaan gedurende de natuurlijke selectie om de structuur te worden die het is. Biologen en neurologen suggereren dat we zowel domein-specifieke als domein-aspecifieke neurobiologische systemen overgeërfd hebben, waarbij aangeboren en door de omgeving verworven mechanismen betrokken zijn: Lamarck mag tegenwoordig weer een beetje: aangeleerde kenmerken kunnen doorgegeven worden. Evolutionaire hypothesen klinken aannemelijk, maar zijn niet direct aan te tonen. De biologen en neurologen gaan uit van voortdurende interacties met de omgeving. Ze leggen de nadruk op neuroanatomie en neurobiologische functies en verwerpen aangeboren informatieverwerkingsprocessen en Cartesiaans gereedliggende modules.

(2) De psychometrici Borsboom en Dolan (2006) interpreteren de G als een latente trek, een theoretisch attribuut dat voldoet aan vereisten van een psychometrisch model. Het veroorzaakt verschillen in activiteiten die we intelligente handelingen noemen. De trek is een resultante van persoon x omgeving interacties.

Intelligentie wordt kortom verschillend geïnterpreteerd en dat zal zo blijven. Het is een speeltuin van lezingen hoe menselijk kennen en leren door psychologen wordt opgevat. Is het een psychometrisch construct (exploratief komt intelligentie uit een analyse over

samenhangen), een latente trek, een informatieverwerkende systeem, al of niet in de evolutie tot stand gekomen om te overleven, is ze zowel specifiek als aspecifiek en berust dat op onze biologische en neurologische hardware?

De vraag over het aantal factoren dat nodig is om rekenschap te geven van de *positive manifold* is verschillend beantwoord door de VS psychologen Thurstone en Guilford.



Louis Leon Thurstone (1887-1955) was wiskundige en werkte in een van de laboratoria van Thomas Edison, de uitvinder van elektriciteit en vele apparaten. Hij zag dat Edison niet goed was in wiskunde, maar wel creatief, slim, inventief en commercieel. Dat was onder meer aanleiding om te denken dat er verschillende intelligenties moesten zijn. Dit bracht hem mogelijk op de zeven *Primary Mental Abilities* (PMA).

Thurstone leefde in een land met veel immigranten die verschillende talen spraken, allerlei beroepen hadden uitgeoefend, nieuwe probeerden te leren en uiteenlopende vaardigheden meebrachten. Naast de vondst van de PMA heeft hij bijgedragen aan *attitude scaling* (*Law of comparative judgment* en *Law of equal appearing intervals*, zie Edwards, 1957). Deze methoden zijn voorlopers van het Rasch model. Hij wijst de Spearman G- en S-factoren niet af. Hij beschouwt de G eerder als een abstracte factor die 'vrij' boven interessante, inhoudelijke factoren 'zweeft'. Bij hogere-orde analyses komt je altijd op een algemene G uit. Hij onderscheidt zeven inhoudelijke factoren. Dit is het begin van de Amerikaanse *multiple intelligence* traditie. Thurstone respecteert de grote variatie in vaardigheden en probeert de wezenlijke eruit te halen. In 1941 liet hij door een legertje rekenaars - er was nog geen PC - 56 tests analyseren. Hij wilde factoren trekken die gelijke items verbonden en verschilden van items op de andere factoren.

Spearman trok een eerste factor die zoveel mogelijk variantie verklaarde. Thurstone roteerde de factoren naar *simple structure*'. Deze is eenvoudig te interpreteren want het is een profiel van zeven onafhankelijke factoren. Elk item laadt hoog op één factor en laag op alle andere. Een genuanceerde interpretatie van de *Big Five* (AB5C-model) berust precies op het betrekken van ladingen van telkens twee factoren bij de interpretatie. Het is misschien eenvoudiger om factorzuivere intelligentie-items te schrijven dan factor-zuivere persoonswoorden. Intuïtief ligt het echter voor de hand dat bij het beantwoorden van intelligentietaken veel vaardigheden en kennis betrokken zijn. Sommige veroorzaken in een steekproef geen variantie, bijvoorbeeld een snelheidstest met letters en getallen bij middelbare scholieren. Deze zijn zo perfect geleerd dat ze geen verschil meer maken tussen deze leerlingen. De zeven primaire mentale vaardigheden zijn:

I Verbal Comprehension

II Word Fluency

III Number

IV Space

V Associative Memory

VI Perceptual Speed

VII Reasoning

Uit het voorafgaande blijkt dat de keuze voor het type factoranalyse niet neutraal is. Het meest gebruikte type - de *default* optie in SPSS - gaat uit van een aantal orthogonale (loodrechte), onafhankelijke, ongecorrleerde factoren. Er wordt verondersteld dat de gegevens zo adequaat beschreven worden. Deze veronderstelling leent zich voor confirmatieve factoranalyse. De AB5C-aanpak bij de *Big Five* wijkt daarvan af. Later is aanbevolen om te beginnen met scheve rotaties en daarna te bekijken of de orthogonale oplossing een goed alternatief is (Fabrigar et al., 1999).

Guilford (1897-1987) was geïnteresseerd in creativiteit. Hij was hoogleraar aan de Universiteit van Kansas en hielp om rekruten te selecteren en personeel te (her)plaatsen. Hij liet de G-factor los en werkte zijn SOI, *Structure Of Intellect* uit.



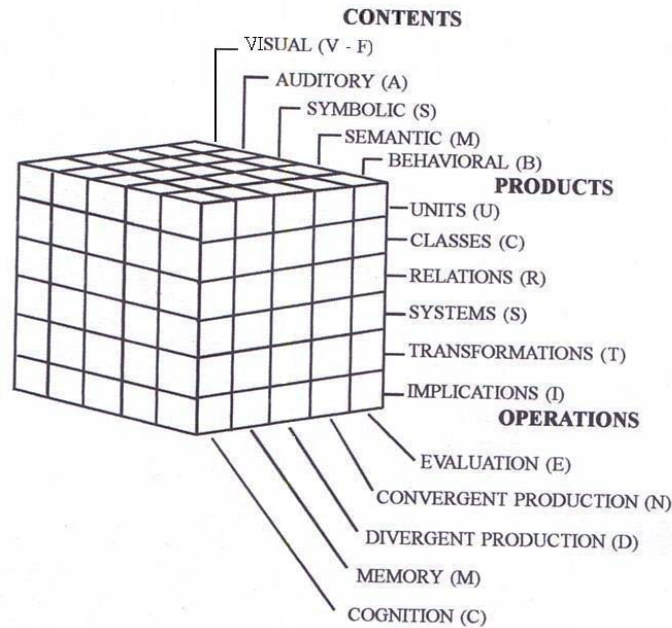
John P. Guilford: Structuur van de Intelligentie: 120 of zelfs 150 factoren.

De eerste versie bevatte 120 soorten intelligentie en wordt gepresenteerd als een rechthoekige doos met ribben van 4 : 5 : 6 (Figuur 1). Intelligentie omvat:

Vier typen inhoud: figuren, symbolen, semantische betekenissen gedragingen.

Vijf soorten operaties: cognitie, geheugen, convergente productie, divergente productie en evaluatie.

Zes typen producten: eenheden, klassen, relaties, systemen, transformaties en implicaties.



Figuur 1: Guilford's *Structure of Intellect*. De rechthoekige doos bevat inhoud, operaties en producten (4 × 5 × 6) en levert 120 soorten intelligentie op. Elke cel vertegenwoordigt een specifiek soort intelligentie.

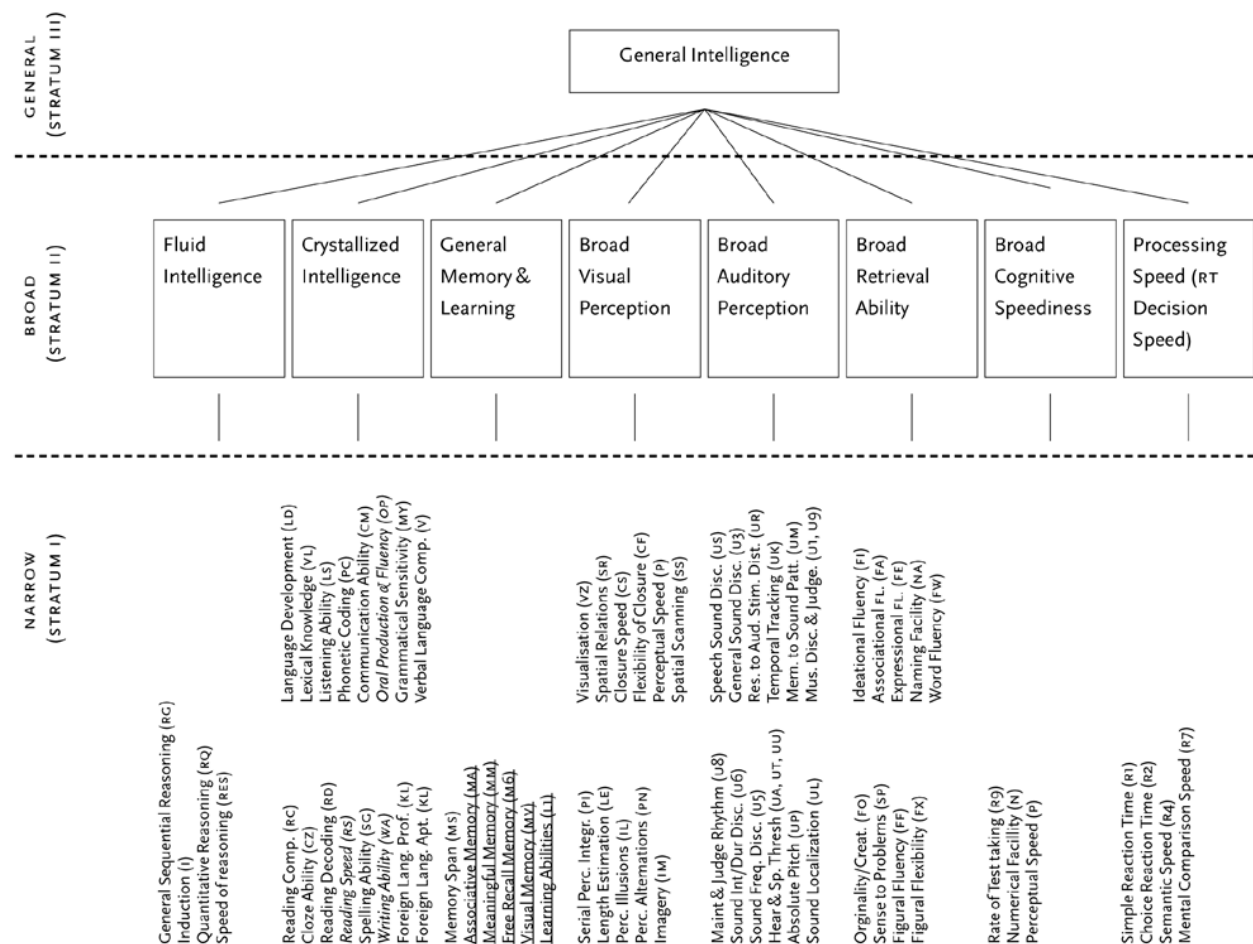
Deze indeling heeft het construeren van nieuwe tests gestimuleerd. Vooral onderzoek naar divergente productie: creativiteit, Guilford's oorspronkelijke interesse, kwam op gang. Er zijn pogingen gedaan om tests te maken voor elke cel in de doos. Het is niet duidelijk of Guilford dat ook voor ogen stond. Volgens Cronbach en Snow (1977) was dat zijn bedoeling niet maar Jensen (1998) beweert van wel. Later heeft Guilford nog eens dertig cellen toegevoegd (5 × 5 × 6). Inhoud werd nu uiteengelegd in gedrag, semantiek, gehoor, symbolisch en visueel. Het aantal operaties en producten bleef gelijk.

De 120 en 150 soorten hoeft je niet op te vatten als een (structurele) theorie. Je kunt ze ook beschouwen als een heuristisch om verschil te maken tussen de vele vaardigheden en bekwaamheden die betrokken bij het oplossen van de grote diversiteit aan problemen. Het model is niet populair meer en is misschien wel het slotakkoord van de multiple intelligentie traditie.

Ordering van intelligentiefactoren De groei van 1 naar 7 en naar 120 en 150 factoren lokt pogingen uit orde te brengen in de veelheid. Vernon (1950, 1961) was een van de eersten die een hiërarchisch model uitwerkte. Vernons model is: Eerst G: Spearman's G-factor; daaronder twee brede groepsfactoren: *Verbal Educational* (bijvoorbeeld Nederlands, Geschiedenis) en *Spatial Mechanical* (bijvoorbeeld auto's repareren); weer daaronder smalle groepsfactoren: *Perceptual-practical abilities*, *Creative*, *Psychomotor*, *Physical*, *Technical abilities* en ten slotte specifieke groepsfactoren: *Reading*, *spelling*, *linguistic*, *numerical*, *science abilities*.

Carroll (1993) deed een poging om de definitieve factoranalyse op tot dan bekende intelligentietests uit te voeren. Hij stelt een hiërarchisch model voor dat drie strata (lagen)

bevat. Het is op te vatten als een overzicht van psychometrische intelligentieonderzoek tot 1993 (Figuur 2).



Figuur 2: Het drie-lagen concept van intelligentie van Carroll (1993). Stratum I-factoren zijn gedifferentieerd als niveaus (levels zonder kenmerken), snelheid en niveau (cursief) en snelheid (onderstreepte factoren). Verklaring van de tekens: G is algemene bekwaamheid; V:ed zijn verbaal-educatieve bekwaamheden; K:m zijn ruimtelijke-perceptuele-praktische bekwaamheden.

De hiërarchie is als volgt opgebouwd: De toplaag of de bron is algemeen: de G: *General intelligence* (Stratum III). De middenlaag is breed en bevat acht factoren: *Verbal ability, reasoning, memory and learning, space, auditory perception, ability to retrieve from memory, cognitive speed* en *processing speed* (Stratum II). De bodemlaag is smal en bevat veel specifieke cognitieve vaardigheden, zoals kwantitatief redeneren, begrijpen van teksten, absolute toonhoogte die men waarneemt (Stratum I).

De meeste tests bevatten taken voor verbale, redeneer-, geheugen- en ruimtelijke vaardigheden. Na analyse van 371 factoren van *Verbal ability* kwam Carroll tot productieve en receptieve beheersing van mondelinge en geschreven taal. Op dezelfde wijze omschreef hij *reasoning ability* als sequentieel of deductief, inductief en kwantitatief redeneren. Dit lijkt op Horn en Cattells vloeibare intelligentie (Gf). De geheugen- en leerfactor komen minder overtuigend uit de meta-analyses te voorschijn maar Carroll hield er aan vast. Vijf

specifieke factoren vullen de volgende laag: geheugenspanne, associatief geheugen (het vermogen om willekeurige associaties tussen stimuli en objecten te leggen en ze te reproduceren), vrije weergave van niet-gerelateerd materiaal, reproductie van zinvol verbonden materiaal en visueel geheugen.

In ruimtelijke vaardigheid worden vijf factoren onderscheiden: het vermogen om visuele patronen onder tijdsdruk te manipuleren, snelheid van manipulatie van eenvoudige patronen, waarnemingsnelheid om nieuwe en bekende visuele patronen te vinden die verborgen zijn in afbeeldingen en herkennen van een bekend visueel patroon.

Factoranalyse kan geen definitief antwoord geven over aantal en aard van de factoren. De analyse van Carroll stelt echter wel orde op zaken. Allerlei factoren die al eerder genoemd zijn, worden in een kader gezet: G en S, PMA, SOI en informatieverwerking. Het laatste is als interpretatie een onderdeel van IQ-tests geworden. Een voorbeeld is Kaufman en Lichtenbergs (2000) interpretatie van de Wechsler-schalen voor volwassenen en kinderen: WAIS, WISC en WIPPSI.

De discussie over het ene of het vele is een oud. Ik maak een zijstap naar de uitwerking van deze vraag in de geschiedenis. Dat we ons nu nog bezighouden met deze vragen, rechtvaardigt het om te zeggen: *we are still confused, but at a higher level.*

Box G, S, PMA, SOI: Het ene of het vele?

Thales van Milete (ca. 624-546 BCE) vroeg zich af wat nu wel het principe was van al wat is: 'de G-factor van het Zijnde'. Wat is eerst, wat is de oorsprong, waar gaat alles naar terug? As tot as, stof tot stof? Het is niet de verzameling van alle verschillende voorwerpen en mensen, er moet iets algemeen aan ten grondslag liggen. Ja, en dat is water. Bij zo'n eenvoudig antwoord kon Thales kritiek verwachten. Maakt het dan niet uit of je in de Himalaya woont of in de woestijn? Is er niet meer: naast aarde, water, lucht en vuur? Of je maakt er een grap over: als je met je geliefde slaapt, heb je dan ongeveer 40 liter water in bed? Of de Thracische vrouw die Thales uitlachte toen hij de hemel bestuderend in een waterput viel? Later werd het aarde, vuur, lucht en water. Deze elementen zijn het tot ver na de middeleeuwen gebleven, zo wordt het steeds verteld, maar de Grieken kenden al meer elementen. Vooral dure materialen, zoals goud kenden ze al.

Demokritos (ca. 460-370 BCE) ging ervan uit dat alles was opgebouwd uit deeltjes die niet verder gedeeld konden worden: atomen. Alles, mensen, dieren, planten en de stoffelijke werkelijkheid was opgebouwd uit atomen. De ziel had weliswaar de gladste atomen maar ze bleef een bundel atomen. Dit is de eerste verschijning van het atomisme en materialisme, waar zogenoemd hogere mentale functies als bewustzijn, de ziel, het bundels van weliswaar mooie atomen zijn.

Descartes wees dit af. Hij isoleerde het bewustzijn van de objecten, van de wereld. Nu gebruiken we het periodiek systeem van Mendelejev en gaan ervan uit dat alles is opgebouwd uit chemische elementen en hun verbindingen: de mens bestaat bijvoorbeeld voornamelijk uit zes belangrijke elementen. Volgens een Amerikaanse geleerde mevrouw is er een zevende: arseen bijgekomen, maar dit kon wel eens een hype zijn. Het doet zich in zo kleine concentraties voor dat het niet uitmaakt. Beoefent zij de homeopathie van de chemische geologie? Atomen bleken later deelbaar, want het gaat verder: neutronen en protonen, enzovoort. Er is voor natuurwetenschappers geen hiërarchie in de kleinste deeltjes.

De één-of-veel vraag leidt niet tot logisch onmogelijke beweringen of antwoorden. Dat zou zich voordoen als dat ertoe zou leiden dat een identiek object of subject één én veel is onder hetzelfde perspectief. Het ene bijvoorbeeld een object of een stofje staat wel veel elementen toe, als in de scheikunde, of veel differentiaties als in de biologische evolutie en bij menselijke gedragingen.

De bijdragen van Spearman, Thurstone en Guilford stimuleerden onderzoekers om langs empirische weg de definitieve intelligentiestructuur te vinden. Bovendien leidden ze tot het ontwikkelen van nieuwe tests voor onvermoede kanten van de intelligentie. Ze gebruikten daarvoor, op Guilford na, items van bestaande tests.

Een nieuwe generatie bracht het intelligentieonderzoek verder met behulp van informatieverwerkingstheorie en moderne testtheorie. Ze gebruikten niet alleen al bestaande items. Ze ontwierpen taken die pasten bij hun intelligentieconcept. Twee voorbeelden worden vermeld.

Robert Sternberg sluit aan bij informatieverwerking. Deze oriëntatie op kennen en cognitie kwam onder de aandacht toen communicatie complexer werd door gebruik van telefoon en telex. Commerciële instellingen zoals de Bell Company in de VS subsidieerden bijvoorbeeld onderzoek naar signaal en ruis in communicatie. Hij onderscheidt verschillende informatieverwerkende processen in intelligent gedrag, zoals coderen, gevolgtrekkingen maken, bepalen van relaties tussen brokjes informatie, het ordenen van relaties in een hiërarchie, het toepassen van kennis, het denken over alternatieven, het vergelijken van oplossingen en het rechtvaardigen van een gekozen oplossing. Hij onderscheidde deze processen na het uiteenrafelen van handelingen en stappen die nodig zijn voor het maken van taken en uitvoeren van opdrachten. Sternberg heeft daar onder meer analogietaken voor gebruikt, waarbij een persoon vragen van het volgende type moet beantwoorden:

.... staat tot vaardig, als ongeschoold staat tot...

1. stuntelig 2. getuigen 3. blindheid 4. aap
- a. bekwaam b. dom c. lelijk d. links

Naast informatieverwerkingsstappen en -strategieën onderscheidt hij meta-componenten. Deze hebben betrekking hebben op het inzicht wélk probleem opgelost moet worden en hóe de componenten achtereenvolgens of tegelijkertijd ingezet moeten worden en wánnere er gestopt moet worden. De theorie vereist metingen van de componenten en strategieën om de veronderstelde relaties te toetsen. Dat is niet eenvoudig. Sternberg vertelde dat zijn verwachtingen hoger gespannen waren dan de resultaten wettigden. Hij denkt dat zijn benadering niet in de praktijk ingevoerd zal worden, omdat het meten van de componenten te veel tijd kost. De kracht van een theorie is afhankelijk van de kwaliteit, de voorbeeldigheid van de operationalisaties. In 1993 ontwierp hij de STAT, een test die bijna alle componenten dekte. Hij maakte negen meerkeuzesubtests met vier items. Ze representeren drie domeinen: het analytische, creatieve en praktische. Elk domein werd op

drie manieren gemeten: verbaal, kwantitatief en met behulp van afbeeldingen en figuren. Naast meerkeuze-items liet hij opstellen schrijven voor de drie domeinen:

- (a) Analytisch-verbaal: De persoon ziet een nieuw niet-bestaand woord (een neologisme) in een tekst en moet de betekenis ontdekken vanuit de rest van de tekst.
- (b) Creatief-afbeeldingen, figuren: De persoon ziet een reeks figuren die veranderen; hij moet de regel ontdekken die de veranderingen verklaart of genereert.
- (c) Praktisch-kwantitatief: de persoon moet een wiskundeopgave oplossen die in het dagelijks leven voorkomt; bijvoorbeeld koop kaartjes zo goedkoop mogelijk in voor een groep, waarvan de leden verschillende prijzen voor een kaartje kunnen betalen.

Daarnaast schreven de deelnemers een opstel over een thema, bijvoorbeeld: 'Ontwerp een nieuw curriculum voor toegepaste psychologie, dat beter is dan bestaande'. De meerkeuzevragen en opstellen werden gescoord en gefactoranalyseerd. Hij vond een zwakke algemene factor en er was methodevariantie: de meerkeuzevragen hingen bijvoorbeeld samen, afgezien van het domein waar ze betrekking op hadden. De drie domeinen waren niettemin te onderscheiden. De test voorspelde schoolprestaties niet beter dan een algemene intelligentie test (WAIS, WISC of Raven). Een belangrijk criterium werd dus niet beter voorspeld met deze theoriegeleide informatieverwerkingstest. De test biedt wel een gedifferentieerd en subtiel beeld van intelligent handelen. De aanpak en het onderzoek laten zien dat de twee oriëntaties: individuele verschillen en de gemanipuleerde context bij onderzoek naar informatieverwerking elkaar niet versterken of aanvullen bij het prediceren van criteria.

Susan Embretson (1998) startte eveneens met een informatieverwerkingskader. Als psychometricus aanvaardde zij alleen constructen die unidimensionele (Rasch) schalen vormden. Een test was bedoeld om abstract redeneren te meten. Ze legde dat uiteen in processen zoals deductie, inductie, hiërarchisch categoriseren (bijvoorbeeld met elkaar insluitende, steeds omvattender categorieën), gebruik van logische regels (bijvoorbeeld $A > B, B > C \rightarrow A > C$) en het gebruik van 'alsdan' regels. Ze onderscheidde facetten aan abstract redeneren. Haar aanpak lijkt op Guttmans (1971) voorstel voor een structurele theorie. Daarin wordt een poging gedaan meten te passen op de veronderstelde structuur van de aspecten (facetten) van intelligentie. Er werden voorspellingen gedaan over relaties van abstract redeneren met criteria. De facetstructuur werd getoetst met confirmatieve multivariate technieken. Als items niet aan het meetmodel voldeden, werden ze vervangen door nieuwe. De toets voor de factoroplossing (een stressmaat) moest aanvaardbaar zijn. Daarna werd de voorspellende waarde van het theoriegeleide instrument bepaald door de resultaten in een steekproef te correleren met cognitieve prestaties. Deze aanpak leidde ook niet tot hogere predictieve validiteitscoëfficiënten dan klassieke tests. Embretsons stappenplan om individuele verschillen in abstract redeneren te verklaren aan de hand van informatieverwerkingscapaciteiten ziet er als volgt uit:

- a. Exact omschrijven van de doelen: wat te meten?
- b. Beschrijven van alle taakkenmerken (het wat: het domein), het hoe (bijvoorbeeld algemeen of specifiek, onder welke condities, en welke vorm wordt gekozen om de vragen, items aan te bieden).
- c. Ontwerpen van een cognitief model (stappenplan) met gebruikmaking van theorie van informatieverwerking, toetsen van de schalen met IRT, veranderen model.
- d. Evaluatie van het cognitieve model en van het niveau van complexiteit. De test kan geconstrueerd worden uit de items die voldoen aan het gekozen psychometrisch model.
- e. Schrijven van taken, items die aan psychometrische eisen voldoen. Vaak voldoen items niet en moeten vervangen worden. Dit vereist conceptuele analyse van informatieverwerking bij ieder item.
- f. Evaluatie van de cognitieve en psychometrische kenmerken van de items en revisie als niet aan de vereisten van beide kenmerken is voldaan.

Dit stappenplan lijkt vruchtbaar voor het ontwerpen van theoriegeleide tests en het is inderdaad een enkele keer benut, bijvoorbeeld door Primi (2014) om een test voor Gf (*Fluid intelligence*) te construeren. Items werden ontworpen met als kenmerken: Hoeveelheid informatie (aantal elementen en regels), Soort regels (inhoud van de relatie) en Perceptuele complexiteit (het in het oog springen van de elementen). Deze sluiten aan bij de Gf componenten Werkgeheugen (visueel ruimtelijk; verbaal auditief), Uitvoering (*executive*): doel controle, coördinatie van gelijktijdige mentale activiteiten, monitoren en overzien van mentale activiteiten, adaptieve en flexibele aanpassing) en Abstractie (selectieve aandacht). De item-moeilijkheidsgraad werd voorspeld door de structuurkenmerken van de items. Het stappenplan en de theorie over Gf werd met (enig) succes getoetst bij 2.797 undergraduate studenten van een privé universiteit van São Paulo (Brazilië).

De voorbeelden zoals uitgewerkt door opvolgers van de psychometrische intelligentietheoretici Sternberg en Embretson laten dus zien dat een nieuwe generatie cognitieve- en intelligentieonderzoekers aansluit bij de contemporaine opvattingen over intelligentie - bijvoorbeeld het Gf-Gc onderscheid en informatieverwerking en over meten: IRT. De uitwerking van bijvoorbeeld Primi is gesofisticeerd en vergt samenwerking van psychometrici en intelligentietheoretici. Het resultaat vindt schoorvoetend zijn weg naar de testgebruiker. Niettemin is Spearman's opvatting verbreed want er worden verwerkingsprocessen aan toegevoegd en wordt moderne testtheorie te hulp geroepen om betere tests te maken. De IRT is duidelijk over de meetschaal. Dit hoeft en levert ook niet winst op voor predictieve validiteit. Deze auteurs proberen theoriegeleide instrumenten te ontwikkelen maar ze vinden hun weg naar de praktijk nauwelijks.

Sternberg (1991, p. 263-265) vroeg zich toen al af of zulke inspanningen de moeite waard zijn. In een artikel met de titel: *Death, taxes and bad intelligence tests* concludeert hij dat net als bij de twee andere verschijnselen slechte tests onvermijdelijk lijken.

Pragmatiek: de Wechsler schalen De aanpak van David Wechsler laat zien dat er naast de theoretische *top-down* benaderingen er een *bottom-up* benadering is die heeft geleid tot de

best verkochte test ter wereld. Bij *bottom-up* geconstrueerde tests wordt *achteraf* een verbinding gelegd met theorievorming, die op dat moment populair is. De factoren - eerst verbaal en perfoormaal IQ - van de WAIS en WISC worden nu geïnterpreteerd als informatieverwerkingscapaciteiten. Er zijn studies waarin IQ met zulke processen verbonden wordt:

(1) Danner et al. (2011) gebruikte bij 173 werknemers van enkele bedrijven twee IQ tests, twee dynamische beslistaken en twee impliciete leertaken. Als verwacht waren alle gecorreleerd met het IQ. Snelheid van informatieverwerking werd opgevat als het centrale intelligentieproces. Het bond een substantiële hoeveelheid variantie in de tests. Het IQ voorspelde inkomen, sociale status en hoeveelheid en kwaliteit van genoten onderwijs. De auteurs vonden daarnaast dat resultaten van dynamische beslistaken de oordelen van de chefs over de kwaliteit van de werknemers beter voorspelden dan hun IQ.

Werkgevers waarderen eerlijkheid, gehoorzaamheid en loyaliteit aan het bedrijf meer dan het IQ. Chefs en managers waarderen daarnaast flexibiliteit en snel kunnen beslissen. De hier voorgestane rijke beschrijving van het 'probleemveld' is bij managers, chefs en bestuurders doorgaans te veel werk. Ze hebben belangrijkere dingen aan hun hoofd. Een wrang voorbeeld is het beoordelingssysteem van de NZa: de Nederlandse Zorgautoriteit (ontleend aan Gotlieb, 2014: de ambtenaar die zelfmoord pleegde na tegenwerking en onvoldoende beoordelingen die postuum ongeldig verklaard zijn). Het is een kwadrant met twee kenmerken die beoordeeld worden door een chef. Er vindt geen betrouwbaarheidsonderzoek plaats en er valt de facto niet - de jure wel - tegenin te gaan. De dichotome kenmerken van de werknemer zijn: Toegevoegde Waarde (TW 0,1) en Groei Potentieel (GP 0,1). Dit resulteert in 'learners' (TW 0, GP 1), 'werkbijen of 'bakens' (TW 1, GP 0), 'High Potentials: hipo's' (TW 1, GP 1) en 'achterblijvers' TW 0, GP 0). De laatste werden in NZa-terminen lieden die de payroll bevullen genoemd.

(2) Vock et al. (2011) omschreven als cognitieve kernkwaliteiten: redeneren (een onderdeel van de Binet-test), divergent denken (creativiteit uit Guilfords SOI-model), mentale snelheid (MS, al door Spearman gesuggereerd) en korte-termijn-geheugen (STM dat al sinds Plato en daarna de Middeleeuwen met intelligentie in verband is gebracht). Zij beschouwden de complexe capaciteiten van redeneren en divergent denken als bemiddeld, 'gemedieerd' door simpele capaciteiten als STM en MS in een steekproef van 1135 adolescenten van 12 tot 16 jaar. Een ingewikkeld structureel vergelijkingsmodel ondersteunde de veronderstelling dat de simpele vaardigheden STM en MS de complexe medieerden: STM en MS hadden een indirect effect van $\beta = .22$ en $\beta = .24$ op prestaties op de complexe vaardigheden. STM had ook een direct effect op die prestaties: $\beta = .22$. STM is een subtest van de WAIS en de WISC en daarom is de relatie niet verrassend. Je zou zelfs een hogere β -waarde verwachten.

(3) Benson et al. (2013) benutten een normeringssteekproef van 730 subjecten van 6 tot 16 jaar. De steekproef was gestratificeerd naar leeftijd, (stem)regio, sekse, opleidingsniveau van de ouders en etniciteit. Ze gebruikten CFA bij hun procesbenadering van intelligentiefactoren en het Horn-Cattell onderscheid tussen Gf en Gc. De WISC-IV-Integrated versie (Wechsler et al., 2004) werd afgenomen. Deze bevat alle kern en toegevoegde subtests van de IV-versie. Het resultaat was een *positive manifold* die liet zien dat de WISC-IV Integrated de G-factor goed representeerde. Het Gf-Gv onderscheid was behulpzaam bij het interpreteren van de subtests: *Block Design* en *Elithorn*

doolhoven bleken Visuele informatieverwerking te meten (Gv). De Perceptueel Redeneren subtests maten beide: Gc en Gf. Geheugenspanne en Werkgeheugen zijn facetten van *Memory*.

(4) Nelson et al. (2013) onderzochten de structuur van de WAIS-IV in een klinische steekproef: 300 studenten die naar de universiteitskliniek gingen voor een *comprehensive psycho-educational evaluation*. De meeste subjecten waren 1^{ste} - 3^{de} jaars (75%) en de meesten blank (86%). De gemiddelde leeftijd was 22.6 jaar, SD = 7.56. Ze hadden de diagnose: Leermoeilijkheden (LD), Altijd Druk Heel Druk (Attention Deficit Hyperactive Disorder: ADHD) en co-morbide LD/ADHD. Hoewel alle studenten de middelbare school doorlopen hadden en tot de universiteit waren toegelaten was hun IQ gemiddeld: WAIS hele schaal: 97.74. De scores op de factoren waren: Verbaal: 102.94, Perceptueel 99.41, Werkgeheugen: 93.91, Verwerkingssnelheid: 94.02. De SDs bleven binnen het bereik van een normeringssteekproef (12.00 tot 13.62) SD is in de normeringssteekproef gesteld op 15.00. Gegeven deze kengetallen is het een normale representatieve steekproef. De CFA ondersteunde de structuur van de WAIS-IV. In een hiërarchisch model was er een eerste algemene factor die 32.9% variantie bond van de 52.9% gemeenschappelijke variantie. Voor de verbale factor was dat 10.9% en 17.4%, voor perceptueel redeneren 2.8% en 4.5%, voor Werkgeheugen 5.7% en 9.2% en voor Perceptuele snelheid 10.2% en 16.4%. De CFA liet een sterke eerste factor zien: de G. De processen kunnen gezien worden als tests voor specifieke informatieverwerking: de Spearman S factor. Hoewel deze steekproef binnen het normale bereik ligt is er een (dun) spoor van stoornis te ontwaren: verbaal is gemiddeld maar werkgeheugen en processnelheid ligt ongeveer 0.5 SD onder het gemiddelde van 100. Overigens is dit de groep waarvoor Wechsler zijn test oorspronkelijk gemaakt heeft als medisch psycholoog in het New Yorkse Bellevue Hospital.

De houding van WAIS onderzoekers en -gebruikers is flexibel: het oorspronkelijke WAIS systeem van Performaal en Verbaal IQ met sterke G factor (algemene intelligentie genoemd) kon bij werknemers gelezen worden als snelheid en flexibiliteit, als samenstellingen van andere factoren met de bemiddelende processen van mentale snelheid en STM, als de G en de Gf en Gc en als een hiërarchisch geheel met een G, verbale factor (door te Nijenhuis, altijd de sterkste factor bij IQ genoemd) en een gedifferentieerde perceptuele factor.

Oorzaken van individuele verschillen Naast vernieuwing op basis van informatieverwerking en IRT is bij analyse van intelligentie (IQ) en cognitie gespeculeerd over oorzaken van individuele verschillen. Galton zocht de oorzaak in genetische verschillen. Deze opvatting is een *come back kid*. Gedragsgenetische studies gaan over het verdelen van de geobserveerde variantie in IQs in een representatieve steekproef. De bronnen zijn genetische verwantschap (mono-, dizygote tweelingen, broers en zussen, ouders en kinderen) en verschillen tussen omgevingen: binnen gezinnen, tussen gezinnen, buurten en bevolkingsgroepen, landen). Het resultaat is meestal dat 50% (variërend van 25% tot 75%) aan genetische variatie wordt toegeschreven. De overige variantie wordt verdeeld over omgevingen: ongeveer 15% wordt gebonden aan verschillen binnen gezinnen en 5% aan verschillen tussen gezinnen. Er blijft een deel over dat we niet thuis kunnen brengen.

Eerst, wordt een benadering van Kanazawa uitgewerkt en vervolgens worden voorbeelden gegeven van opvattingen over de relatie tussen IQ en onze biologische uitrusting.

(1) Kanazawa (2008, 2010) beweert dat twee derde van de variantie in nationale IQs verklaard kan worden door *evolutionary novelty*. Dit wordt geïndiceerd door de fysieke afstand tussen de omgevingen van huidige generaties met die van ouders en voorouders. De afstand is het gevolg van mobiliteit. Daarnaast noemt hij klimaatveranderingen, vooral temperatuur. Een koud klimaat met felle winters zou individuen met een laag IQ weg geselecteerd hebben. Deze bronnen zet hij naast gedeelde omgevingseffecten die alle leden van een gezin gemeen hebben maar die verschillen tussen gezinnen, bijvoorbeeld op basis van verschil in sociale klasse. Dit effect wordt meestal laag (5%) geschat maar ook hiervoor is een tegenvoorbeeld: Haworth et al. (2010) rapporteerden in een meta-studie een waarde van .16 voor de eerste levensjaren. Behandelen ouders kinderen niet gelijk? Dit is nagegaan door geboorte-volgorde in onderzoek te betrekken: gemiddeld genomen is het IQ van de eerstgeborene 3 of meer IQ punten hoger dan het tweede kind. Deze bevinding is door de psycholoog Piet Vroon de soep wordt dunner hypothese genoemd. Hij baseerde zich op Raven IQs van broers, gekeurd voor militaire dienst, lang geleden. De gezinssamenstelling (minder kinderen!) en de behandeling van kinderen en adolescenten zijn sindsdien wel enigszins gewijzigd. De IQ verschillen binnen een gezin waren en zijn wellicht nog steeds substantieel; wel tot driekwart standaarddeviatie.

De variaties in nationale IQs zijn volgens Kanazawa voor een deel toe te schrijven aan deze *evolutionary novelty*. Hij doelt op de afstand van de omgeving van (voor)ouders en op temperatuurverschillen. Zo'n bewering kan alleen gesteund worden door *circumstantial evidence* want er is geen direct bewijs te verkrijgen. En, moet een individueel verschillen variabele noodzakelijk te maken hebben met een onderliggende adaptatie? De G-factor geeft niet direct informatie over welke domein-specifieke adaptaties er voor nieuwheid bestaan. Er zijn meer mechanismen te bedenken dan alleen evolutionaire nieuwheid.

(2) Kanazawa schaar roken en drinken onder *evolutionary novelty*. Hij heeft het gebruik van deze middelen in verband gebracht met de nationale IQs in het VK en de VS. De studie is van Kanazawa en Hellberg (2011). *Evolutionary novelty* wordt de 'Savanna-IQ hypothese' genoemd. De veronderstelling is dat intelligente personen gemakkelijker evolutionair nieuwe activiteiten en waarden verwerven dan mindere intelligente. Dit leidt tot de hypothese dat de intelligentere meer alcohol, tabak en andere drugs consumeren dan de minder intelligente. Ze konden data krijgen uit grote longitudinale studies in de VS en het VK. Ze menen steun te hebben voor hun hypothese: intelligente VS kinderen maken een grotere kans om tabak te consumeren dan hun minder intelligente landgenoten. In het VK loopt de intelligentere groep meer kans (illegale) drugs te gebruiken. De auteurs vermelden ongeveer 20 associaties: verbanden tussen demografische variabelen enerzijds en tabak-, alcohol- en drugsgebruik, anderzijds. De correlaties zijn laag en volgens Cohen verwaarloosbaar maar wel significant bij deze grote aantallen. Ter illustratie: In het VK: algemene intelligentie van kinderen is met een waarde van $r = .08$ en $.06$ gecorreleerd met tabak- en drugsgebruik ($p < .001$; p.386). De VS waarden zijn $.04$ (alcohol, $p < .001$), en $.03$ (tabak, $p < .01$). Er was geen relatie met drugsgebruik (p. 389). Zou het ook te maken kunnen hebben met het kunnen betalen van deze middelen? En, hoe wordt iemand geïdentificeerd als gebruiker? Vertelt iedereen je dat waarheidsgetrouw?

Kanazawa zoekt de verklaring van intelligentieverschillen dus onder meer in de fysieke omgevingen. Aan het organisme schrijft hij een evolutionair bepaald streven naar *novelty*

toe. Dit laatste kenmerk was een tijdlang tot de jaren 1975-1980 een kenmerk waar bevolkingsgroepen in Z. Afrika op vergeleken werden.

Er zijn directere bronnen en oorzaken van intelligentieverschillen te bedenken dan klimaat en evolutionair verleden. De biologische, fysiologische en neurologische uitrusting wordt daarbij gezien als basis voor intelligentieverschillen via processen van informatieverwerking. IQ wordt in verband gebracht met kenmerken van het centraal zenuwstelsel. Significante correlaties zouden dat ondersteunen. Er is geen discussie zoals die door Oppenheimer & Putnam en Fodor gevoerd is over de relaties tussen de niveaus van de biologische en neurologische wetenschappen en de psychologie. Er zijn geen 'brugwetten' geformuleerd. Recent is er bijvoorbeeld aandacht voor de groei van het puberbrein. Het feit van doorgaande groei en het nog niet tot stand komen van verbindingen zou deze periode tot een wisselvallige en risicovolle periode maken. Dit vergt empirisch onderzoek als bewijs. De oorzaken voor IQ-verschillen en voor processen van informatieverwerking zijn al geruime tijd gezocht in onze bio-, fysio-, neurologische uitrusting. Aanwijzingen voor oorzaken werden ontleend aan conceptuele analyses, geval- en correlatiestudies. Voorbeelden:

(1) Hebb (1949) bracht IQ-verschillen in verband met het functioneren van het centraal zenuwstelsel. Hij onderscheidde drie soorten intelligentie. Type A verwijst naar het beschikken over een *good brain* gekenmerkt door een goede neurale stofwisseling. Subjecten die hoog scoren op de Raven (een cultuurarme, G-verzadigde IQ test) en op een computerspel blijken een lagere metabolische snelheid te hebben dan zij die laag scoren. Ze hoeven zich minder in te spannen. Type B verwijst naar het gemiddeld prestatieniveau van een volwassene. Type C is de score op een intelligentietest. Het ging Hebb om type A. Zijn idee is dat intelligente mensen meer bewerkte celstructuren tot hun beschikking hebben. De basis voor IQ-verschillen ligt zo in het metabolisme en de kwaliteit van celstructuren. De studie van de verbinding tussen hersenen en informatieverwerken kreeg een impuls door gevalstudies, vooral gewonden uit de wereldoorlogen:

(2) Goldstein (1918-1965) was een van de pioniers. Hij was arts en gespecialiseerd in neurologie en psychiatrie. Hij ontwikkelde een testbatterij om hersenschade te diagnosticeren. Hij was beïnvloed door de Gestaltpsychologie en lette op de schade én op de betekenis daarvan voor patiënt en zijn sociale context. Hij gebruikte reactietijdtaken en waarnemingsnelheid om de schade aan mentale processen zoals geheugen, aandacht, optische verstoringen, alexie, agnosie en leerstoornissen vast te stellen. Hij benutte rekentaken waarbij de patiënt steeds twee opeenvolgende getallen moest optellen, gedurende een heel uur! Daarnaast werden fijn motorische taken voorgelegd. Hij was zich bewust dat wat iemand in een laboratorium toonde niet hoefde samen te vallen met het dagelijkse doen van de patiënt. Het ging hem om *lebenswahre Leistungsprüfungen* (Goldstein, 1967).

(3) De Russische fysioloog en psycholoog Luria (1902-1977) maakt onderscheid tussen drie hersengebieden met verschillende functies:

De hersenstam en de middenhersenen zijn verantwoordelijk voor 'arousal' (prikkeling, opwekking) en aandacht.

Delen buiten de hersenstam, vooral delen van de midden- en frontale hersenen zijn betrokken bij het geschikt maken van informatie voor verwerking van de zintuiglijke (sensorische) input door *encoding en processing*).

De voorhersenen (frontale cortex) zijn verantwoordelijk voor organisatie, planning en monitoren (volgen of bewaken) van gedachten en handelingen. Dit onderscheid is terug te vinden in de handboeken. Luria (1973, 1980) heeft er steun voor gevonden bij personen met hersenletsel. In de Sovjet Unie werd na de Tweede Wereldoorlog veel onderzoek verricht bij soldaten met hersenletsel.

(4) De theorie over de plaats van de functies in de hersenen wordt in *tests* gebruikt: Das en Naglieri (1995) nemen in hun intelligentietest gelijktijdig en opeenvolgend sequentieel coderen op. Er zijn twee subtests voor aandacht en planning. Ze gaan alertheid en het ordenen van gedachten en handelingen. Het instrument heet *Cognitive Assessment System* (CAS). Het idee van 'plekken, plaatsen' (loci) in de hersenen x functies is ook te herkennen in het onderscheid tussen de hersenhelften. Gazzaniga (1989) veronderstelt dat de rechterhersenhelft georganiseerd is in een aantal betrekkelijk onafhankelijk werkende modules. De linkerhersenhelft is actief bij het integreren en interpreteren van deze onbewuste processen.

(5) Niet alleen loci in de hersenen worden verbonden met IQ en informatieverwerking, ook hersenactiviteit zelf. Het elektro-encefalogram (EEG) meet de elektrische activiteit van de hersenen en levert een complex plaatje op. Dit wordt bewerkt met een computerprogramma om getallen te verkrijgen die met IQ scores, bijvoorbeeld met de Raven, gecorreleerd worden. Jensen en Reed (1992) rapporteren lage waarden ($r = -.10$ tot $-.20$), maar Kline vermeldt een waarde van $r = -.45$. Meta-studies kunnen tot een redelijke schatting van de correlatie leiden. Als onderdeel van het EEG worden *Evoked Potentials* (EVPs) op dezelfde complexe manier gemeten en aan IQ-verschillen gekoppeld. De resultaten zijn ongeveer gelijk.

Voor psychometrië is de kleinste eenheid het item op een latente schaal. Spearman begon met het idee van *apprehension of experiences* als een basisproces voor *mental power*. We zouden dat nu *encoding* noemen. Bovendien werd *mental power* verbonden met *snelheid*.

(6) Intelligentieonderzoek leidde tot nieuwe kleine eenheden om IQ-verschillen in informatieverwerking te beschrijven: *Elementary Cognitive Tasks* (ECTs). Daar worden taken mee bedoeld die processen uitlokken die nodig zijn om de taken van IQ tests uit te voeren. De processen worden aan het zenuwstelsel verbonden en staan los van IQ-items. Voorbeelden zijn het onderscheiden van twee stimuli of objecten, zo snel mogelijk kiezen tussen twee stimuli, heel snel het korte geheugen afzoeken, iets snel ophalen uit het langetermijngeheugen. Een bekende operationalisatie van een ECT is inspectietijd (IT). Op een computer worden, bijvoorbeeld gedurende zeer korte tijd twee evenwijdige lijnen van iets verschillende lengte aangeboden. De proefpersoon moet zo snel mogelijk zeggen welke de langste is. Dit wordt enige keren herhaald om een stabiele schatting van de tijd te verkrijgen die een proefpersoon nodig heeft om te reageren. Correlaties tussen de IT op deze taak en IQ variëren tussen de $r = -.30$ en $r = -.50$. De waarden zijn negatief, omdat korte reactietijd gecorreleerd is met hoog IQ (Jensen, 1998, Nettlebeck, 1987). Hutton et al. (1997) vonden een soortgelijke relatie tussen de Raven scores en IT bij 8 tot 13-jarigen. Ze controleerden voor het verschil in aandacht. Een verwante operationalisatie van elementaire processen is reactietijd (RT). Zoals gezegd suggereerde Spearman al een relatie van *mental power* en het gemak en de *snelheid* waarmee neurale circuits zich aanpassen aan veranderende omgevingen (Garlick, 2002). RT wordt ook uitgedrukt als de SD bepaald over herhaalde metingen in een reactietaak de RTSD: reactietijd standaarddeviatie. Dit drukt de consistentie van de afzonderlijke RTs op stimuli uit. Jensen (1987) vermeldt een $r = -.23$ tussen RTSD en de G-factor. De waarden lopen opnieuw uiteen: Fink en Neubauer (2001) vonden een $r = -.62$ tussen de Raven en RTSD bij 97 studenten. Ze veronderstelden dat de subjectief ervaren informatiedichtheid samenhangt met

intelligentie. Deze ervaren informatielast blijkt onafhankelijk van RTSD en draagt bij aan het voorspellen van het Raven IQ. De correlatie bedraagt $r = -.48$. In 2011 maten Klein et al. de attention *blink* (AB), dat wil zeggen het in staat zijn om twee objecten of targets te benoemen die heel kort na elkaar verschijnen. Het vermogen de twee targets te zien is met een r van $-.41$ tot $-.43$ gecorreleerd met psychometrische intelligentie.

(7) Het vermogen om complexe problemen op te lossen (*CPS ability*) is in verband gebracht met de tijd dat een subject bezig met de taak (*Time On Task: TOT*) door Scherer et al. (2015). Hautamäki leverde voor deze studie de Finse representatieve gegevens door onderzoek bij 2000 *nineth graders*. De gegevens werden in een structureel vergelijkingsmodel gepast dat CPS, cognitieve- en motivatie metingen en schoolprestaties bevatte. Lange tijd met de complexe taak bezig zijn was met een p van $.40$ gecorreleerd met CPS vermogen; *CPS-time on task* was wisselend gecorreleerd met prestatie en motivatie van de leerlingen: de β voor redeneren was $.25$ (bescheiden), voor motivatie tot presteren $-.06$ (verwaarloosbaar) en voor doelgerichtheid $.13$ (verwaarloosbaar). CPS vermogen was gecorreleerd met redeneren met een waarde van $.54$ (hoog), met motivatie tot presteren $.03$ (verwaarloosbaar), met doelgerichtheid $.03$ (verwaarloosbaar). *CPS time on task* voorspelde de schoolprestaties significant ($n = 2000$) met β s van $.12$ tot $.22$ (verwaarloosbaar en laag).

De IQ correlaties en met RT, RTSD en AB en CPS met snelheid (*time on task*) variëren van verwaarloosbaar, via bescheiden tot gemiddeld. Bij uitzondering wordt een hoge samenhang vastgesteld. Wijlen Arthur Jensen (2011) kritiseerde dit type studies omdat bewegingstijd (MT) niet los van de reactietijd (RT) gemeten werd. Hij voerde *decision time* in: een RT meting zonder MT. Deze is ook gecorreleerd met IQ maar de correlaties verschillen niet substantieel. Hij was zo vriendelijk om tegen betaling een methode aan te bieden om RT 'puur' te meten. De methode is te verkrijgen bij het *Jensen Institute of Mental Chronometry*. Het tijdschrift *Intelligence* bood hem de gelegenheid reclame te maken voor zijn instituut in een *peer-reviewed* artikel.

(8) Nog een operationalisatie van een ECT-basisproces is een poging de snelheid van zenuwgeleiding *direct* te meten. De Nederlandse oogarts Donders (19^{de} eeuw) deed dat als een van de eersten. Vernon en Mori (1992) probeerden de snelheid te meten door elektrodes op de mediane zenuw van de arm te plaatsen. Ze rapporteren correlaties van rond de $r = -.40$ met IQ scores. Het ophalen van overgeleerde informatie wordt als een elementair cognitief proces beschouwd. Een voorbeeld daarvan is de studie van *lexical access*. Verbaal IQ werd gecorreleerd met de snelheid van het vergelijken van letters (a-A tegenover A'', a-A''). Dit werd snelheid van toegang tot deze overgeleerde kennis genoemd. Deze snelheid (RT) correleerde $-.30$ met verbaal IQ (Hunt et al., 1975). Jensen (1998) heeft verschillende ECTs gecombineerd tot één getal en rapporteert multiple correlaties tussen de $-.50$ en $-.70$ met de G-factor.

De correlaties van ECTs met IQ (meestal de G-factor) variëren van laag, via gemiddeld tot hoog, kortom, het hele scala. Sommige zijn even hoog als de relaties tussen IQ en schoolprestaties. Een meta-studie kan tot een verantwoorde schatting van de relatie leiden en het lijkt erop dat het betrouwbaarheidsinterval groot zal zijn. Deze stand van zaken heeft er *niet* toe geleid dat elementaire taken en snelheidsmetingen de klassieke en tijdrovende IQ test vervangen. Een pragmatische reden is dat dure apparatuur en complexe computerprogramma's nodig zijn die de EEG, EVP en DNA metingen prepareren en die bewerkte metingen moeten geïnterpreteerd worden. Lezak et al. (2004) hebben een dik boek geredigeerd met tests die kwaliteit van hersenfuncties afleiden uit geheugen, waarneming, verbale capaciteiten en conceptvorming. We kunnen concluderen dat snelheid van waarnemen en discrimineren een rol speelt bij het tot stand komen

van individuele verschillen in intelligentie. Een praktische vraag is of ECTs, enzovoort criteriumprestaties beter voorspellen. Dat is niet uitgezocht. Mijn vermoeden is dat het precies zo gaat als met de informatieverwerkingsaanpak van Sternberg en IRT-tests: geen noemenswaardige verbetering ten opzichte van de klassieke IQ tests.

De oorzaak of basis van intelligentieverschillen is divers gelezen en uiteenlopend onderzocht bij normale en gewonde (oorlogen) personen. Ze variëren van metabolische kwaliteit, mentale snelheid, simultane en sequentiële verwerking van stimuli en snelheid van zenuwgeleiding. De correlaties tussen IQ en allerlei metingen van reactietijd variëren eveneens: van verwaarloosbaar tot hoog, waarbij Jensen de hoogste waarden vermeldt (.70), maar er zijn ook β s van .06 gevonden. Zowel de basis of oorzaak als de correlaties tussen biologische en neurologische parameters en IQ variëren ruim. Het biedt geen eenduidig beeld en de gemiddelde waarde zegt niet zoveel bij een dergelijke variantie.

Ter afsluiting van onderzoek naar de oorzaak/ basis van intelligentieverschillen en de naar de associatie van IQ met biologische parameters worden overzichtsstudies vermeld van Nisbett et al. (*American Psychologist*, Januari 2, 2012). In deze studie ook de behandeling van bekende thema's over IQ verschillen, bijvoorbeeld blank en zwart en tussen landen, van Deary (*Annual Review of Psychology*, 2012) en van Deary et al. (2010) over *neuroscience* en de G-factor. Ten slotte wordt het verbinden van IQ met de hersenen gerelativeerd in een essay van Dar-Nimrod & Heine (2011):

(1) Nisbett et al. (2012) stellen dat IQ en Gf en Gc het best school- en werkprestaties voorspellen. Hun artikel sluit aan bij twee klassiekers: Herrnstein en Murray's controversiële boek (1994) met de boodschap dat IQ vooral erfelijk bepaald is en van interventies kan niet verwacht kan worden dat IQs van verschillende groepen gelijk worden en Neisser et al.'s (1996) artikel met de boodschap dat de omgeving een bijdrage heeft. Ik herhaal: de *nature-nurture* variantieverdeling (h^2) varieert tussen de 25% en 75%. De auteurs wijzen erop dat de waarden samenhangen met kenmerken van de steekproeven. De relatie tussen SES en IQ is in de VS bijvoorbeeld hoger dan in Europa. In een studie van Turkheimer et al. (2003) is de waarde $r = .46$. In een studie van Asbury et al. (2005) is de correlatie .20. Dit zijn de *correcte* waarden. In het originele Nisbett et al. artikel worden *andere* waarden gerapporteerd (zie *American Psychologist*, February-March, 2012, p. 129).

Aantal en soort genen verantwoordelijk voor IQ verschillen zijn groot. Zo dragen 262 genen bij aan mentale retardatie en bij lengte - een goed te meten fysieke variabele - zijn al 700 genetische kenmerken betrokken. Ze concluderen (p. 6): *It may simple be that the number of genes involved in an outcome as complex as intelligence is just as small as the number of genes is large, and thus very difficult to detect without huge samples.* Je kunt opmerken dat de auteurs blijven vertrouwen op multivariate technieken om de bijdrage van alle genen, de omgeving en hun interacties te ontrafelen en kwantificeren. Een compleet model zal een grote hoeveelheid kleine interacterende effecten bevatten. De eruit volgende complexiteit, dat wil zeggen veel lage waarden op alle variabelen zal het verklaren of begrijpen van intelligentie niet dichterbij brengen. Ervaring met complexe pad analyses en structurele vergelijkingsmodellen leert dat de interpretaties lastig zijn en niet nuttig voor practici. Omgevingsfactoren moeten in rekening gebracht worden: adoptiekinderen winnen bijvoorbeeld ongeveer één standaarddeviatie IQ. Er is een gemiddelde correlatie (Cohen) tussen SES en IQ.

Nisbett et al. schatten de effecten van scholing hoog als je ze vergelijkt met *ontbreken* van enige scholing: ongeveer 2.0 SD. De verschillen tussen experimentele- en controlegroepen bij interventies zijn meestal niet groter dan 0,50 SD. Deze waarde doet zich voor onder gunstige omstandigheden, bijvoorbeeld bij een goed uitgevoerde en gemonitorde training. De resultaten zijn heterogeen. De auteurs wijzen verder op kleine verschillen tussen mannen en vrouwen.

Rushton en Jensen (2010) melden dat de verschillen tussen blank en zwart in de VS gelijk zijn gebleven tussen 1972 en 2002. Op jeugdige leeftijd zijn ze beperkt: gemiddeld 5 IQ punten op 7-jarige leeftijd maar groot op 24-jarige leeftijd: gemiddeld 17 IQ punten. De positie van blank ten opzichte van zwart is door de jaren nauwelijks gewijzigd, ondanks of misschien zelfs dankzij stimuleringsprogramma's.

Gegevens laten zien dat Joodse, Chinese en Japanse proefpersonen ongeveer 0,35 SD hoger scoren dan blanken. Dit blijven steeds ruwe vergelijkingen omdat de variantie binnen die grote groepen even groot is dan tussen de groepen.

Lynn (2011) heeft boek over dit onderwerp geschreven: *The chosen people: a study of Jewish intelligence and achievement*. Hij schat het gemiddeld IQ van de Aschkenazi Joden in de VS, Canada, het VK, Polen en Israël op 110 (twee-derde SD hoger dan het populatiegemiddelde). Voor de schatting heeft hij bronnen geraadpleegd. Het is geen gok of schatting. Er worden ondersteunende gegevens gerapporteerd: voor 1930 was ongeveer 0.8% van de Duitse bevolking Joods, maar ze kregen 24% van de Nobelprijzen voor wetenschap en literatuur. In de vroegere USSR en Rusland waren 70% van de Nobelprijswinnaars Joods. In de VS waren 60 van de 200 laureaten Joods. Lynn stelt drie interpretaties voor: Eugenetica, Vervolging (antisemitisme) en Discriminatie (een andere uiting van antisemitisme). Ook al is het niet politiek correct je kunt/mag ook denken aan het netwerk van deze groep en politieke motieven voor toekenning van de prijs. Om het idee van een superieur ras te voorkomen kan erop gewezen worden dat - op dezelfde wijze als het lage IQ van ontwikkelingslanden in Afrika niet toegeschreven kan worden aan ras (blacks) - het hoge IQ ook niet toegeschreven kan worden aan ras, bijvoorbeeld aan Ashkenazi Joden. Mensen vormen één ras. *The chosen people* verwijst niet naar een superieur ras al zijn ze bijbels uitverkoren. Toen Robinson wereldkampioen boksen werd vroeg men hem of dit een overwinning was voor zijn - het zwarte- ras, zei hij: 'Ja: voor het *mensenras*'.

Nisbett et al. kunnen niet om de bekende vragen heen: Wat zijn de centrale processen? Ze vermelden werkgeheugen en uitvoerende functies in relatie tot IQ. Wat is de G? Wat zijn de ingrediënten voor een succesvolle (leer) carrière? Ze noemen zelfdiscipline, uitstel van beloning, omgaan met stress. De auteurs wijzen er op dat deze kenmerken op individuele verschillen én op contextkenmerken berusten.

(2) Deary (2012) vat de kennis over intelligentie bondig samen en stelt vast dat intelligentieverschillen een focus van onderzoek blijven. Het aantal factoren is nog steeds onderwerp van debat. Hij bevestigt de relatie tussen IQ en sensorische discriminatie, inspectie en reactietijd: de rs liggen volgens hem tussen de -.31 en -.36. Het genetisch onderzoek is verfijnder geworden door het beschikbaar zijn van nieuwe technieken en kennis over afzonderlijke en over groepen genen. De rol van onderwijs wordt steeds weer bevestigd en ook de relatie tussen IQ en ouder worden en gezondheid. Het valt hem op dat sekseverschillen de aandacht blijven trekken, ook al zijn de verschillen klein.

(3) Deary et al. (2010) sommen kenmerken van onze hersenen op die zouden samenhangen met de G-factor of de eerste grote factor bij multiple intelligentietests. Er zijn bescheiden samenhangen gevonden met grijze massa dichtheid, witte massa integriteit, schedelomvang, corticale dikte, urine

zuurniveau, lengte, amplitude van het gemiddeld EVP die van het EEG wordt afgeleid, verhoogde neurale activiteit, afnemende neurale activiteit en zenuwgeleidingssnelheid. De meeste zijn hierboven vermeld.

(4) Dar-Nimrod en Heine (2011) waarschuwen tegen het gemakzuchtig leggen van een connectie tussen genen en concreet gedrag. Los van het feit dat de feitelijk correlaties tussen genen en intelligent gedrag zwak zijn, denken ze na over de boodschap die dit type onderzoek aan het publiek geeft. Ze beweren dat het onderzoek een *genetic essentialist bias* kent en bij leken veroorzaakt. Dit zijn cognitieve *biases*, vertekeningen waardoor mensen denken dat genen intelligentieverschillen veroorzaken. Genetische attributies gaan gepaard met het geloof dat de verschillen onveranderlijk zijn, want 'het zit immers in de genen' en het is biologisch en dus natuurlijk. De genetisch essentialistische lens is ook, terecht of niet, werkzaam bij het begrijpen van verschillen tussen rassen, sekse, seksuele voorkeur, criminaliteit, persoonlijkheidsstoornissen en obesitas. De opvattingen spelen een rol in politiek, wetgeving en in ideologische stromingen. Gergen (2010) is een cultureel georiënteerde psycholoog/socioloog en hij verdedigt dat onze hersenen vooral een orgaan zijn om cultureel bepaalde middelen en doelen te construeren. Genen bepalen onze concrete prestaties en activiteiten niet.

Het feit dat er mechanismen uit allerlei wetenschappen bijgehaald worden om intelligentieverschillen te verklaren laat zien dat psychologen slechts beperkt gericht zijn op specifieke psychologische mechanismen. Ze kijken een beetje rond wat een passende trend is om 'iets' te verbinden met het IQ. Om een dergelijke trend tegen te gaan helpt Darwins houding: hij bleef *observeren* om de soorten te begrijpen en werd geen farmacoloog of theoloog. Freud begon als neuroloog maar zijn focus bleef de interactie met de patiënt om zijn denken over de drie lagen van de persoon en psychologische mechanismen aan te scherpen.

De overzichtsstudies vullen de hiervoor vermelde conceptuele en empirische analyses over IQ en hersenen en IQ en reactietijden aan. Er is grote variatie in lezingen over oorzaken en bases de correlaties tussen IQ en biologische parameters.

Geschiktheden Intelligentie en geschiktheid noemde ik 'twee-eiige tweelingen' die zich willen onderscheiden al komen ze uit dezelfde familie. Bij geschiktheid denken we aan het voorspellen van rendement van training en opleiding. Tests voor geschiktheid bevatten niet één maar een aantal factoren want geschiktheden zijn gedifferentieerd naar verschillende taken, scholen, opleidingen, functies en beroepen. Voor een HBO opleiding tot toegepast psycholoog zijn andere geschiktheden nodig dan voor farmacoloog, jurist, bankdirecteur, vuilnisman of effecten- of huizenmakelaar.

Geschiktheden hebben een theoretische en praktische oorsprong. Ze zijn ontwikkeld om de veelzijdigheid van intelligent handelen te beschrijven en bedoeld voor selectie en opleiding voor complexe functies. De opleiding voor zulke functies is duur en de uitvoering van de taken is verantwoordelijk werk; denk aan piloten of operators in een kernreactorcentrum. Er zijn met andere woorden argumenten om kandidaten te selecteren die renderen. Personen verschillen in geschiktheid of hebben verschillende hoeveelheden training nodig om een vereist niveau van geschiktheid te behalen. Waarschijnlijk komen niet alle

kandidaten door training op dezelfde niveaus. Grote verschillen in het begin zullen verminderen maar er blijft daarna variatie tussen personen over. Twee voorbeelden: Fleishman heeft onder meer geschiktheden van piloten in kaart gebracht en Horn en Cattell hebben een geschiktheidsmodel ontworpen dat voor een deel een reactie is op Spearman's twee-factortheorie. Eén van de intelligentiefactoren: G-crystallized (G-cultural) zou gevoelig zijn voor training, onderwijs en culturele invloeden:

(1) Fleishman (1975) heeft een taxonomie van vaardigheden beschreven die verwant is aan geschiktheden. Hij omschrijft een *ability* of geschiktheid als ‘... a more general capacity of the individual, related to performance in a variety of human tasks’ (p. 1131). Zo blijkt dat ruimtelijke visualisering samenhangt met prestaties bij navigeren, lezen van stukken en de kwaliteit van het werk van tandartsen. De auteur verdeelt geschiktheden in globale, bijvoorbeeld perceptueel motorische vaardigheden, lichamelijke conditie en psychologische capaciteiten. De taxonomie ziet er als volgt uit: aan de *top* staan de algemene categorieën, bijvoorbeeld waarnemings- en motorische vaardigheden zoals snelheid van armbeweging, coördinatie van ledematen, mikken op een doel. *Daaronder* een goede lichamelijke conditie, bijvoorbeeld het evenwicht kunnen bewaren bij verstoring, grove lichaamscoördinatie en dynamische en explosieve kracht. *Verder* psychologische capaciteiten zoals ze in de multipiele intelligentietheorieën beschreven zijn, bijvoorbeeld ruimtelijke en verbale vermogens. Fleishman bedacht zelf taken en nam ook een deel van Galton's taken over. Ze gaan van het evenwicht bewaren op een balk en grove lichaamsmotoriek tot het onderscheiden van weinig in gewicht verschillende objecten en visuele waakzaamheid voor een onregelmatig opduikend signaal dat hard/zacht, sterk/zwak kan zijn. Bij psychologische tests is er voorkeur voor numerieke, ruimtelijke, logische en verbale vaardigheden. Deze taken werden gefactoranalyseerd en geïnterpreteerd als brede geschiktheidsfactoren. Ze passen bij functies in de luchtvaart want piloten en verkeersleiders moeten beschikken over accurate waarneming van verschillen, waakzaamheid, motorische coördinatie, alertheid, ruimtelijk inzicht en verbale vaardigheden om een stand van zaken precies te kunnen beschrijven.

(2) Horn en Cattell (1966) hebben een hiërarchisch geschikthedenmodel ontworpen. Het bevat specifieke, primaire factoren zoals verbaal en numeriek vermogen, ruimtelijk inzicht, snelheid van waarneming, geheugen. Deze lijken op de PMA van Thurstone. In twee bekende tests voor geschiktheden de Differentiële Aanleg Test (DAT) en de *General Abilities Battery* (GATB) zien we deze factoren terug. Een tweede-orde factoranalyse op de scores op de primaire factoren levert twee belangrijke en drie minder belangrijke tweede-orde factoren op. De twee belangrijke zijn *Gf: Fluid intelligence* en *Gc: Crystallized intelligence*. *Gf* verwijst naar flexibiliteit in denken, abstract redeneren en oplossen van nieuwe problemen. Het lijkt op de *Eduction of relations and correlates* van Spearman en is een uitdrukking voor de complexiteit van relaties die een persoon aankan, dat wil zeggen kan waarnemen en bewerken wanneer hij geen beroep kan doen op zijn geheugen. Dit vermogen vertoont naar verhouding lage ladingen op culturele inhouden en geleerde taken. De *Gc* is verworven, geconsolideerde kennis en hangt samen met het culturele aanbod, onderwijskansen en persoonskenmerken zoals vasthoudendheid en gemotiveerdheid. Als redenering uit het ongerijmde, kun je beweren dat de *Gf*-verschillen overblijven als personen dezelfde onderwijskansen, een gelijk cultuuraanbod gekregen hebben en even gemotiveerd zijn. De *Gf* en *Gc* zijn niet onafhankelijk. Een voorbeeld van een *Gf* (1) taak en een *Gc* (2) is:

- 1 Temperatuur verhoudt zich tot kou als hoogte tot...
 - a. hitte b. centimeters c. grootte c. lengte
- 2 Bizet verhoudt zich tot Carmen, als Verdi tot...
 - a. Aïda b. Elektra c. Lakmé d. Beethoven e. Tosca

Het eerste item gaat over het vinden en toepassen van een relatie. Het tweede gaat over cultureel verworven kennis want hoe kun je anders weten dat Bizet en Verdi geen profvoetballers of maffiosi zijn maar operacomponisten.

Je bent intuïtief geneigd te denken wat Horn en Cattell veronderstellen dat de Gf een hogere overerfbaarheidscoëfficiënt (h^2) heeft dan de Gc. Jensen (1998) beweert echter dat beide een ongeveer even hoge h^2 -waarden opleveren. Hij is een aanhanger en verdediger van de G-factor. De h^2 is de verhouding tussen de genetische variantie in een kenmerk van een populatie (oogkleur, lengte, IQ) en de totale variantie van dat kenmerk in de populatie. Het onderscheid tussen Gf en Gc wordt regelmatig gebruikt. Cattell heeft er nog drie aan toegevoegd: Gv: visualisatie, het vermogen om informatie in beelden om te zetten en vast te houden, Gr: het vermogen en gemak om informatie uit het geheugen op te halen en Gs: de snelheid waarmee problemen opgelost worden. Horn (1988) heeft de taxonomieën samengevat en geordend met behulp van zes factoren:

I *Short term Apprehension Retention (SAR)*

II *Fluency of retrieval from long-term storage (TSR)*

III *Visual Processing (Gv)*, bijvoorbeeld visueel conceptualiseren, analytische perceptie en *Gestalt closure*

IV *Auditory Processing (Ga) of broad auditory intelligence*, bijvoorbeeld discriminatie van geluiden, rijm kunnen herkennen en onmiddellijk auditief geheugen

V *Processing Speed (Gs)*

VI *Quantitative knowledge (Gq)*

De persoonskenmerken introversie, intellectuele nieuwsgierigheid en Gf en Gc zijn recent door Schmidt (2014) benut om de samenhang van deze variabelen met *academic* functioneren en beroepsinteresse en -prestaties bij volwassenen te beschrijven. Het geeft tevens een beeld van de relatie tussen de Gf en Gc. Hij gebruikte daarvoor vier artikelen die zes variabelen bevatten. De structuur waarin Gc, specifieke beroepsinteresses en gewetensvolheid *academic* en beroepsprestaties 'veroorzaken' kreeg steun in eerder verricht empirisch onderzoek. Gc blijkt een centrale 'oorzaak' voor adequaat mentaal functioneren op oudere leeftijd. De structuur is dat introversie interesse in leren bevordert, de laatste staat onder de invloed van Gf die op zijn beurt Gc beïnvloedt en Gc is gerelateerd aan beroepsinteresses. Samen 'veroorzaken' ze de werkprestatie die op zijn beurt samenhangt met gewetensvolheid. Hij noemt het een *general integrative model for individual differences variables* (p. 214). De variabelen zijn introversie, interesse in leren, Gc, Gf, beroepsinteresse en gewetensvolheid. Het model verschilt niet van een hiërarchische

regressiemodel. Schmidt gaat mee met de tijdgeest door introversie te verbinden aan hormoonproductie.

Cattell zag in tegenstelling tot Jensen het nut van de G-factor van Spearman niet in: ... *there can be no such a thing as a categorical general factor...* p. 87). Categorisch betekent volgens hem dat het om een factor gaat die uniek bepaald is en invariant is over factoranalyses met verschillende groepen tests. Daar heeft hij geen empirische steun voor kunnen vinden. Horn (1988) toont zich wat inschikkelijk en maakt in zijn overzichtsartikel over *Human Abilities* een *top-down* schema dat van de algemene factor G van Spearman via de Gc en Gf overgaat in de zes brede factoren die hierboven vermeld zijn.

De analyses van Fleishman, Horn en Cattell laten dus zien hoe dicht intelligentie en geschiktheid bij elkaar liggen. Horn is evenals Fleishman betrokken bij het ontwerpen van instrumenten om te selecteren voor kostbare gespecialiseerde opleidingen en hoge functies in het bedrijfsleven. Dat verklaart de aandacht voor complexe visueel-motorische taken en hun doel om met behulp van tests te voorspellen of de intensieve en dure trainingen renderen. Dit verklaart ook het gebruik van specifieke tests die laden op één bekende factor, bijvoorbeeld *Processing Speed*. De brede factoren verwijzen ook naar individuele verschillen in snelheid van waarnemen en informatieverwerken. Geschiktheidstests zijn bruikbaar om studenten voor dure opleidingen te selecteren en succes te voorspellen. Een student die medicijnen, farmacologie, toegepaste psychologie, informatica of rechten wil gaan studeren heeft nog geen specifieke vakkennis. Met behulp van de geschiktheidstest willen we met een zekere waarschijnlijkheid kunnen voorspellen of hij de vakken aankan die in die studies onderwezen worden.

In Duitsland is een intelligentiemodel ontwikkeld dat een verbinding vormt tussen Spearman, Guilford, Horn en Cattell en de opvatting over intelligentie als informatieverwerking. Het is het *Berlin Intelligenz Structur* (BIS) met Gf en Gc. Het BIS heeft een facetstructuur (Guttman, 1971). Er is een inhoudelijk facet voor verbale, numerieke en figuur-analytische vaardigheden. Dat is onderscheiden van een tweede, operationeel facet voor snelheid van informatie verwerken, geheugen, creativiteit en hoeveelheid informatie die een systeem aankan. De facetten resulteren in twaalf zogenoemde *structuples*. Dat zijn drie inhouden x vier operaties. Het model is bij 9520 proefpersonen getoetst. Er werd verwacht dat de Gf samenhang met geheugen en snelheid van informatie verwerken en Gc met *fluency* en kennis en wat minder met de hoeveelheid informatie die een systeem kan verwerken. De Gf en Gc kwamen tevoorschijn uit de hiërarchische factoranalyse en de overige veronderstellingen werden gesteund, behalve die van een correlatie tussen Gc en *fluency*. Intelligentie en geschiktheden zijn weliswaar een tweeling maar er is ook verschil: ze kunnen hun identiteit behouden.

Voorgaande onderscheidingen gaan voorbij aan de generalist die een diagnosticus vaak en noodgedwongen is. Als hij echter voor specifieke groepen werkt (selectie van piloten, operators, enzovoort) dan zijn de onderscheidingen zinvol. Dit type tests maakt dan deel uit van het protocol. Als de diagnosticus een algemene indruk van het niveau van een cliënt

nodig heeft zal hij een G-verzadigde test kiezen en een totaal IQ berekenen. Als hij moet differentiëren, bijvoorbeeld naar verschillende opleidingen voor een cliënt, liggen tests voor geschiktheden voor de hand.

Prestatie verwijst naar wat een persoon geleerd heeft van een afgebakende hoeveelheid stof. Het is ontwikkelde capaciteit of geschiktheid. Leerkrachten zijn geïnteresseerd in huidige prestaties van leerlingen en in vorderingen die ze door de tijd heen zullen maken. We denken bij prestaties aan basisvaardigheden zoals lezen, rekenen, taal en werk- en studievoordigheid. Daarnaast gaat het over prestaties in opleidingen, van basisonderwijs, vmbo, assistent in opleiding (AIO) tot PhD cursussen en over opleidingstrajecten bij instellingen en bedrijven. Er is onderwijspsychologische theorievorming over hoe de prestaties bij vaardigheden als lezen, rekenen en taal tot stand komen. De theorievorming over technisch lezen kan aansluiten bij onderzoek naar perceptuele discriminatie, analyse en synthese van symbolen en klanken. Bij begrijpend lezen past theorievorming over het ontwikkelen van inzicht in concepten, ontwikkeling van de semantiek en de grammatica. Bij rekenen is er theorievorming over het getalbegrip en rekenkundige bewerkingen. Deze spelen bij het maken van toetsen een beperkte rol want ze zijn al 'opgenomen' in de ordening en inhoud van de doelstellingen van het domein, zoals lezen, rekenen, biologie, natuurkunde, Engels, enzovoort. Er moet een niveau gehaald worden voor lezen, rekenen en schrijven. Deze doelen leiden tot curricula met passende toetsen. Ze stellen ons in staat te bepalen of leerlingen geleerd hebben wat in de doelstellingen staat.

Voor nieuwe vakken in het voortgezet en hoger onderwijs geldt hetzelfde. Er bestaat theorie over het verwerven van inzicht in biologische, natuurkundige, scheikundige en psychologische verschijnselen. Ze speelt een beperkte rol bij de formulering van onderwijsdoelstellingen en het construeren van de studietoetsen. Het meten van onderwijsprestaties staat onder het regime van Cito Rasch modellen. Items 'mogen' opgenomen worden als ze passen in een Rasch model. Het lijkt op wat Henry Ford zei: Alle kleuren voor de T-Ford zijn geoorloofd als die maar zwart is.

Box Het denken over leren en prestatie verandert voortdurend. Het is niet bij te houden!

In het begin van de 19^{de} eeuw was de Britse scholing beperkt tot inhoud en was uniform in het Britse Imperium. Scholing moest immers voor ambtenaren zorgen die in het hele rijk ingezet konden worden. De curricula waren gelijk: lezen, schrijven en rekenen. De school moest identieke klerken, bedienden afleveren die zowel in Londen als Afrika en India konden werken. Prestatietoetsen gingen daarom over leesbaar en foutloos schrijven, foutloos rekenen en gearticuleerd lezen. Sedert 200 jaar heeft de machine veel zwaar werk overgenomen en massaproductie mogelijk gemaakt. Werknemers werden ook opgeleid om identiek te zijn. Ze konden aan iedere machine verder werken in de tweede, derde shift. Ze deden in de fabrieken hetzelfde werk in dezelfde kleding: bankbedienden en ambtenaren droegen vroeger een grijze stofjas. Kinderen droegen op school een uniform en dat was niet alleen om de verschillen tussen kinderen niet zichtbaar te laten zijn. Het ging er ook om gelijke kinderen te creëren voor de taken in de samenleving. Rijken van het *Britannia rules the waves* en

het strikt communistische type bestaan bijna niet meer. Het is mogelijk dat grote bedrijven nieuwe *Empires* zijn, maar er zijn verschillen. Soms mogen werknemers zelfs innovatief zijn, zelfs op de universiteiten. De nadruk ligt niet meer alleen op massaproductie, bijvoorbeeld korte artikelen in Engelstalige tijdschriften. Het neemt niet weg dat veel bedrijven in China en Azië die uniformiteit tonen. De lage lonen landen hebben nog wel identieke, dat wil zeggen inwisselbare werknemers: voor jou 1000 anderen.

Wie bepaalt *nu* wat we moeten leren? Wie bepaalt of we aan de eisen voldoen? De staat en voor een deel ook de ouders die vol verwachting zijn over hun kinderen en hopen dat ze slim, goed opgeleid en gelukkig zijn. De staat moet ouders daarin tegemoetkomen want zij zijn het eerste station om bruikbare burgers af te leveren. Hetzelfde geldt voor werkgevers. Zij zorgen voor banen maar werknemers moeten hen niet al te zeer tegen de haren instrijken. De SATs en Cito-toetsen meten prestaties. Daarnaast zijn oordelen van chefs aanvaarde selectie- en beloningsmiddelen.

Recent lijkt het dat ouders minder greep krijgen op wat kinderen leren omdat de IT revolutie iedereen en vooral de jongsten toegang geeft tot informatie waar ouders niets van weten of zelfs niets van willen weten. Kinderen zijn handiger met nieuwe apparatuur dan ouders. Leraren moeten ook accepteren dat zij niet meer zo nodig zijn. Leerlingen weten zelf de weg vaak al om aan de nodige informatie te komen. Ook de staat kan niet meer controleren wat geleerd wordt ook al probeert die daar controle op te houden via SATs en Cito-toetsen. Leerlingen vinden het vaak saai wat ze op school leren. Ze zoeken het zelf wel uit. Er ontstaat een andere houding ten opzichte van kennisverwerving en het daarop getoetst worden. Een voorbeeld: Enkele Harvard studenten schreven hun essays door van alles van het internet te plukken en leverden dat in bij hun docenten. Die vonden dat bedrog maar de studenten konden dat *niet* begrijpen. Ze voelden *niet* dat ze iets illegaals of onoorbaars hadden gedaan. Waarom dat allemaal moeten kunnen oplepelen als het op het internet staat? Nog een voorbeeld: een student die gevraagd werd of zij een Spaanse brief kon lezen en vertalen, zei daar op 'ja' terwijl ze nooit Spaans geleerd had. Ze had echter een vriend die tweetalig was: Spaans-Nederlands. Ze vatte dat op als een voor haar beschikbare vaardigheid. En wat te denken van een solliciterende universitaire docent die op de vraag of ze statistiek en methoden voldoende beheerste, 'ja' antwoordde want ze had een partner die M & T professor was? Is het nog nodig dat docenten programma's inzetten om na te gaan of stukken van het internet geplukt zijn? Mogen ze zulke werkstukken weigeren? Als ik een partner heb met specifieke kennis kan ik die dan als fort opvoeren als ik beloof de partner 'aan te houden', zolang ik bij deze baas werk?

Tijden veranderen: De eerste apparaten waren zwaar en niet te dragen: denk aan de eerste radio's en TVs. Rond 1960 was er de Walkman: je kon met je radio rondlopen. Daarna kwam de MP3 speler, een nog beter apparaat maar het *verdween*. Dat gaat steeds verder: denk aan voorraden van boeken voor onderwijs, banken (we lopen nooit meer een bank in om geld te halen) en het wordt ontraden om met geld op zak te lopen. De enige tegenstanders van deze verandering zijn de bedelaars. Hoe gaat dit doorwerken op de scholen? Kunnen we niet net zo goed achter een zoekmachine gaan zitten? Houdt de traditionele school op te bestaan? Mogen kinderen leren wat ze willen? Er is geen afstand meer. Met Skype kun je over de hele wereld communiceren. Waarom moet Robbert Dijkgraaf nog naar de studio of de collegezaal komen? We kunnen zijn lunch-discussies met alle collega's op Princeton volgen via het internet. Leraren zijn niet meer nodig en boeken worden vervangen door e-boeken. Het enige wat ze nog zouden kunnen doen is leerlingen helpen zichzelf te organiseren en de leeromgeving met goede apparaten en software vorm te geven maar het is de vraag of zij daarvoor de aangewezen personen zijn.

Hoe moeten we nu die 'in het wild' verworven kennis en vaardigheden meten? En, heeft dat nog zin, als je om de vijf jaar weer iets anders moet weten en kunnen. Overleven in de toekomst alleen nog flexibele studenten en werknemers? Verdwijnen de Cito-toetsen omdat ze altijd achter de feiten aanlopen? Gaat de IT revolutie zo hard dat het meten van prestaties achterblijft? Of is dit een herhaling van het vooruitgangsgeloof van de 19^{de} eeuw.

En is informatie hetzelfde als kennis en inzicht? Nee, zeggen sommigen: kennis en inzicht zijn het resultaat van kiezen, selecteren, verwerken en scheppen van informatie over onderwerpen. En is de IT revolutie niet slechts één van de vele revoluties, religies, ideologieën? Van sommige kennen we de afloop. Neem bijvoorbeeld de 17^{de} eeuwse Franse dokter De la Mettrie die de mens als machine opvatte dat wil zeggen met een *zero-level-subjectivity*. Dat was een provocatie maar het hielp de wetenschap wel vooruit en was tegelijkertijd een 'halve waarheid'. Of neem Hegel, de 19^{de} eeuwse filosoof die de hele geschiedenis opvatte als het verschijnen van de Geest. Zeker, de wereld van gebeurtenissen, objecten, verschijnselen, zoals we die kennen bestaat niet zonder een organiserende geest, verstand, maar die wereld is er niet een van een Hegeliaanse complete subjectiviteit. Deze provocatie hielp om te laten zien dat de organiserende geest iets oplegt aan de werkelijkheid maar dat is steeds opnieuw weer een 'halve waarheid'.

Prestatietoetsen zijn aan vak inhouden gebonden en worden niet primair vormgegeven vanuit een theorie die verklaart hoe en met welke processen en mechanismen een bepaald niveau van vakkennis verworven wordt. De diagnosticus heeft deze theorievorming soms nodig, bijvoorbeeld als een cliënt onderpresteert met rekenen, lezen, wiskunde, taal, enzovoort. In dat geval krijgt hij de vraag voorgelegd wat de oorzaak is. De vorm en inhoud van prestatietoetsen zijn gevoelig voor de tijdgeest en staan onder invloed van de technologische vooruitgang. Misschien brengt de IT revolutie vernieuwingen over hoe en wat geleerd wordt en bijgevolg voor het toetsen van school- en werkprestaties.

Samenvatting en conclusie

Spearman volgde Galton op en ontwikkelde zijn twee-factortheorie van individuele verschillen in intelligentie. Hij vond empirische steun in de *positive manifold*: de matrix met positieve correlaties tussen tests of items (Carroll, 1993). De G verwees naar *activities of the mind of brain power* en bestond uit *Eduction, Deduction of relations* en *Correlations*. Thurstone nam verschillende vaardigheden waar en stelde zeven *Primary Mental Abilities* voor. Met zijn *simple structure* factoranalyse kreeg hij steun voor het onderscheid in de zeven: elke afzonderlijke test laadde hoog op één factor en laag op de zes andere. Guilford bracht zijn *Structure of Intellect* (SOI) onder in een doos met drie ribben met 4 (later 5) × 5 × 6 cellen. Dit leidde tot de constructie van tests met als doel alle cellen van de doos zuiver te meten. Voor de oorspronkelijk 120 cellen zijn ongeveer 70 à 80 tests geconstrueerd. Vernon bracht ze in een hiërarchisch model onder met de G-factor aan de top, twee brede groepsfactoren daaronder en daaronder smalle groepsfactoren. Carroll vatte in zijn factoranalyse het onderzoek met intelligentietests tot 1992 samen in een drie-lagen model: laag III: algemene intelligentie, II: brede factoren (*G fluid, G crystallized, Memory, Visual Auditory, Retrieval, Cognitive Speed, Decision and Processing Speed*) en I vele specifieke

factoren. Tests voor al deze factoren worden zelden gebruikt. De diagnosticus benut meestal G-verzadigde tests en tests voor geschiktheden die aansluiten bij de PMA.

De interpretatie van de psychometrische intelligentie loopt langs dezelfde lijnen als die van de *Big Five* en de trektheorieën. Is het een evolutionaire aanpassing mede beïnvloed door het klimaat (Kanazawa), is het een latente trek, zijn het informatieverwerkende processen, zijn het biochemische processen, gelokaliseerd in bepaalde delen van het CZS? Zijn er kleine eenheden (ECTs) aan te wijzen die verantwoordelijk zijn voor individuele verschillen in IQ scores? Er zijn correlaties gevonden met informatieverwerkende processen. Dat 'hersenen' iets met intelligentie te maken hebben is door de neurologen, fysiologen en psychologen als Hebb en Luria bestudeerd en bevestigd. Ze probeerden bepaalde functies aan verschillende plekken ('loci') in het brein te koppelen: bijvoorbeeld de hersenstam voor attentie en alertheid, de voorhersenen, frontaalkwab voor planning en organisatie en de overige met als functie de stimuli te coderen en te verwerken. Dit idee is ook aanwezig in de verschillende functies van de twee hersenhelften. Snelheid van zenuwgeleiding werd al sinds de 19^{de} eeuw in verband gebracht met cognitief functioneren. De variatie is groot. Dat geldt voor conceptuele interpretaties van de oorzaak, de basis van IQ-verschillen en voor correlaties tussen IQ scores en biologische kenmerken van de hersenen. Elementaire cognitieve taken (ECTs) zijn kleinere eenheden dan het item en de latente trek van de psychometrici. Indexen van consistentie en snelheid van het onderscheiden van stimuli werden gecorreleerd met IQ. De correlaties variëren van de $r = -.20$ tot $-.50$. Kenmerken van het EEG en EVPs zijn in verband gebracht met het IQ. De waarden zijn gemiddeld (Cohen). Het is niet waarschijnlijk dat deze metingen de klassieke IQ tests gaan vervangen. Recente ontwikkelingen in meetapparatuur voor hersenactiviteit van biologische, neurologische en misschien zelfs de DNA structuur zullen niet gemakkelijk door diagnostici gebruikt worden. Die berusten op een ander vak. Het is bovendien niet aangetoond dat ze efficiënter zijn en criteria beter voorspellen. De vooronderstelling bij dit type studies lijkt dat je denkt meer te begrijpen als je 'dieper gaat', dat wil zeggen zoeken naar biologische en neurologische parameters. Er is geen analyse over de samenhang, 'brugwetten', tussen deze wetenschappen.

Intelligentie en geschiktheid zijn beschreven als een twee-eiige tweeling. Geschiktheid verwijst naar het tot ontplooiing brengen van een aanwezige geschiktheid (*aptitude*) door scholing en training. Anders als intelligentie wordt geschiktheid meestal uiteengelegd in een aantal factoren. Fleishman onderscheidt brede geschiktheidsfactoren in onderzoek bij onder meer piloten. Horn en Cattell vinden in hun analyse van primaire factoren twee belangrijke en drie minder belangrijke factoren. Van de twee belangrijke, *G-fluid* (vloeiende intelligentie) en *G-crystallized* (uitgekristalliseerde intelligentie) verwijst de eerste naar flexibiliteit en abstractie. Deze zou afnemen met het ouder worden. De tweede hangt samen met cultuur en scholing en blijft stabiel of neemt soms nog (een klein beetje) toe gedurende het leven. Horn heeft later zes brede geschiktheidsfactoren onderscheiden die berusten op factoranalyses van uiteenlopende selectietests voor complexe functies en dure opleidingen: Opnemen en selecteren van informatie in het kortetermijngeheugen, Visueel

en Auditief correct verwerken van informatie, Snelheid van verwerken van informatie, Kwantitatieve kennis en Gemak waarmee je uit je langetermijngeheugen informatie ophaalt. Horn erkent, in tegenstelling tot Cattell, het bestaan van de G-factor. Het *Berlin Intelligenz System* (BIS) brengt intelligentie en geschiktheidsfactoren onder in één model .

Het IQ is met vele lichamelijke en persoonlijkheidskenmerken in verband gebracht. Ook de verschillen tussen sociale klassen en landen zijn veelvuldig vastgesteld en geïnterpreteerd. Westerlingen overschatten hun IQs naar verhouding het meest.

De vele onderscheidingen en variërende correlaties tussen IQ en hersenprocessen zijn voor de diagnosticus niet direct van belang. Hij kiest voor een G-test of IQ-totaalscore als hij een algemeen beeld van de intelligentie van een cliënt wil hebben. Moet hij differentiëren voor een school of beroep, dan ligt een geschiktheidstest voor de hand.

Theorie speelt een beperkte rol bij het beschrijven en verklaren van prestaties. De resultaten van onderwijs, trainingen en scholing worden gemeten met toetsen. De inhoud daarvan berust niet op theorie over de basisvakken op school (lezen, rekenen, schrijven) van vakken (wiskunde, biologie, taalkunde, aardrijkskunde, enzovoort). De doelstellingen en inhouden van curricula bepalen wat er geleerd moet worden en aan welke criteria de kennis moet voldoen. Als een diagnosticus te maken krijgt met cliënten bij wie het reguliere onderwijs niet tot het gewenste resultaat leidt dan heeft hij wel theorievorming nodig die helpt te verklaren waarom een vereiste prestatie niet geleverd wordt. Het wat van school- en werkprestaties loopt mee met ontwikkelingen, vooral van de IT technologie.

2. Ontwikkeling

De individuele verschillen oriëntatie van intelligentie, geschiktheid en prestatie wordt gedomineerd door psychometrische intelligentie. Ontwikkeling van intelligentie, geschiktheid en prestatie wordt voornamelijk opgevat als onderzoek naar stabiliteit en continuïteit. Naast individuele verschillen zijn de ontwikkelings- en sociale-contextoriëntatie onderscheiden als gezichtspunten waaronder elk gedrag bestudeerd kan worden. Ten eerste heeft de ontwikkelingsoriëntatie op cognitie, kennen een gezicht gekregen door het werk van Jean Piaget (1896-1980). Ten tweede is er geen expliciete theorie over de ontwikkeling van geschiktheden en prestaties, maar deze ontwikkelen zich evenzeer door de tijd heen. Het systeem van Piaget wordt geïntroduceerd, kritiek vermeld en er wordt gespeculeerd of de invloed van de Piagetiaanse benadering blijvend is.

Jean Piagets theorie Zijn oorspronkelijk doel was de studie van de ontwikkeling van het westers denken. Hij kon daartoe de geschiedenis van de westerse wijsbegeerte bestuderen. Als gevolg van zijn opvatting dat de geschiedenis van het denken zich herhaalt in de ontogenetische ontwikkeling van baby tot volwassene, bestudeerde hij kinderen tussen de 0 en 18 jaar. Dit Haeckeliaans herhaalprincipe werd en wordt niet algemeen aanvaard maar er is mogelijk een Lamarckeaanse kentering op komst. We kunnen ook geleerde kenmerken overerven. Bovendien is niet duidelijk hoe die herhaling opgevat moet worden. Is het letterlijk zo, of is het een metafoor? De gedachte is op zich niet vreemd of bij voorbaat

kansloos, bijvoorbeeld in de natuurkunde ziet men vergelijkbare processen op de kleinste schaal in de deeltjesfysica en op de grootste in de sterrenkunde.

In Piagets organismische theorie ligt de nadruk op ontwikkeling van binnenuit. De sociale context (scholing, training, cultuur) kan de ontwikkeling beperkt beïnvloeden. De stadiumgewijze ontwikkeling vindt plaats door adaptatie (accommodatie) - aanpassen door de cognitieve structuur te veranderen - of aanpassen door assimilatie - het nieuwe in een bekende structuur opnemen. Hij wees intelligentietests en factoranalyse af als middelen om cognitie te bestuderen. Hij bedacht allerlei taakjes en experimentjes om de stadia in de ontwikkeling te demonstreren. Hij verborg voorwerpen voor baby's en observeerde wat ze deden. Gingen ze zoeken of hield het object voor hen op te bestaan als het uit het oog verdween of gingen ze zoeken, en waar? De conservatieproeven gaan ook over een vorm van stabiliteit of soliditeit van objecten, bijvoorbeeld het feit dat hoeveelheid, aantal en volume gelijk blijven ondanks een andere verschijningsvorm: een pond veren is even zwaar als een pond lood.

Zijn theorie had een groot bereik. Zij beperkte zich niet tot het kennen van fysische wetten, ook de ontwikkeling van het concept rechtvaardigheid en het standpunt van een ander kunnen innemen werden begrepen vanuit de algemene stadiumtheorie. De kernvraag voor Piaget was: hoe wordt de mens van een afhankelijk biologisch wezen, uitgerust met een paar reflexen tot een volwassen formele denker die logica, wiskunde en experimenteel onderzoek begrijpt? Zijn vraagstelling ligt op een raakpunt van cognitieve psychologie en filosofie. Hij was er trots op dat er naar zijn werk werd verwezen in filosofische compendia. Hij werkte in Genève waar zijn studenten jaarlijks een nieuw onderwerp als taak kregen. Zij ondervroegen veel kinderen en legden hen allerlei proefjes en taken voor. De verslagen leidden bijna jaarlijks tot een nieuw boek van zijn hand. Regelmatig op maandagochtenden nodigde hij sprekers uit andere vakgebieden (wiskunde, logica, taalkunde) uit om hun licht te laten schijnen over zijn thema's. Hij begon met observeren van zijn eigen drie kinderen en zodra ze talig waren, ondervroeg hij ze naar aanleiding van zijn proefjes. Dit noemde hij het klinisch interview. Door zijn studenten zijn op dezelfde manier veel kinderen en adolescenten onderzocht. Hij onderscheidde de volgende vier stadia van cognitieve ontwikkeling:

In het sensomotorische stadium (0-2 jaar) coördineert het kind sensorische waarnemingen en motorische vaardigheden om kennis van de buitenwereld te verkrijgen. Dit stadium is verdeeld over zes substadia die empirisch een 'natuurlijke' ordening vertonen in normale en geretardeerde steekproeven (Rensen, 1994). In deze periode ontdekt het kind dat een object blijft bestaan ook al ziet het dat niet omdat het onder een doek verborgen is. Het kind ontdekt de eerste invarianties, de soliditeit, stevigheid van de wereld van objecten en mensen. Als het een object *niet* ziet vallen, als het losgelaten wordt, toont het kind verbazing.

Het preoperationele stadium (2-6 jaar) bevat het internaliseren, het verinnerlijken van handelingen, die het al kan uitvoeren. Een kind kan bijvoorbeeld één kenmerk van een probleem of object vasthouden. Als bij de taak van het conserveren een zelfde hoeveelheid water uit een breed kort glas in een smal lang glas wordt gegoten, houdt het één kenmerk vast, bijvoorbeeld de hoogte van

het water en zegt dat het hoogste glas het meeste water bevat. Dit geldt ook voor sociale waarneming: het kind ziet één perspectief, dat van zichzelf, niet dat van een ander.

Het concreet operationele stadium (7-12 jaar) bevat het in staat zijn om twee onafhankelijke kenmerken tegelijk te benutten en in relatie te brengen. Het kind kan bijvoorbeeld objecten ordenen volgens twee onafhankelijk dimensies, bijvoorbeeld de grote-rode schijven uit een verzameling schijven met kleuren en grootten. Nu worden de lengte en de hoogte van het glas gelijktijdig betrokken in het oordeel of er evenveel water in twee glazen zit. Dit is een ontdekking van de invariantie van kenmerken van objecten en mensen, ook al is de verschijningsvorm verschillend. Dit stadium staat nog niet het oplossen van abstract wiskundige problemen toe.

Het formeel operationele stadium (12-18-21 jaar) staat het oplossen van wiskundige en logische problemen toe. Natuurwetten kunnen ontdekt worden. Als een adolescent bijvoorbeeld de kleur rood moet maken uit vijf stoffen dan gaat hij systematisch de tien mogelijkheden af in plaats van zomaar wat uit te proberen. Hij 'weet' of handelt zodanig dat hij in tien stappen tot de oplossing kan komen: de combinaties van 2 uit 5: $n(n-1)/2$.

Kritiek Er is uiteraard kritiek op Piagets werk. Eerder zijn al verschillende interpretaties van het stadiumbegrip besproken. Wat is de grondslag: een formele structuur, een hermeneutisch verworven inzicht in onderliggende betekenis of een empirisch ondersteunde factor? Er is geen winnende interpretatie ook al beweert Brainerd (1978) dat hij gewonnen heeft en beweert dat stadia geen empirisch bestaan hebben. Zijn houding en empiristische interpretatie van stadia hebben het stadiumbegrip gemarginaliseerd. Naast kritiek op het bestaan van stadia zijn bezwaren tegen de leeftijden waarop gemaakt. Piaget (1972) erkende dat de formele operaties niet bij ieder op 12 tot 14 jaar leeftijd bereikt zijn. Nu legt men wel de grens rond het achttiende levensjaar. Dit wordt tegenwoordig gekoppeld aan het groeiende brein dat instabiliteit zou veroorzaken bij adolescenten. Het zou waar te nemen zijn aan het experimenteren in de zijn van risico's onderschatten en niet kunnen overzien van gevolgen van hun handelingen op lange termijn. Bovendien zijn de formele operaties niet zozeer een structuur los van inhoud, maar domeinspecifiek: in je beroep ben je formeel operator, maar daarbuiten niet. Ieder zijn vak.

Enkele onderzoekers wilden er niet aan dat de cognitieve ontwikkeling zo abrupt en zo vroeg eindigt. Zij stelden een *postformele* periode voor (Perry, 1970; Zajonc, 2006). Kramer (1983) verbond het idee van postformele operaties (*adult thinking* genoemd) met 'relativistisch-dialectisch denken'. Dit 'stadium' houdt in: (a) het besef van de relativistische, niet-absolute aard van kennis (b) het accepteren van contradicties en (c) integratie van een contradictie in eenheid van een hogere orde. Hoewel er een doorgaande ontwikkeling, dat wil zeggen verbreding, verdieping in cognitie is door de tijd heen, is dit (zou dit) geen nieuw stadium in de ogen van Piaget (zijn). We zouden dan ook moeten kijken of we soortgelijke uitbreidingen bij de concrete operaties vinden en die 'post concrete operaties' noemen. Dat is nooit gebeurd. Kallio (2011) probeert de conceptuele verwarring rond de postformele operaties op te lossen door het concept *integrative thinking* te introduceren. Het verwijst naar modellen van expertise, kritisch denken en relativistisch-dialectisch denken. Voor het Piaget systeem is dit geen vernieuwing, maar een uitwerking.

Piagetiaanse psychologie nu Na zijn dood in 1980 is Genève geen centrum meer voor een brede studie van de ontwikkeling van cognitie. Er is een Jean Piaget Society die zijn denkstijl en wijsgerige bijdrage levend houdt. De belangstelling is verschoven. Op kleine schaal inspireert zijn theorie nog steeds tot empirisch onderzoek. In de VS bijvoorbeeld in het werk van Robert Siegler. Een concreet voorbeeld is de balanstak die gebruikt is door Boom et al. (2002). Kinderen moeten ontdekken dat het evenwicht van de balans een functie is van de afstand tot het middelpunt en van het gewicht. De onderzoekers leiden uit de antwoorden van de kinderen stadia af. Deze gaan van een willekeurig antwoord naar het inzicht dat het evenwicht een vermenigvuldiging is van de eenheden afstand en aantal gelijke gewichtjes. De stappen vormen een ontwikkelingsschaal. De kinderen isoleren uiteindelijk de twee relevante variabelen en hun relatie, die van vermenigvuldigen. Uitwerking van de theorie vindt op beperkte schaal plaats. De handeling gebaseerde theorie: de sensomotoriek als basis voor kennen, wordt nog steeds geanalyseerd (Allen & Bickhard, 2013).

Ondanks de afname van belangstelling zijn Piagets bijdragen van belang voor de diagnosticus. Het is in de Piagetiaanse traditie gewettigd om kleine experimentjes te doen, kinderen vragen te stellen en evenveel te leren van hun foute als van hun goede antwoorden. Het onderscheid in stadia is zinvol hoewel het zelden in instrumenten is ondergebracht. Deze traditie verruimt het arsenaal aan diagnostische middelen.

In het onderwijs is er weerstand (geweest) tegen deze opvatting van cognitieve ontwikkeling. Ze gaat niet direct over het schoolcurriculum en de invloed van training en instructie zou te beperkt geacht zijn. De centrale stelling is dat bij leren en onderwijzen rekening moet worden gehouden met het ontwikkelingsniveau van het kind. Hieraan wordt alleen lippendienst bewezen als er geen inschatting van dat niveau van ontwikkeling plaatsvindt.

Leerkrachten zijn niet opgeleid als cognitief psychologen. Piaget pleitte ervoor dat leraren in het basis- en middelbaar onderwijs een dergelijke scholing kregen. Aan de Universiteit Utrecht is een academische opleiding voor basisschoolleraren gestart. Het is afwachten of de studenten deze waarderen en of het geen kloof schept tussen hen en de aanwezige Pabo-leraren.

Ontwikkeling van geschiktheden en prestaties Voor de ontwikkeling van geschiktheden geldt voor een deel hetzelfde als voor intelligentie. De psychometrische basis sorteert voor naar stabiliteit en continuïteit en gedifferentieerde factoren worden niet opgevat als het resultaat van cognitieve ontwikkeling. Ze worden bovendien gebruikt voor leeftijdsgroepen die het stadium van de formele operaties gepasseerd zijn. De ontwikkeling van prestaties op school, in opleidingen en beroep is rechtstreeks verbonden met het curriculum en zijn doelen. Het ligt voor de hand dat niet iedere inhoud geleerd kan worden op elke leeftijd. Daar wordt rekening mee gehouden zonder dat er een ontwikkelingsidee aan ten grondslag ligt. Er zijn leerlingvolgsystemen ontworpen met moderne psychometrische (IRT) modellen door het Cito. Deze laten zien in hoeverre een leerling de lijn en stappen van het curriculum

volgt of ervan afwijkt. Men probeert - gegeven de trouw aan het idee van één latente trek (een Rasch schaal) - de items van bijvoorbeeld rekenen over leerjaren heen te tillen. Zo kunnen de prestaties vanaf groep 3 tot groep 8 rechtstreeks vergeleken worden. Er is daar geen sprake van verschillende stadia.

Samenvatting en conclusie

Intelligentie, cognitie en geschiktheden ontwikkelen gedurende het leven en prestaties worden geacht beter te worden als gevolg van scholing en instructie. De bekendste theorie van cognitieve ontwikkeling is van Piaget. Deze gaat niet over individuele verschillen maar over de ontwikkeling van het kennen: van reflexen van de baby tot de formele operaties van de adolescent en van de professional die zijn vak beoefent. Er worden vier stadia in de cognitieve ontwikkeling onderscheiden: sensomotorisch (0-2 jaar), pre-operationeel (2-7 jaar), concreet operationeel (7-12 jaar) en formeel operationeel (12-18 jaar of misschien nog later: 21 jaar). De stadia zijn kwalitatief onderscheiden, volgen elkaar op volgens een vast patroon en hebben een vast en eender eindpunt. De interpretatie verschilt: een onderliggende formeel weer te geven structuur, een hermeneutisch te vormen onderliggende betekenis of een empirisch af te leiden factor. Er is geen winnende interpretatie en er is kritiek op alle drie. Piagets theorie staat na zijn dood in 1980 niet meer centraal. Zijn werkwijze (klinisch interview, proefjes en taken) wordt nog wel gebruikt. In onderzoek naar regels bij het oplossen van natuurkundige en socio-morele vraagstukken leeft zijn idee van stadiumgewijze ontwikkeling nog voort (Kohlberg). De diagnosticus maakt weinig gebruik van de cognitieve-stadiumtheorie, hoewel het vaststellen van het stadium van diagnostische waarde kan zijn. Bovendien kunnen de wijze van ondervragen en het voorleggen van proefjes en taken nuttig zijn als diagnostisch middel.

Geschiktheden veranderen of ontwikkelen onmiskenbaar. Ze worden echter gedefinieerd als latente trekken en er wordt van stabiliteit en continuïteit uitgegaan. De 'trekken' worden bijvoorbeeld gepast in een Rasch model dat de prestaties van groep 3 tot groep 8 vergelijkbaar maakt. Daarnaast worden prestaties bepaald ook bij personen die de periode van de formele operaties gepasseerd zijn.

De ontwikkeling van prestaties is 'vastgelegd' of 'gestipuleerd' in het curriculum en de onderwijsdoelen. Er ligt geen expliciete ontwikkelingstheorie aan ten grondslag. Wel is er het formele Rasch model dat prestaties over de basisschoolperiode vergelijkbaar maakt: het leerlingvolgsysteem van het Cito. Het ligt overigens voor de hand dat niet iedere inhoud op elke leeftijd geleerd kan worden.

3. Sociale context

Context verwijst naar een willekeurig gemanipuleerde of naar een natuurlijk aanwezige sociale en fysische omgeving. Intelligentie is met veel contextvariabelen gecorreleerd. Bekende thema's zijn: genen en IQ, sekse en IQ, inkomen en IQ, mortaliteit en IQ, IQ verschillen tussen landen, staten, en toename van het IQ over generaties.

Genen en IQ: De rol van de sociale en fysische omgeving bij het verklaren van variantie in intelligentie (IQ) is ten eerste het thema van de gedragsgenetica (GG). Dat bestaat uit een lineair model dat varianties over bronnen verdeelt en optelt. De helft gaat naar de genetische variatie en ongeveer 20% wordt toegeschreven aan verschillen binnen en tussen gezinnen. De SES van gezinnen bepaalt voor een groot deel de sociale context van de cliënt, bijvoorbeeld de opvoedingsstijl van ouders, de school die men bezoekt en de buurt waar men woont. Omdat de GG-modellen optellen tot 100%, gaat een afname in SES variantie (als mensen gelijk worden) gepaard met een toename van variantie toegeschreven aan genetische factoren. SES hangt weer samen met gezinsinkomen en toegang tot cultuur.

Sekse en IQ: Het bestuderen van sekseverschillen is net als bij de BF bijna entertainment, maar daar is iedereen het niet mee eens (zie bijvoorbeeld Eagle et al., 2012). Calvin et al. (2010) compileerden onderzoek bij 175.000 elfjarige Engelse schoolkinderen. De resultaten zijn voorbeeldig te noemen. Calvin et al.'s gegevens bevestigen de bekende sekse verschillen in IQ en cognitie:

Sekse verschillen zijn klein en de G-factor verklaart de sekseverschillen in lees-, schrijf- en rekenprestaties niet. De score op de G-factor is iets hoger bij meisjes, maar met een verwaarloosbare effectgrootte (Cohens $d = .01$). Meisjes behalen hogere scores op de verbale residuele factor (na verwijdering van de G-factor variantie): 28% behaalt 1.0 SD hoger en jongens hogere scores op de kwantitatieve residuele factor (28% behaalt 1.0 SD hoger). Er is geen verschil op een niet-verbale residuele factor. In Europa en de VS behalen meisjes de *governmental targets* (de vereiste SAT of Cito-scores beter dan jongens: 10% meer meisjes op taal (Engels), niveau 4 en 5% meer meisjes op rekenen niveau 5. De G-factor is een goede voorspeller van de onderwijsfactorscore ($r = .83$). Dit is een hoge waarde, meestal tref je waarden tussen de .40 en .70 aan. Het zou nuttig zijn als de onderzoekers de betrouwbaarheidsintervallen rond de waarden zouden vermelden.

Inkomen en IQ: De hoeveelheid verdiend geld en IQ correleren in een groot aantal landen tussen de .31 en .37 (Lynn & Vanhanen, 2002, 2012). Stolarski (2013) combineerde intelligentie en persoonlijkheid in verschillende staten en brachten de twee in verband met economisch succes. Eén en vijftig landen van regio's als Europa (protestant en katholiek), Engeland, voormalige communistische landen, Latijns Amerika. Midden Oosten, Zuid Azië, Oost Azië en Zuid Afrika waren erbij betrokken. Extraversie, Openheid en Aangenaamheid waren op nationaal niveau significant ($r = .52$) en positief gecorreleerd met nationaal IQ. Een model met alleen Extraversie en Openheid was eveneens significant gecorreleerd ($p = .005$). Een regressieanalyse liet zien dat het Bruto Nationaal Product (BNP) in 50 landen significant voorspeld kon worden vanuit de nationale IQs en de Grote Vijf ($p = < .001$). Ze verklaarden 70% van de variantie in het BNP. De belangrijkste voorspellers waren IQ en Extraversie. Er was grote variantie in IQs van de landen en ze verschilden aanzienlijk in toegankelijkheid tot goed onderwijs. Dat zorgt voor uitbijters die de correlatie 'opblazen'. Dat maakt de grote hoeveelheid verklaarde variantie mogelijk. Het blijft lastig om nationaal IQ (het IQ van Latijns Amerika, etc.) betrouwbaar te schatten. Niettemin is er steun voor de hypothese dat

IQs van landen, die op hun beurt gecorreleerd zijn met opleiding en persoonlijkheid ('de Extraversie van de Europeaan, etc. '), economisch succes en inkomen voorspellen.

Sterftcijfers en IQ: Barnes et al. (2013) vonden in 267 provincies uit 37 landen met IQs tussen de 86 en 116 een relatie tussen IQ en sterftcijfer: kindersterfte: $r = -.38$, sterfte tussen 1 en 4 jaar: $r = -.19$, 4 tot 15 jaar: $r = -.18$, 15-25 jaar $r = -.21$; AIDS en sterftcijfer en IQ: $r = -.44$. Er is dus een lage tot gemiddelde correlatie tussen IQ en sterftcijfer op verschillende leeftijden in 37 staten. Er werd een achterstandsmaat voor elk land gemaakt die berustte op het percentage inwoners dat van de staat afhankelijk is, percentage zwarte inwoners en het percentage alleen door vrouwen geleide huishoudens. Deze maat correleerde $r = .49$ met IQ. Het hoogste sterftcijfer kwam voor bij huishoudens waar de vrouwen er alleen voorstonden: $r = -.74$.

Gezondheid en IQ: Wraw et al. (2015) benutten het gegevensbestand van de *National Longitudinal Study of Youth* (NLSY cohort 1979; 5793 deelnemers) om de relatie tussen pre-morbide intelligentie en gezondheid te bepalen op 50-jarige leeftijd. Er werden 16 maten voor gezondheid gebruikt met twee hoofdklassen: lichamelijke gezondheid en functionele beperkingen. De meeste maten waren objectief, vier bestonden uit zelfrapportages. Als te verwachten en zoals iedere huisarts je kan vertellen heb je meer kans op een gezond leven als je uit een hogere socio-economische klasse komt. In de studie bleek hogere pre-morbide intelligentie gecorreleerd met een betere fysieke conditie op 50-jarige leeftijd en met een lager risico op chronische ziekten. In SD hoger IQ (115) verhoogde de kans op goede, zeer goede of excellente gezondheid (odds ratio 1.70; betrouwbaarheidsinterval 1.55-1.86). Dertien gezondheidsmaten waren significant negatief gecorreleerd met IQ in de jeugd (odds ratio van .85 voor diabetes tot .65 voor herseninfarct per 1 SD hogere IQ score). De SES op 50-jarige leeftijd hing samen met kans op ziekten. Dat was minder het geval voor de SES op jeugdige leeftijd. Je kunt immers van SES veranderen in een periode dat de middenklasse groter wordt door brede economische vooruitgang. Door de economische crisis vanaf 2008 kan de middenklasse weer verarmen en dat doet de kans op ziekten weer toenemen. Lagere intelligentie is dus een risicofactor voor slechte gezondheid en draagt bij aan gezondheidsongelijkheid.

Salahodjaev en Azam (2015) bepaalden de samenhang tussen *body mass indices* (BMI) en IQ van 187 landen. Dat is een lastig te interpreteren verband. Als je de plots bekijkt is er een zwak positief verband: hoe dikker hoe intelligenter; je kunt ze ook lezen als omgekeerde U: bij laag IQ BMI onder de 25, tussen IQ 80 en 90 iets boven de 25, en rond de 105 weer onder de 25. Een BMI van 25 is gemiddeld. Het keerpunt voor BMI (mannen) ligt bijvoorbeeld bij een IQ van 83.7. Bij de 187 landen zijn ook onderontwikkelde waar honger heerst. Wellicht is het handiger niveaus van obesitas te bepalen in de verschillenden landen en daarvan de IQs te vergelijken.

IQs in verschillende landen: Lynn en Meisenberg (2010) bepaalden de correlatie van IQ met onderwijsresultaten op *math*, *science* en *reading* in 86 landen. Zij komen tot $r = .917$, dat wil zeggen dat ze vrijwel samenvallen. Ze vergeleken de IQs van 108 naties of landen en vonden uiteraard grote verschillen. De volgende trends zijn zichtbaar: 0,33 tot 0,50 SD boven het gemiddelde scoren inwoners van China, Hong Kong, Japan, Zuid Korea, Taiwan en Singapore. Onder het gemiddelde (1,0 SD tot 2,0 SD) scoren inwoners uit Ghana, Malawi, Mozambique, Namibia, Tanzania, Uganda, Zimbabwe en Zambia. Europa is gemiddeld, score rond de 100. Deze data suggereren een hogere correlatie dan de .31 en .37 van Lynn en Vanhanen. Ze laten ook en vooral de rol van scholing voor het IQ zien. Op school leren we immers *math*, *science* en *reading*. Zodra inwoners van de landen toegang tot hoogwaardig onderwijs krijgen is een faire vergelijking tussen hun IQs mogelijk. De Lynn en Meisenberg resultaten zouden zelfs kunnen suggereren dat ze verdwijnen want de correlatie tussen onderwijsresultaten en IQ is in hun studie bijna perfect.

Goddard sprak meer dan 100 jaar geleden over immigranten uit *Europa* naar de VS als *morons*, een term even denigrerend als ons 'debiel'. Zij kwamen de VS binnen aan het eind van de 19^{de} en begin 20^{ste} eeuw. In 1927 legde Goddard in een 'wetenschappelijk' tijdschrift (*Scientific Monthly*) uit waardoor een *moron* gekenmerkt was. Lynn en Meisenberg leggen een eeuw later de nadruk *niet* op kenmerken van personen maar op de *toegang tot onderwijs*.

Deze voor de hand liggende interpretatie wordt al 30 jaar betwist door Rushton en Jensen. Ze geven de voorkeur aan een genetische interpretatie van IQ verschillen, inkomen, misdaad en HIV/AIDS. Een voorbeeld is een studie van Templer & Ruhston (2011) waarin de gegevens van 50 staten van de VS verwerkt zijn. Ze berekenden correlaties tussen 11 variabelen, onder meer IQ, huidskleur, aantal geboorten, kindersterfte, levensverwachting, HIV/AIDS, geweldsmisdrijven en *state income* (GPD), dat er als volgt uitziet:

Gross private consumption + private consumption + gross investment + government spending + (export - import).

De *positive manifold* van deze variabelen leverde een eerste principale component op die 33% van de variantie voor zijn rekening nam. De mediane factorlading op de eerste component (factor) was .34. De auteurs wijzen op de correlaties tussen HIV/AIDS en misdrijven met moorden en berovingen ($r = .92$) en huidskleur (percentage zwart) en moord ($r = .84$). Hun voorkeur voor een genetische interpretatie blijkt uit de vergelijking, waarin het verschil tussen de correlaties van totale geweldsmisdrijven met huidskleur - een biologisch bepaalde variabele - en GPD: $r = .55$ en $r = -.17$ (p. 441). Hun interpretatie van deze resultaten is gekritiseerd. Wicherts et al. (2010) vonden ongeveer dezelfde resultaten in een vergelijking tussen 60 landen. Zij interpreteerden de covariaties als resultaat van achterstand in ontwikkeling. Het werk van de Canadese evolutionaire psycholoog Rushton is soms als opruiend, pseudowetenschappelijk en racistisch gekarakteriseerd. Gottfredson (2012) probeert balans te brengen in de bijdragen van Rushton en wijst op autoritaire

argumenten, vooroordelen en politieke meningen ten opzichte van zijn werk. De feiten blijven overeind: welvaart, IQ, inkomen, opleiding, enzovoort zijn *ongelijk verdeeld* tussen en binnen landen.

Roivonen (2010) vergeleek de WAIS III prestaties van Europeanen en Amerikanen. De WAIS Perceptuele Organisatie Index (factor) was hoger (5-10 punten; 1,0 SD = 15) bij de Duitse, Spaanse, Britse en Finse deelnemers ten opzichte van de Amerikaanse. De auteur schrijft het verschil toe aan verschillen in onderwijs en houding tegenover testen. Dit resultaat laat zien dat testnormering van een land niet voor een ander land gebruikt kan worden.

Toename van IQ over generaties: Intelligentie en prestatie worden gewaardeerd door ouders, leraren en werkgevers. Er zijn programma's ontwikkeld om het gemiddelde niveau van intelligentie en prestaties te verhogen. De concepten voor deze interventies zijn: 'compensatie' dat verwijst naar een tekort. Dat kan aangevuld worden door bepaalde activiteiten. 'Verrijking' betekent dat het niveau verhoogd kan worden door de sociale context en de materialen te verrijken. Deze programma's zijn ontwikkeld als reactie op het ontdekken van achterstanden bij bepaalde groepen en op het idee dat groepen tekortkomen, bijvoorbeeld hoogbegaafden. De effectgrootten van deze interventies zijn matig tot gemiddeld. *Het Flynn effect* gaat rechtstreeks over de toename van IQs over generaties. Globaal neemt het IQ per generatie tussen de 0.5 en 1.0 standaarddeviatie toe volgens Flynn. Het wordt opgevat als een effect van de omgeving maar het is natuurlijk gecompliceerder.

Geschiktheid staat direct in verband met training en instructie. Geschiktheidsniveaus worden gebruikt om studenten en personeel te selecteren. De scores op SATs, Cito-toetsen helpen vast te stellen of en hoeveel training nodig is. Ze worden ook gebruikt om succes van een training te voorspellen en te evalueren. Als er geselecteerd en toegelaten wordt aan de hand van geschiktheidstests krijgt een beperkte groep een opleiding of training. Dit leidt tot het *restriction of range* probleem want de geschiktheidsscores zijn een segment van de scores. Bij complexe taken blijkt de correlatie tussen intelligentie/ geschiktheid en werkprestatie tussen de $r = -.45$ en $r = .67$ te liggen (Lynn & Vanhanen, 2002).

Prestatie is direct verbonden met scholing en opleiding. Het doel van de interventies is het bereiken van doelstellingen, omschreven als niveaus van beheersing van leerstof en specifieke vaardigheden. Leerlingen, studenten en werknemers moeten vooraf vastgelegde niveaus behalen. Geschiktheid en vooral prestaties behoren tot het domein van de onderwijspsychologie. Daar is veel kennis over de effectiviteit van curricula, scholing, instructie en training. Als voorspeller van prestatie wordt het IQ het meest gebruikt. De correlaties tussen IQ en schoolprestaties liggen tussen de .57 en .72 (Lynn & Vanhanen, 2002). Deze correlaties zijn berekend in nationale steekproeven. Er is ook al vermeld dat de relatie tussen IQ en toegang tot onderwijs bijna perfect is: er is een waarde gerapporteerd van $r = .92$.

Naast de theorie over de bijdragen van de omgeving aan de variantie in IQ in de GG-studies, is de relatie van kenmerken van de natuurlijke *sociale context* met intelligentie, geschiktheid en prestatie bepaald. Deze omgevingen zijn niet bij voorbaat stabiel en men kan er vaak slechts beperkt invloed op uitoefenen. Voorbeelden zijn gezinskenmerken (SES, opvoedingsstijl), school (stijlen van leraar-leerling interactie, moreel klimaat op school, niveau van het personeel), leeftijdgenoten, klas en leraar. Het onderzoek naar de relatie van ontworpen en aangetroffen sociale contexten met IQ, geschiktheid en prestatie heeft als doel prestaties te verhogen. Men neemt meestal niet aan dat intelligentie (de G-factor en G 'fluid') gemakkelijk beïnvloed kan worden.

Er is geen taxonomie van de natuurlijke sociale contexten en ze zijn niet in een kader gezet. Interventies zijn op het behalen van succes gericht. Vaak worden ze ontwikkeld als er een achterstand wordt waargenomen. De programma's zijn kostbaar en van lange duur. Methodologen worden er regelmatig achteraf bij gehaald om nog iets van de resultaten te maken. Was de interventie wel de oorzaak van de verandering? Bovendien lijkt het erop dat dure interventies niet *mogen* mislukken. Dit type onderzoek is gefragmenteerd en de effectgrootten zijn meestal bescheiden.

Samenvatting en conclusie

Gemanipuleerde en natuurlijke omgevingen zijn in verband gebracht met intelligentie, geschiktheid en prestaties. Gedragsgenetica verdeelt de IQ-variantie ongeveer half om half over genetische en de gedeelde + niet-gedeelde omgevingsfactoren + onbekende factoren. Het doel van interventies is het verhogen van prestaties en het vergroten van geschiktheden. De samenhangen en effectgrootten zijn bescheiden tot gemiddeld. Het IQ wordt meestal als een absoluut en relatief stabiel vermogen beschouwd en betrekkelijk ongevoelig voor interventies. Bestaande omgevingskenmerken van gezin, leeftijdgenoten en school hangen significant, maar matig tot gemiddeld samen met prestaties en geschiktheden. Sekseverschillen zijn klein. Meisjes behalen in geïndustrialiseerde landen de doelen van de staat (SATs en Cito-scores) wat beter dan jongens. Er wordt steeds verschil gemaakt tussen talige en ruimtelijke, wiskundige geschiktheden. IQ is gemiddeld gecorreleerd met inkomen. De grote verschillen in IQ tussen landen hangen samen met de toegang tot goed onderwijs. Het Flynn effect drukt uit dat er per generatie (ongeveer 25 tot 30 jaar) een halve tot een hele standaarddeviatie winst gemaakt wordt in IQ; per decade wordt doorgaans 0,3 SD vermeld. Prestaties zijn uiteraard gecorreleerd met scholing en training.

4. Reflectie en evaluatie

Uit ervaring, literatuur en onderzoek weet de diagnosticus dat intelligentie geen neutraal begrip is. Het is een individueel verschil variabele waaraan men waarde hecht. Westerlingen overschatten hun IQ met ongeveer een halve standaarddeviatie ten opzichte van het gemiddelde van 100 bij een SD van 15. Het begrip sluit voor leken praktische en interpersoonlijke vaardigheden in om problemen uit het dagelijks leven het hoofd te bieden.

Het begrip reflecteert vooroordelen over wie intelligent is en wie niet en suggereert soms sociale mobiliteit. De nadruk op praktische intelligentie klinkt door in de uitspraak dat men rijk kan worden zonder opleiding.

Expliciete psychometrische theorie over individuele verschillen levert G en S factoren, PMA en SOI en hiërarchische modellen voor intelligentie en geschiktheden op. Dit leert evenveel over factoranalyse als over intelligentie. Er is geen definitief systeem of dwingend idee om de correlaties tussen de tests en taken te interpreteren. Een door ieder aanvaarde architectuur is ook niet nodig. De diagnosticus weet voor welk probleem of vraag van de cliënt hij tests gebruikt. Is het complex en gearticuleerd in onafhankelijke onderdelen dan ligt een algemene G niet voor de hand. Een keuze uit 120 soorten is niet realistisch. Hij moet een tussenweg vinden. In de praktijk is de keuze soms al vastgelegd in protocollen van bedrijven en instellingen.

De aanvulling van informatieverwerkingstheoretici is interessant maar niet meteen bruikbaar. Ze voorspellen niet beter dan de klassieke tests. Ze stellen in staat deelprocessen van intelligent handelen te onderscheiden. Metingen van EEG, EVPs hersenscans en misschien het DNA vallen voornamelijk buiten het diagnostisch handwerk. Deze uitslagen zijn niet zonder specifieke training te lezen. Bovendien vergt het dure apparatuur die niet iedereen zomaar kan kopen en bedienen. Het is niet realistisch om op korte of middellange termijn iets van deze metingen te verwachten. Bovendien zijn de correlaties met criteria niet hoger dan die met klassieke instrumenten.

Intelligentiemeting (G, S en multiple factoren) is nuttig. Hij voorspelt schoolresultaten (*grade point average*: GPA) met een geschatte r van .50, het aantal jaren opleiding en het behalen van diploma's: r geschat = .55. Voor succes in het beroep ligt de correlatie ook rond de .50 na correctie voor administratieve fouten. Er zijn overigens steeds aanzienlijke variaties in de vele studies. De relatie tussen intelligentie en schoolresultaten toont grote verschillen over landen. Ze zijn volgens Lynn en Meisenberg toe te schrijven aan verschil in scholing in rekenen, lezen, taal en *science*.

De geschiktheidsmeting lijkt op de multiple intelligentiemeting. Er is theorie om de factoren te interpreteren. Voor deze theorievorming waren complexe en realistische taken voor beroepen als piloot, operator, enzovoort het uitgangspunt. Dat heeft ertoe geleid dat naast de bekende factoren motorische, perceptuele en fysieke vaardigheden als bouwstenen van geschiktheden erkend worden. Voor sommige taken is de abstracte, schoolse, formele intelligentie niet genoeg. Er is een concrete goed te meten aanvulling nodig op de rationalistisch geïnspireerde intelligentie- en geschiktheidsmetingen. Dat blijkt moeilijker dan verwacht (zie het volgende hoofdstuk).

De discussie over de overerfbaarheid van IQ en geschiktheid leidt tot de verdeling van de varianties over genetische en omgevingsfactoren. De laatste worden dicht bij huis gezocht: verschillen tussen tweelingen, broers en zussen en ouders. Aan de schatting van de variantiecomponenten heeft een diagnosticus niet zoveel. Wat doet die ertoe voor de diagnose naar aanleiding van een vraag of probleem van een cliënt? De kennis over hoezeer cliënten op hun familie lijken en wat de 'huiselijke omgeving' bijdraagt behoren tot het

kennisbestand van de diagnosticus, maar biedt geen directe hulp voor diagnostiek. Kinderen 'erven' bovendien van hun ouders niet alleen genen, maar mogelijk ook een epigenetisch patroon en zeker een ecologische niche (sociaal, biologisch en fysisch) en voorsortering in hoe ze iets waarnemen en zich gedragen (culturele *beliefs* en praktijken).

Theorie over prestatie is er in de vorm van analyses van leren lezen, rekenen, taal, grammatica en wetenschappelijke vakken. Deze speelt voor de diagnosticus echter een beperkte rol. De metingen van prestaties die hij onder ogen krijgt zijn het resultaat van experts op een inhoudelijk gebied en vooral van IRT deskundigen. De laatsten beslissen of de items voldoen aan een model en verwijderen items die niet voldoen. Dat levert de diagnosticus psychometrisch verantwoorde schattingen van de prestaties van een leerling, sollicitant of cliënt op. Over de inhoud van een vak en de leerdoelen is niet altijd overeenstemming; bovendien zijn de itemschrijvers geneigd een beperkt aantal doelen te 'dekken', bijvoorbeeld nadruk op kennis en minder op toepassing, inzicht en oordeelsvorming. De diagnosticus kan er geen direct advies uit afleiden over wat te doen als de prestaties afwijken omdat de instrumenten geen informatie bevatten over de oorzaken van de afwijkingen.

Piagetiaanse ontwikkelingstheorie en het stadiumconcept spelen nauwelijks een rol in de diagnostiek van intelligentie en cognitie. De klinische methode dat wil zeggen dat de diagnosticus relevante taken bedenkt om gedrag van een cliënt te beschrijven kan vruchtbaar zijn. De kans is op toepassing is niet groot want het is geen gestandaardiseerde aanpak en lokt gauw het verwijt uit dat de diagnosticus subjectief te werk gaat. De kennis uit de theorie over de kwalitatieve verschillen bij cliënten tussen de 1 en 18/21 jaar bij het uitleg geven over verbale (bijvoorbeeld categorisering), morele, logische en wiskundige taken is interessant. Het is echter onwaarschijnlijk dat een diagnostisch rapport daarover informatie verstrekt. Kortom, de Piagetiaanse ontwikkelingstheorie speelt een geringe rol in het diagnostisch handwerk.

Manipuleerbare en natuurlijke sociale contexten zijn niet strikt gescheiden. Een diagnosticus kan immers aanbevelen om kenmerken van de natuurlijke omgeving van een cliënt te wijzigen. De gedragsgenetica schrijft de helft van de populatievariantie toe aan genetische verschillen. Het overige is voor rekening van de sociale context. Geschiktheden worden bepaald met het oog op de vraag of opleiding en training het beoogde effect zullen hebben. Er is veel en gefragmenteerde kennis over de relaties tussen contextkenmerken en intelligentie, geschiktheid en prestaties. De gepubliceerde correlaties zijn significant en variëren van bescheiden tot gemiddeld. De diagnosticus heeft daar kennis van op grond van zijn ervaring en uit de literatuur maar die is niet zomaar toe te passen een concrete cliënt. Er zijn vele en uiteenlopende correlaties tussen kenmerken van gezin, broers, zussen, buurt, klas, school, subcultuur en intelligentie en geschiktheden en prestaties. Deze zijn niet te verwaarlozen maar de massa aan correlaties en het ontbreken van een taxonomie maken toepassing op concrete groepen cliënten moeilijk, laat staan op individuen. Dit kan tot gevolg hebben dat de diagnosticus zich meer richt op de meting van de kenmerken van de persoon. De relevante kenmerken van de context moet hij zelf uitvinden en vaststellen.

Evenals bij de analyse van typerend gedrag van de cliënt kun je bij intelligentie, geschiktheid en prestatie niet om de vraag heen wat ze *zijn*. Kennen en inzicht zijn uitdrukking van een *relatie* tussen kenner en buitenwereld. In de zelf betrokken wetenschap van de psychologie moeten we daar aan toevoegen: de relatie van de kenner met andere personen en met zichzelf. 'Ken uzelf', zeiden de Griekse filosofen al. U kunt een onnadrukkelijke uitwerking van de relatie terugvinden in de individuele verschillen oriëntatie: Spearman legt uit wat de kenner doet met de werkelijkheid van verschijnselen, bijvoorbeeld de eductie en reductie. We halen er iets uit en we beperken wat we waarnemen. Hij kritiseerde Binet (niet theoretisch genoeg) en Thorndike (behavioristisch en naïef realistisch). Borsboom erkent het feit dat kennen intelligentie altijd betrekking heeft op een relatie tussen mij, de werkelijkheid en mijzelf. Het gaat bij een dergelijke constatering natuurlijk om de precieze uitwerking. Thorndike koppelde evenals Horn & Cattell en Guilford kenmerken van de kenner aan formele en inhoudelijke aspecten van de werkelijkheid. Er zou een interpretatieve studie en conceptuele analyse nodig zijn om de verhouding die deze auteurs onnadrukkelijk verwoorden expliciet te maken. Binnen de psychologie en de filosofie wordt dit nauwelijks gethematiseerd.

Bij informatieverwerking staat de relatie tussen kenner en werkelijkheid centraal. Wat komt van mij en wat van de (sociale) werkelijkheid? Het begrip *affordance* van de Gibsons zegt iets over die relatie. Het is er een van betrokkenheid en uitdaging. De werkelijkheid is niet alleen lengte, breedte en hoogte. Voor Sternberg geldt dat ook. Hij werkt uit hoe we informatie verwerken als kenner. Embretson heeft een rationalistische inslag met haar stappenplan. De kenner volgt niet alleen stappen. Hij is grillig.

Bij Wechsler regeert de pragmatiek in tweevoudige zin: a. maak een test, geschikt om bij bezoekers van een groot hospitaal de intelligentie vast te stellen en b. past het resultaat, de interpretatie aan bij heersende verklaringsvoorstellen van genetisch tot de G-factor tot informatieverwerkende mechanismen.

De verklaring met behulp van de kleinste eenheden is wellicht ontstaan door aansluiting bij de biologie met haar eenheden en structuren en bij de successen van de natuurkunde: bijvoorbeeld de deeltjesfysica: het kleinste deel dat aan de basis van al wat is staat. Het is misschien ook de overtuiging als ik alles van het kleinste weet, kan ik veilig de complexiteit van intelligent gedrag met behulp van een architectuur weergeven en bijgevolg doorzien. Het berust op een overtuiging over het beste aggregatieniveau. Dat was in de natuurkunde al een vraagstuk: het gedrag van de kleinste eenheden zou eerst onder probabilistische formules gevangen kunnen worden na een zekere aggregatie (Schrödinger, 1944).

Piagets ontwikkelingstheorie heeft nadrukkelijk een wijsgerig uitgangspunt: hoe de ontwikkeling van het kennen (epistemologie) te beschrijven. De bijzonderheid bij hem is dat hij van een wijsgerig vraagstuk een psychologisch onderwerp heeft gemaakt, zoals omgekeerd Popper van probleemoplossen een wijsgerig vraagstuk heeft gemaakt. Dat is in methodologieboeken terecht gekomen, denk aan De Groot's *Methodologie*.

De contexttheorie bevat uitwerkingen van haar relatie met intelligentie en cognitie. Die is nogal wisselend. Veel onderzoek gaat over de correlatie tussen de natuurlijke context,

bijvoorbeeld ouderlijk SES en IQ en de correlatie tussen de gefabriceerde context, bijvoorbeeld scholing en IQ. De Gibsons werken de relatie, als gezegd, uit als *affordances*. Brunswik is een voorbeeld van een bestemming van de relatie kenner/waarnemer en werkelijkheid met representatieve designs en ecologische validiteit.

Het is wellicht de moeite waard naast de typerend gedrag ook intelligent handelen en cognitie te onderzoeken op de wijsgerige vooronderstellingen. Dat laat zien dat er niet één of een dominant perspectief is dat intelligentie en cognitie afdoende en bevredigd beschrijft en verklaart. We beschikken immers nooit over een *régard survolante*.

IV Alternatieve theorieën van intelligentie en prestaties

Alternatieve theorieën zijn een reactie op expliciete theorievorming. Er wordt bijvoorbeeld beweerd dat de inhoud eenzijdig gedefinieerd is, er iets aan de methodologie ontbreekt of weer anders, dat testen in onderwijs en hulpverlening te veel aandacht krijgt.

Welke alternatieve opvattingen over individuele verschillen in intelligentie en geschiktheid zijn er en op welke gebreken wijzen ze? Zijn er alternatieve opvattingen voor prestatie? Is er kritiek op onderwijsdoelstellingen en schoolse vaardigheden? Sluiten deze aan bij vaardigheden die er in het dagelijks leven toe doen? Zijn SATs en Cito-toetsen aanvaardde en betrouwbare indices voor prestaties?

Welke zijn de alternatieven voor de zwakke en sterke ontwikkelingstheorieën? Ligt de nadruk bij intelligentie en cognitie niet eenzijdig op informatieverwerking? Wat betekent het sociaal constructivisme voor de opvatting over de sociale en culturele context met het oog op intelligentie, cognitie, geschiktheid en prestatie?

1. Individuele verschillen

Intelligentietheorie is gericht op schoolse, rationele, logische (*academic*) vaardigheden en op het vermogen tot abstraheren. De nadruk bij alternatieve theorievorming ligt op praktische, sociale en emotionele capaciteiten en inzicht in en aanvoelen van culturele vereisten. Voorbeelden: Natuurlijke Intelligentie (Gardner), Competentie (bijvoorbeeld Howard & Driscoll), Leerpotentieel (Grigorenko), Succesvolle intelligentie (Sternberg) en Emotionele Intelligentie (Mayer et al.).

Naturalistic Intelligence Gardner formuleerde (1983, 1984, 1993) een naturalistische benadering van intelligentie die psychologen en leken heeft aangesproken. Hij gaat evenals Thurstone uit van meervoudige intelligenties en wil deze recht doen door gedragingen in de natuurlijke omgeving gedurende langere tijd te observeren en te analyseren. Met behulp van kwalitatief en kwantitatief onderzoek beschrijft hij individuele verschillen in motivatie en initiatief, in muzikale, motorische en sociale vaardigheden en in interesses en culturele prestaties. Gardner gebruikt objectieve tests én observeert en ondervraagt proefpersonen bij het uitvoeren van taken die hen interesseren. Intelligentie bestaat volgens hem uit zeven factoren of componenten. Hij komt tot de 'factoren' op grond van bestaande theorie en zijn conceptuele analyse:

I Linguïstisch

II Logisch-wiskundig

III Ruimtelijk

IV Muzikaal

V Lichamelijk-motorisch (bijvoorbeeld dans)

VI Interpersoonlijk

VII Intrapersoonlijk

De laatste vier zijn een aanvulling op de psychometrische intelligentie factoren. Dit is gerechtvaardigd volgens Gardner omdat we producten die we in onze cultuur waarderen verwaarloosd hebben als te verwerven intellectuele vermogens, vaardigheden. De zeven componenten zijn zelfstandig maar functioneren niet geïsoleerd. De G-factor bestaat wel maar je hebt er niet veel aan zegt Gardner in navolging van Thurstone. Volgens hem zijn cultuur en scholing van invloed op de individuele verschillen in intelligentie. De zeven factoren klinken aantrekkelijk maar zijn een nachtmerrie voor testconstructeurs. Hoe moet hij ze operationaliseren en op testtheoretisch verantwoorde wijze meten? En zijn ze stabiel als de culturele context zozeer van invloed is? Er zijn beperkte pogingen gedaan om de factoren te meten. Gardner is ervan overtuigd dat deze typen van belang zijn in het onderwijs. Dat heeft ertoe geleid dat er bij zijn werkgever, de Harvard Universiteit, een school voor hoogbegaafden is opgericht onder leiding van zijn echtgenote met een curriculum dat op deze intelligentiefactoren is gebaseerd. Is dit Amerikaans 'vermarken' van zijn eigen interpersoonlijke intelligentie? Construct en predictieve validiteit van de factoren zijn in beperkte mate onderzocht. De resultaten wijken niet af van eerder gevonden uitslagen, vooral voor I tot en met III (factoren uit de PMA), voor VI en VII (ze komen dicht

bij de BF factoren, respectievelijk E en A en N). IV en V zijn in ander verband gemeten, bijvoorbeeld in toelatingsexamens voor muziekscholen en conservatoria en dansacademies. Deze staan echter ver van het psychometrische regime zoals we dat kennen voor tests en vragenlijsten.

Competenties Bij het selecteren van personeel en van studenten voor het hoger onderwijs zijn rond 1990 geschiktheden gelabeld als competenties. Deze moesten breder zijn dan de bekende geschiktheidsfactoren en beter op de praktijk gericht zijn dan de klassiek IQ- en vaardighedentests. Als voorbeeld worden de competenties zoals geformuleerd door Greenspan en Driscoll (1997), beschreven. Deze zijn vooral bestemd voor bedrijfs- en personeelspsychologie. Een tijdlang zijn ze door managers en personeelsfunctionarissen gebruikt om werknemers te selecteren en beoordelen. Bij de competentiebenadering ligt de nadruk op het oplossen van praktische problemen en op professionele prestaties. Intelligentie speelt een rol maar competenties zouden iets toevoegen. De auteurs onderscheiden vier domeinen:

- (1) Persoonlijke competenties, fysieke, bijvoorbeeld visuele scherpte, hartcapaciteit, sensomotoriek en lichamelijke kracht en coördinatie.
- (2) Affectieve competenties bijvoorbeeld temperament, emotie, afleidbaarheid, karakter: sociale oriëntatie en plezier in het functioneren in een groep.
- (3) Alledaagse competenties zoals sociale en praktische intelligentie en inzicht in het leven van alledag.
- (4) Academic competenties zoals conceptuele kracht, oplossen van abstracte vraagstukken, analytische intelligentie, communicatief vaardig en het doorzien van communicatietrucs. In het personeelsselectie jargon heet dit analytisch vermogen. Deze zijn voor een groot deel terug te vinden in de bekende G-S, PMA, SOI, GATB en Gf en Gc tests. Het meten van dit domein sluit aan bij bekende theorie en instrumentatie. De overige competenties worden meestal beoordeeld door deskundigen, vooral personeelsfunctionarissen.

Onderzoek dat *incremental validity* van competenties ten opzichte van klassieke tests laat zien is schaars. De meta-studie van Schmidt en Hunter (1998) voegde 75 jaar voorspellen van werkprestaties samen en kwam tot een $r = .50$ van de klassieke tests na voor de administratieve fouten gecorrigeerd te hebben. Het zal niet eenvoudig zijn om deze waarde te overtreffen met behulp van competenties. De boodschap van managers en personeelsfunctionarissen is niettemin duidelijk. Er is meer dan abstracte intelligentie nodig en die moeten we leren kennen en vaststellen om goede werknemers te selecteren. Boeken- en abstracte kennis is niet genoeg. Er zijn andere competenties nodig voor werknemers als HBO-verpleegkundigen, HTS-ingenieurs, HEAO-ers, enzovoort. De studenten op de universiteiten zijn tot nu buiten schot gebleven hoewel geneeskundestudenten te maken krijgen met competenties als in 1, 2 en 3. Er zijn nog geen evaluatiestudies die laten zien of de competenties door deze studenten getoond worden en vervolgens of deze competenties ertoe bijdragen dat ze bijvoorbeeld betere dokters worden.

Leerpotentieel Sternberg en Grigorenko (2002) achten de klassieke opvatting van een vast en onwrikbaar IQ uit de tijd. Ze leggen de nadruk op het meten van leerpotentieel. De conventionele tests meten volgens de auteurs maar een klein deel van de zich ontwikkelende inzichten en ervaringen die nodig zijn voor aanpassing en succes in werk en in het leven. Ze stellen dynamisch testen voor en dat houdt in dat de geteste al tijdens de testafname getraind wordt, informatie krijgt waarom een antwoord fout is en tips krijgt hoe fouten te vermijden. Er zijn specifieke oefeningen voor het benutten van metacognitieve vaardigheden. De diagnosticus stelt vast, meet en oefent tegelijkertijd. De aanpak stamt van Feuerstein die successen van zijn benadering claimde bij normale, gezonde proefpersonen en bij kinderen met een beperking.

Vergelijking van de resultaten onder dynamische en conventionele condities moet verschillen te zien geven en resulteren in het beter voorspellen van criteriumgedrag. Dit type studies is schaars. Budoff trainde kinderen in het abstract redeneren en zegt vooruitgang in IQ te hebben gevonden. Campione en Brown oefenden zwakke studenten door steeds tips te geven bij het oplossen van problemen. In het Oost-Duitsland van vóór 1989 zijn substantiële resultaten geclaimd met intensieve pre- en posttest trainingen en hulp tijdens het maken van taken en opdrachten (Guthke). Resultaten uit de DDR werden gewantrouwd. Op congressen verschenen steeds dezelfde psychologen die hun succesverhaal vertelden: 95% - 100% succes. Een aarzelende houding blijft misschien op zijn plaats ondanks een case studie van Schönplflug en Lüer (2013) over het *XXII International Congress of Psychology* in Leipzig (1980) onder het DDR regime. Ze beweren dat de wetenschappelijke standaarden gehandhaafd werden ondanks de controle van de communistische partij. Niettemin zijn na de val van de Muur in 1989 veel psychologiedocenten vervangen door West-Duitse. Politieke regimes claimen succes en bieden een mooi perspectief als gevolg van hun beleid. Resultaten van training zijn meestal niet opzienbarend en tonen kleine stapjes. Bovendien is generalisatie - een kernonderwijsdoelstelling - zelden vastgesteld. Sternberg en Grigorenko's overzicht lijkt een afsluiting van het thema dat in de jaren 70 van de 20^{ste} eeuw begon. In Nederland gaat dit onderzoek nog even door (zie bijvoorbeeld Reesink, 2006).

Leerpotentieel is een poging om resultaten tijdens de rit te verbeteren. De vraag is of de training beklijft en later bij andere taken tot winst leidt. Bij leren gaat het niet alleen om de specifieke inhoud zelf. Het gaat ook om overdracht naar andere taken. Dat is beperkt geslaagd bij de leerpotentieel aanpak. Dat oefening kunst baart is niet verrassend: om ergens goed in te worden van een sport tot een vak en zelfs in de wetenschap is minstens 10.000 uur gestage oefening nodig. Soms is dat zelfs niet genoeg.

Er is een vorm van oefening die je kunt opvatten als inzetten van leerpotentieel. Er worden bijvoorbeeld zeer veel oefentoetsen ter beschikking gesteld die in een centraal databestand opgeslagen zijn. De resultaten op die toetsen worden meteen gegeven (snelle feedback). Dit is een ingenieuze uitwerking van het idee van 'oefening baart kunst'. Een voorbeeld is 'Rekentuin' van de vakgroepen Methodologie en Ontwikkelingspsychologie van de

Universiteit van Amsterdam die op internet te vinden is. Het is een voorbeeld van herhaalde en directe instructie en onderzoek moet laten zien dat dit substantieel succes oplevert.

Succesvolle intelligentie Sternberg (1997) heeft steeds de nadruk gelegd op andere capaciteiten naast de schoolse, abstracte en academische. Hij wilde praktische en succesvolle intelligentie onderscheiden van de psychometrische. Voor succes in het dagelijks leven, in beroep en het zakenleven zijn IQ, geschiktheden en schoolprestaties niet genoeg. Succes hangt af van het zo goed mogelijk inzetten van de sterke kanten van een persoon en het compenseren voor de zwakke kanten. Van Lieshouts concept van de succesvolle persoonlijkheid is een analogie van Sternbergs succesvolle intelligentie. Het begrip wordt uitgelegd aan de hand van zijn drietrappentheorie van intelligentie. Die bestaat uit:

- (a) De interne wereld met mentale mechanismen die intelligent gedrag ondersteunen: snel, selectief en adequaat informatie verwerken over problemen en mensen.
- (b) Persoonlijk ervaring die de interne wereld en zijn mechanismen verbindt met de buitenwereld vol met niet altijd duidelijke opdrachten en taken en het grillige onvoorspelbare gedrag van anderen.
- (c) De externe wereld waar de persoon zich aan aanpast met de cognitieve mechanismen van assimilatie en accommodatie.

Leerpotentieel is niet zo populair geworden bij onderzoekers en professionals als bijvoorbeeld Emotionele Intelligentie en Positieve Psychologie. Het is beperkt onderzocht door Sternbergs groep. Er is bescheiden steun voor *incremental validity*. Meestal beperkt de diagnosticus zich tot de traditionele IQ- en geschiktheidstests en dat is gegeven het succes van leerpotentieel onderzoek ook terecht.

Emotionele intelligentie De termen (EI) of emotioneel quotiënt (EQ) zijn gemunt door Mayer en Salovey (1995). Ze werden bekend door een boek van de journalist Coleman (1995). Waarom is het een populair begrip? Iedereen begrijpt het en het gaat in tegen 'koude kennis' van feiten en inzicht in abstracte ideeën. In huwelijksadvertenties en op datingsites wordt EI (al afgekort: zo bekend is het) als een fort van je persoon vermeld. Professionals zagen er heil in want een deel van de variantie in belangrijke criteria wordt niet verklaard door IQ en misschien is EI/EQ wel een kandidaat. Naast *cold cognition* kan *hot cognition* wellicht een deel van de variantie in criteriumprestaties verklaren. Mayer et al. (2000) probeerden het concept binnen het wetenschappelijk discours te houden door het te vergelijken met de standaarden voor intelligentietests. Er moet theorie gemaakt worden, inclusief psychometrisch verantwoorde instrumenten. EI/EQ moet de grens tussen de cognitieve en niet cognitieve kant van intelligente en adaptatie verhelderen. Het concept sluit socialisatie en adaptatie in en dat gaat zowel over cognitieve als niet cognitieve kenmerken. Gegeven haar korte bestaan is EI/EQ uitvoerig onderzocht. Het is immers populair. Voorbeelden:

- (1) Eerst moet er een instrument zijn dat EI/EQ meet. Schutte et al. (1998) gebruikten het concept van Mayer en Salovey om een vragenlijst te maken. Het is een zelfrapportage lijst bestaand uit 33

items. De interne consistentie en test-hertest waarden waren goed: respectievelijk $r = .88$ en $r = .78$ (interval: twee weken). EQ vormde volgens hen één factor en bevatte items als: 'Ik verwacht dat ik alles haal waar ik me voor inzet'; 'Ik doe de dingen die me blij maken'; 'Het is moeilijk voor me om te begrijpen waarom mensen voelen zoals ze voelen'. Het EQ correleerde bij 63 eerstejaars studenten psychologie $r = .32$ met studiesucces in het eerste jaar. Er was geen informatie die het mogelijk maakte om te zien of EQ *incremental* validiteit had ten opzichte van een conventionele intelligentietest.

(2) Sjöberg (2001) gebruikte het instrument tezamen met andere vragenlijsten (BF, MPI en de Myers Briggs-schalen). EQ was volgens hem een *secundaire factor* en laadde positief op empathie en zelfactualisatie en negatief op Machiavellisme. Dat is een schaal die een manipulatieve en cynische houding tegenover instellingen en mensen meet. Sjöberg benoemde de factor als: socio-emotioneel doeltreffend zijn: *self-efficacy in social and emotional domains* en betwijfelde de één factoroplossing van Schutte et al.

(3) In 2002 veranderde de inhoud van het concept, nu door onderzoek van Mayer et al. Het bevat (a) het omgaan met emoties door ze adequaat waar te nemen (b) emoties gemakkelijk kunnen bedenken (c) ze begrijpen en (d) ermee om kunnen gaan. Ze noemen dit het *four-branched ability model for EI* en de test heet: *Mayer-Salovey-Caruso-Emotional-Intelligence Test* (MSCEIT).

(4) In een studie van Føllesdahl en Hagtvet (2009) met behulp van Guttman's facetanalyse bleek dat drie van de vier *branches* een lage generaliseerbaarheidcoëfficiënt opleverde. Dit is een gegeneraliseerde betrouwbaarheidscoëfficiënt. De waarden zijn: r -generalizability = .37, .50 en .46. Ze onderzochten het EQ bij 111 Noorse managers. Dit resultaat zegt dat drie van de vier scores niet gegeneraliseerd kunnen worden. Hoewel de studie bij een selecte groep is uitgevoerd, bij managers, is de betrouwbaarheid of generaliseerbaarheid bij drie van de vier twijfelachtig.

(5) Mayer et al. (2008, p. 503) verdedigen het EQ als een construct dat verwijst naar een

'... capacity to carry out sophisticated information processing about emotions and emotionality-relevant stimuli and to use this information as a guide to thinking and social behavior'.

De bijdrage van EI/EQ aan het voorspellen van succes in een criteriumgedrag is verschoven naar een instrument voor waarnemen en hanteren van emoties.

(6) Gignac (2010) stelde voor om in EI *zeven* factoren te onderscheiden:

I Emotional 'self-awareness'

II Emotional expression

III Emotional awareness of others

IV Emotional reasoning

V Emotional self-management

VI Emotional management of others

VII Emotional self-control

Niet alle factoren kwamen even sterk uit de analyse te voorschijn, onder andere *Emotional reasoning*. Niettemin liet de confirmatieve factoranalyse aanvaardbare stress waarden zien. Dit is mogelijk want je kunt een gedragsverschijnsel uiteenleggen in aspecten. Gignac's studie biedt steun voor de beweringen van Mayer et al. (2008). De herziening of herinterpretatie van EI facetten krijgen steun in een onderzoek van Parker et al. (2011, zie hieronder). Gignac wordt tegengesproken door

Maul (2011). Hij gebruikte twee EI instrumenten, paste verschillende modellen en vond dat *één* factor de covariantie tussen de meeste van de twintig items van de twee instrumenten goed dekte. Waarneming, gebruik, begrip en managen van emoties waren substantieel gecorreleerd.

(7) Het EI construct verkrijgt een plaats als het iets toevoegt aan het voorspellen van cognitieve criteria. Daar was het in eerste instantie ook voor bedoeld. Lanciano en Curci (2014) probeerden met het nieuwe MSCEI construct het eerste doel van EI - toegevoegde validiteit, de meerwaarde - te realiseren dat wil zeggen naast IQ (en persoonlijkheid) variantie verklaren in prestaties. EI was significant gecorreleerd met prestaties in hun studie bij 89 eerstejaars psychologiestudenten (62% vrouwen: M = 19.21 jaar, SD = .97) van een Italiaanse rijksuniversiteit. Het gaat om de toegevoegde validiteit van EI aan prestaties. Ze gebruikten hiërarchische multiple regressie analyse (HRA) met als afhankelijke variabelen het aantal behaalde tentamens, GPA en voorbereidingstijd voor tentamens. De voor ons doel relevante HRA bevatte geslacht, leeftijd in jaren, vier MSCEI facetten, Raven en de EPI (Eysenck Personality Inventory) als voorspellers van drie prestatie-indices. Hoe meer variabelen meegenomen worden in de HRA hoe meer variantie doorgaans verklaard wordt in criteria. Het HRA model is significant en bindt tot 9% variantie mét MSCEI; zónder de EI facetten 3% voor het aantal examens en tot 10% met de Raven en het EI facet: Waarnemen van emoties zonder deze is 4% gebonden. Er is geen significante verandering bij de tentamenprestaties (GPA) wanneer EI facetten toegevoegd worden. Er was een significante relatie tussen aan studie bestede tijd en het EI facet Waarnemen van emoties bij mannen ($\beta = -.28$) maar niet bij vrouwen ($\beta = .07$). De auteurs concluderen dat er gedeeltelijk steun is voor de meerwaarde van een enkel EI facet bij het voorspellen van prestaties bij eerstejaars psychologie studenten.

(8) Parker et al. (2011) namen deel aan het 'een of meer factoren' debat rond EI/EQ. De gegevens van 2508 BA studenten stonden een psychometrische analyse van de structuur en van de interne betrouwbaarheid van de verkorte versie toe. De auteurs zeggen steun te hebben voor een meerdimensionele structuur. Elke dimensie was intern consistent en stabiel over tijd: test-hertest: 6 maanden tussentijd. Bij wijze van validering werd de correlatie berekend met de mate van ADHD: de rs varieerden tussen de -.20 en -.39. Correlaties met andere instrumenten voor hetzelfde construct waren bescheiden, matig en bijna goed: rs tussen de .16 en .56 (pp. 770-771). Men kan zich afvragen of de vuistregels van Cohen over r- en d-waarden passen bij correlaties tussen *gelijke of dezelfde* constructen gemeten met verschillende instrumenten? Past hier niet het idee van een parallel test? Het nut van EI en EQ wordt verdedigd door te wijzen op *incremental* validiteit voor het voorspellen van sociale gedragingen. Daarvoor is het gunstigste onderdeel van het instrument gebruikt: redeneren over emoties. Ze bevelen aan om over te stappen naar kennis van emoties: herkennen van emoties aan gelaatsuitdrukkingen, emotionele zelfregulatie: zie Mayer et al.'s aanbeveling hierboven en EI/EQ op zich te onderzoeken met de multi-method multitrait methode van Campbell en Fiske (1959). Hun resultaat is bescheiden: ongeveer dezelfde constructen correleren niet hoger dan $r = .56$.

(9) Als er een eenvoudig en kort instrument voor individuele verschillen is, wordt het al gauw gecorreleerd met andere constructen en concreet gedrag. DeBusk en Austin (2011) correleerden een facet van EI: 'emotieherkenning bij anderen' met sociale waarneming in een reactietijdtaak: inspectietijd. Semantisch is de connectie *face valid* maar EI was niet gerelateerd aan het herkennen van gezichtsuitdrukkingen. Mikolajcak et al. (2012) vonden dat emotioneel intelligente chefs naar verhouding meer management kwaliteiten lieten zien en minder gestreste werknemers onder hun hoede hadden.

Er is kortom geen eenduidige omschrijving met een dominant meetinstrument voor EI/EQ. Het aantal factoren varieert tussen de 1 en 7. De instrumentatie heeft een nieuwe impuls nodig, gegeven de betrouwbaarheden die Føllesdahl & Hagtvet (2009) melden. Gedurende de rit is de inhoud van het construct geherinterpreteerd of wordt er een veelbelovende dimensie uitgepikt. De correlaties met verwante constructen zijn bescheiden. Studie met het oorspronkelijk doel op het oog: meerwaarde maken ten opzichte van het IQ bij het voorspellen van prestaties is er nauwelijks. Het resultaat is bescheiden tot soms gemiddeld, zoals een metastudie van Murphy en Haal (2011) toont. Intelligentie is met een r van .19 gecorreleerd met *interpersonal decoding*: het kunnen waarnemen van emoties bij anderen. De groeiende populariteit van EI - onder meer onder managers - berust niet op de klinkende onderzoeksresultaten over de structuur en inhoud van het construct en de bijdrage aan het voorspellen van criteria. De hoge correlatie tussen de *General Personality Factor* en EI (r = rond de .70) van Van der Linden et al. (2010) laat zien dat EI iets aan de BF toevoegt. De toegevoegde waarde voor IQ en cognitie is beperkt.

Alternatieve constructen zijn uitgevonden voor individuele verschillen in intelligentie en geschiktheden omdat ze niet het hele verhaal vertellen over de variantie in prestaties. Alternatieven voor school- en werkprestaties ademen dezelfde geest maar hebben niet tot uitgewerkte concepten geleid. Metingen van deze prestaties worden beperkt genoemd omdat ze gericht zijn op *academic* succes: lezen, schrijven, rekenen. Schooltoetsen missen het leren van vaardigheden die je in het leven van alledag hebt: sociale en praktische intelligentie of handigheid.

Samenvatting en conclusie

Individuele verschillen in intelligentie (IQ), geschiktheid (PMA/DAT) en prestatie (GPA, Cito-toetsen) worden in de alternatieve benaderingen eenzijdig genoemd. Gardner voegt nieuwe factoren toe: muzikale, kinesthetische, en motoriek en andere methoden: observeren, ondervragen en producten waarderen. Greenspan en Driscoll stellen dat het IQ en geschiktheden niet echt raken aan waar het om gaat bij prestaties van werknemers en managers. Zij voegen competenties toe zoals fysieke kracht en affectieve en praktische vaardigheden. De leerpotentieel benadering ziet het statische en stabiele IQ als achterhaald en geeft de voorkeur aan dynamisch testen. Daarbij wordt de geteste geholpen met uitleg over testitems en doet hij metacognitieve oefeningen. Je kunt er continue oefening met feedback aan toevoegen zoals in 'Rekentuin'. Sternberg voegde aan de klassieke intelligentie praktische en inter- en intrapersonlijke vaardigheden toe. Emotionele intelligentie (EI/EQ) is oorspronkelijk bedoeld om meer variantie in criteria te verklaren in aanvulling op de IQ- en geschiktheidstests. Dit is verschoven naar een beschrijving van processen: emoties waarnemen en onderscheiden, informatieverwerken en emoties kunnen reguleren. Het construct is niet uitgekristalliseerd en bevat de al bekende *self-efficacy* en *self-actualization*. Het benadrukt aandacht dat te lezen is als een zich naar binnen keren, naar het zelf. Aandacht kan algemener gelezen worden als een houding van niet wegstijgen en niet weg- of goedpraten. Het lijkt op het Heideggeriaanse *Sein Lassen*: iets

laten zijn zoals het is of op het platonische/thomistische: iets toelaten terwijl het beperkt kan tonen wat het is.

Alternatieven voor school- en werkprestatie ademen dezelfde kritische geest als die voor intelligentie en geschiktheden. Ze zijn eenzijdig gericht op *academic* succes op school en in het werk. Er zijn echter geen nieuwe constructen uitgewerkt. De alternatieven zijn voor de diagnosticus geen vaste waarde geworden naast de bekende IQ, geschiktheids- en schoolprestatietests. Dit neemt niet weg dat ze populair zijn en ook buiten het circuit van de wetenschappelijke tijdschriften regelmatig genoemd worden als aanvulling op het IQ.

2. Ontwikkeling

Er zijn nauwelijks alternatieve voorstellen om de ontwikkeling van intelligentie, geschiktheid en prestatie te beschrijven. Het alternatief dat de organismische verklaring verlaat voor het verwerken van de informatie uit de gestructureerde buitenwereld (identificeerbare enkelvoudige en complexe stimuli) is een aanvaard paradigma: informatie wordt gecodeerd in stukjes uiteengelegd, gecategoriseerd, onthouden op korte en voor een deel op lange termijn en kan steeds weer opgehaald worden. Deze stappen verlopen doorgaans wat minder serieel, lineair dan wordt opgeschreven. Spontane processen zijn grillig, maar wellicht te lineair te maken door training. En iets opschrijven is lineair. De enige uitzondering is de experimentele letterkunde maar dat leest behalve Neerlandici die wel moeten, niemand.

Informatieverwerking De verklaring van stadiumovergangen door middel van analyse van stappen en training, bijvoorbeeld van pre- naar concreet operationeel is vrijwel van de onderzoeksagenda verdwenen. Informatietheoretisch bestaat ontwikkeling uit verandering in de processen van coderen, bijvoorbeeld het gebruik van steeds grotere eenheden en categorieën (*chunks*), decoderen en korte- en langetermijngeheugen. Ontwikkeling, uitgelegd als structurele verandering, is geen onderdeel van informatieverwerking. Er wordt verondersteld dat feitelijke kennis en de organisatie met behulp van strategieën toenemen met het ouder worden. De twee benaderingen zijn geen onderwerp meer van gesprek of debat.

De Canadese ontwikkelingspsycholoog Case heeft geprobeerd de informatieverwerking en stadiumgewijze ontwikkeling te *combineren*. Hij gebruikt Piagetiaanse inzichten over het hanteren van twee onafhankelijke dimensies om *memory space* te meten. Een kind kan bijvoorbeeld een verzameling objecten tegelijkertijd ordenen op kleur én op grootte. Volgens de Piagetiaanse theorie ontstaat dit vermogen rond het zesde jaar. Case (2000) neemt op grond daarvan aan dat de *memory space* toeneemt omdat het kind twee onafhankelijke dimensies kan onderscheiden en tegelijkertijd gebruiken bij het ordenen. Het meten van de *memory space* blijkt moeilijk en speelt geen rol in de diagnostiek. En volgens Kail (2000) groeit de geheugenruimte niet maar de snelheid van informatieverwerking. Hij sluit aan bij Thurstone en Thurstone (1941) waarin verondersteld wordt dat kinderen met het ouder worden sneller informatie verwerken. Snelheid betekent ook dat het geheugen

effectiever gebruikt wordt en dat heeft weer een gunstig effect op de kwaliteit van redeneren. Snelheid is zo geen aparte intelligentiefactor maar een basisproces dat het werk van het geheugen en redeneren vergemakkelijkt. Kail houdt dit standpunt overigens *niet* strikt vol, als hij in 2013 de integratiepoging van Demetriou en collega's van Ontwikkeling x Context omschrijft als de 'volledigste theorie over cognitie' die hij kent (hoofdstuk 4). Verlaat hij zijn eerdere interpretatie of is het een vrijblijvend Amerikaans compliment?

Om deugdelijkheid van de informatieverwerking als beschrijving en verklaring van ontwikkeling van cognitieve vaardigheden en intelligentie te onderzoeken, wordt het proces experimenteel in kleine eenheden uiteengehaald. De stappen hebben een logische structuur, bijvoorbeeld een opbouw van eenvoudig naar complex. Hier doen zich dezelfde vragen voor als bij stappen in een ontwikkelingsproces: bestaan ze, zijn ze vast te stellen en zijn ze goed te onderscheiden? Doorlopen proefpersonen de onderdelen, stappen, als verondersteld? Aan elk stappenplan liggen rationalistische en atomistische veronderstellingen ten grondslag. In de werkelijkheid is het gedrag van proefpersonen misschien niet zo keurig ingedeeld, verloopt het niet stapsgewijs en kan de tijd van verblijf in een stap wisselen. Soms verkeren ze in twee of zelfs drie stappen tegelijk. De processen zijn in experimenten beschreven en er is geen diagnostisch instrument voor geconstrueerd met als gevolg dat de diagnosticus een informatieverwerkingsaanpak van ontwikkeling van intelligentie zelden benut.

Alternatieve theoretici zijn gevoelig voor het feit dat ontwikkeling van schoolprestaties en het testen van effecten van curricula verbonden zijn met politieke en maatschappelijke *beliefs* en belangen. Sommige onderwijskundigen, bijvoorbeeld Wiggins (1993) vestigen er de aandacht op dat de ontwikkeling van prestaties en de meting daarvan door SATS eenzijdig belicht zijn. Sociale vaardigheden worden ondergewaardeerd en de tests zijn in het voordeel van midden en hogere sociale klassen.

Samenvatting en conclusie

Alternatieve theorievorming voor de ontwikkeling van intelligentie, geschiktheid en prestatie is beperkt. De studie naar het versnellen van ontwikkelingstransities door leren is nu geschiedenis. De benadering van de informatieverwerking is een erkend onderdeel van expliciete theorievorming. Intelligentie en cognitie ontwikkelen door verandering in selecteren, coderen, bijeenvoegen van informatie uit de buitenwereld. De aanpak bestaat uit een logisch geordend complex stappenplan en dat heeft niet tot diagnostische instrumenten geleid. Het werk van Das en Naglieri (1994, zie verderop) is een uitzondering en is geen veel gebruikte aanpak bij diagnostici. Het vaststellen van het bestaan van de stappen en het in volgorde verwerven ervan, plaatsen de onderzoeker voor dezelfde problemen als het toetsen van Piagetiaanse stadia. Er is om de zoveel tijd terugkerende kritiek op het eenzijdig vaststellen van ontwikkeling van cognitieve, schoolse en abstracte prestaties en het op het formuleren van eenzijdige cognitieve onderwijsdoelstellingen, vooral de nadruk op herkennen en niet op toepassen. Dit heeft evenmin tot nieuwe theorievorming en instrumentatie geleid die de diagnosticus kan gebruiken.

3. Sociale context

Een alternatieve opvatting over de werking van de sociale context is te vinden bij sociaal constructivisten. Zij sluiten aan bij het symbolisch interactionisme, dat vooral in de sociologie is uitgewerkt. Ze kunnen voor een deel tot de interpretatieve en voor een deel tot de kritische epistemologische stroming gerekend worden. Er wordt gesteld dat prestaties niet alleen een uitdrukking zijn van een met verstand begiftigd individu maar ook en vooral van de invloed van de sociale en culturele context. Deze schrijft immers voor welke concepten en doelen een individu moet beheersen en nastreven. Het is een belang van elke samenleving om personen te differentiëren naar kenmerken, zoals intelligentie, geschiktheid en prestatie. De samenleving is gericht op 'de werken van het verstand, de ratio' en op rationele productie van goederen en diensten. Sommige sociaal constructivisten zien alle sociale groepen en alle theorie als gelijkwaardig en een gevolg daarvan is dat er geen hiërarchie tussen impliciete en expliciete theorie is. Anderen gaan ervan uit dat sommige groepen, op basis van wat dan ook, bijvoorbeeld afkomst, beogaafdheid, rijkdom, enzovoort meer recht op de waarheid en goederen en diensten kunnen doen gelden dan anderen. Dit verschijnsel is al zo oud als de geschiedenis. Bij de Grieken waren de filosofen de vertolkers van de waarheid. In religies waren het de priesters. In de moderne consumptiemaatschappij zijn het bankiers, economen, ondernemers, reclamemakers, politici en wetenschappers. Waarheid is daarbij niet het ontbreken van leugen maar economisch gewin. Dat laatste vereist soms regelrechte leugens zoals in de reclame, bankproducten en beloften op succes. Een recent voorbeeld is de bijdrage van psychologen aan efficiënte ondervragingstechnieken in Guantánamo Bay (Abu Ghraib gevangenis), waarop de ethische commissie van de Amerikaanse psychologenvereniging nogal laat reageerde (zie de documentaire *Doctors of the Dark Side* met Stephen Behnke, de toenmalige Ethiek voorzitter van de APA: www.journeyman.tv).

Volgens sociaal constructivisten verwijzen psychologische concepten naar belangen van sociale groepen en instituties. Intelligentie helpt om onderscheid te maken tussen personen en hen een plaats te geven in een groep. De sociale representatie van intelligentie heeft een functie en schept een symbolische realiteit voor een sociale groep, die minstens zo krachtig is als de werkelijkheid of het IQ zelf. Het is geen mensenrecht om niet bedrogen, in het ootje genomen, op het verkeerde been, enzovoort gezet te worden. Iemand krijgt een plaats in de gemeenschap door die symbolische werkelijkheid met anderen te delen. Sociaal constructivisten voeren meestal kwalitatieve studies uit, al hoeft dat niet. Voorbeelden:

(1) Carugati (1990) vroeg zich af hoe de sociale representatie van individuele verschillen in intelligentie tot stand komt in een groep. Sommige leden van een groep zijn bekend met de verschillen in intelligentie want zij zijn betrokken bij het vormen van de representatie. Dat zijn onder meer psychologen en pedagogen, onderwijsambtenaren, leraren en ouders. De laatsten moeten zich een beeld vormen van de mate waarin intelligentie van leerlingen en kinderen te beïnvloeden is.

(2) Mugny en Carugati (1989) hebben volwassenen gevraagd te omschrijven wat algemene en specifieke intelligentie is en wat ze moeten ze doen om beter te presteren bij bijvoorbeeld

wiskundeopgaven, talen en tekenen. En verder, hoe gedraagt een prototypisch intelligent kind zich en hoe kan men intelligentie vaststellen? De representatie van individuele verschillen in intelligentie in deze sociale groep toont volgens de auteurs 'rijke en genuanceerde dimensies':

- a. Volwassenen vragen zich af waar individuele verschillen in intelligentie vandaan komen en zijn nieuwsgierig of de wetenschap deze vraag kan beantwoorden.
- b. Zij vragen zich af of intelligentie ook betekent dat je de sociale regels van de samenleving leert kennen.
- c. Intelligentie wordt door de informanten gezien als een product van hogere processen waarvoor logica, abstraheren, wiskunde en de computer de symbolen zijn. Als je met de computer uit de voeten kunt, ben je volgens de maatstaven van ouders in 1989 intelligent.
- d. Intelligentie houdt volgens de volwassenen in dat kinderen zich aanpassen aan de fysische en sociale omgeving, vooral aan de schoolomgeving. Als oorzaken of redenen voor verschillen wijzen zij op de sociaal economische status van het gezin, de overerfbaarheid van intelligentie en de invloed van leraren. Ten slotte vragen ze zich af of je door onderwijs de prestaties in wiskunde kunt beïnvloeden en of intelligentie een gave ('gift') is.

(3) Sociaal constructivisten noemen twee mechanismen die werkzaam zijn bij het tot stand komen van de sociale representatie van intelligentie: (1) bekendheid met intelligente prestaties en (2) processen om sociale identiteit in stand te houden. Ze proberen deze mechanismen terug te vinden in interviews met ouders over de intelligentie van hun kinderen. Gergen et al. (1989, p. 145) hebben vastgesteld dat naarmate mensen minder vertrouwd zijn met individuele verschillen, zij eerder geneigd zijn om intelligentie als een *natural gift* op te vatten. Verschillen in intelligentie zijn volgens hen dan ook niet of weinig te beïnvloeden. Bovendien vermelden de auteurs dat buitenshuis werkende moeders vaker dan thuiswerkende moeders zeiden dat intelligentie zich ontwikkelt volgens een biologisch rijpingsproces. Ze legden niet zoveel nadruk op de bijdrage van de moeder voor de taal-, denk- en sociale ontwikkeling van de baby's. De auteurs interpreteren dit als een poging haar identiteit als 'goede moeder' te behouden.

Andere sociale groepen representeren hun intelligentie en prestaties op een andere wijze om die in overeenstemming te brengen met een gewaardeerde identiteit. Dat verschilt, bijvoorbeeld tussen leden van de Mensa - een groep die zich hoogbegaafd noemt en waarvoor door het behalen van een ingangstoets een lidmaatschap kan worden verworven - en een groep Haagse jongeren die op een Grieks eiland feest viert van wie een jongere vanwege zijn tatoeage 'Sterretje' genoemd opmerkte: '... op school was ik vooral goed in pauzes'. De sociaal constructivisten leveren weerwoord en tegenspel aan maatschappelijk aanvaarde intelligentie als individueel verschil door de belangen van en sociale praktijken van gevestigde groepen aan het daglicht te brengen. Volgens hen is iedere waarheid gecontextualiseerd.

Sociaal constructivisten maken hun opvattingen over intelligentie als sociaal, cultureel en symbolisch verschijnsel aannemelijk aan de hand van kwalitatief onderzoek. Ze hebben daarbij oog voor wie belang heeft bij een specifieke symbolische constructie.

Samenvatting en conclusie

Intelligentie, geschiktheid en prestatie zijn volgens sociaal constructivisten geen trek of een adaptief informatieverwerkend mechanisme. Het zijn symbolische constructies die het belang van een sociale groep dient en zijn identiteit bevestigt. Groepen, bijvoorbeeld

ouders, construeren intelligentie als een proces dat onder meer blijkt uit begrip van wiskunde en handig zijn met computers in dit onderzoek van de jaren 80 en 90 van de 20^{ste} eeuw. Als ze weinig kennis hebben over de constructie van intelligentie door experts, construeren ze het als een gave. Constructivisten wijzen er bijvoorbeeld op dat intelligentie, geschiktheid en prestatie verschillend geconstrueerd worden in verschillende groepen en de identiteit van de groep bevestigen.

4. Reflectie en evaluatie

Alternatieve theorievorming wordt geïnspireerd door de intuïtie dat bepaalde vaardigheden, capaciteiten en prestatiemetingen in IQ-, geschiktheids- en prestatietests ontbreken. De nadruk op abstractie, schoolse en boekenkennis wordt gekritiseerd. Sociale vaardigheden, competenties, het vermogen te leren, natuurlijke intelligentie, snelheid van informatieverwerking, codeer- en geheugenprocessen worden gemist. Dit leidde tot het benadrukken van natuurlijke intelligentie (niet die kunstmatige testitems!), emotionele intelligentie (niet die koude, rationalistische en abstracte kennis!), succesvolle intelligentie (een hoog IQ is geen garantie voor succes in de samenleving!), de potentie van intelligentie (niet een vastliggend, maar een voortdurend voor verbetering vatbaar vermogen!) en competenties om goed te functioneren in een beroep (geen nerds met abstracte boekenkennis, maar emotioneel en communicatief vaardige medewerkers graag). De concepten hebben *face* validiteit maar het bewijs dat ze *incremental* validiteit tonen ten opzichte van de IQ en prestatietests is nauwelijks geleverd. Deze alternatieven dragen niet bij aan theorievorming over intelligentie, leveren geen nieuwe efficiënte methoden om te diagnosticeren en bieden weinig om (on)gewenst gedrag van de cliënt beter te voorspellen, als ze toegevoegd worden aan IQ tests.

De informatieverwerkingsbenadering van de ontwikkeling van intelligentie en cognitie is onderwerp van experimenteel onderzoek. Het is geen kritisch alternatief maar een aanvaard paradigma. De analyses van de processen en hun groei door de tijd heen zijn weliswaar interessant maar hebben tot op heden niet geleid tot nieuwe instrumenten of werkwijzen voor de diagnostiek van de cliënt. Ze kunnen bijdragen om intelligentie, geschiktheid en prestatie gedetailleerd te beschrijven.

Theorie en waarheid zijn soms verbonden met belangen van een elite, een institutie en de bourgeoisie. De sociaal constructivisten laten zien dat intelligentie en prestatie geen neutrale concepten zijn. Sommigen gaan zover dat ze de (sociale) wetenschap als uitdrukking van belangen van een groep, vooral de elite en soms ook de middenklasse, beschouwen. Ze vatten iedere theorie als gelijkwaardig op. Dit kan leiden tot een ver doorgevoerd en uiteindelijk onvruchtbaar relativisme. Constructivisten vormen soms een militante groep die zich verzet tegen de bourgeois mentaliteit van de gevestigde orde die tests als middel gebruiken om hun middelmatig burgerleven ongestoord voort te zetten. Ze schrikken die burger graag af. De symbolische constructie van intelligentie en cognitie vervult een functie voor hun belangen en identiteit.

De diagnosticus is vertrouwd met deze opvattingen. Hij zal die overigens niet vaak bij cliënten aantreffen. Sociaal constructivisten hebben niets met diagnostiek en ze verwachten er nadelen van. Ze zullen niet bij hem te rade gaan. Als professional is hij vertrouwd met de functie en werking van psychologische concepten voor cliënten. Sommige sociaal constructivisten zien alle systematische concepten als uitdrukking van belangen van een groep en elke groep is gelijkwaardig in de toegang tot de waarheid: alle theorieën even veel waard. Deze militante opvatting komt van tijd tot tijd op als verzet tegen de bourgeois- en elitecultuur. Het constructivisme biedt geen nieuwe concepten of werkwijzen die een diagnosticus kan benutten voor analyse en een rijke beschrijving van het gedrag van de cliënt. Het laat wel zien dat deze vorm van sociologie van kennis kan bijdragen tot het relativeren van psychologische constructen maar ze kan niet de vervanger zijn voor kennisverwerving zelf (Williams, 2002). Het constructivisme is een voorbeeld van een welwillend (Sloterdijk, 2009) uitgewerkt alternatief dat aanzet tot reflectie op de bestaande opvattingen over sociale functies van intelligentieconcepten. Daarnaast bevat het enige afschrikking. In een kwaadaardig uitgewerkte versie zou welke inhoud dan ook irrelevant zijn. Dit zet de deur open voor bewuste misleiding en dat doet zich voor in sommige doctrines die geformuleerd zijn door onaantastbare leiders.

V Integratie van expliciete theorieën van intelligentie en prestaties

Er zijn drie bronnen en drie oriëntaties onderscheiden die de theorieën over intelligentie/cognitie en context dekken. De behoefte aan integratie is bij intelligentie en cognitie minder aanwezig dan bij theorievorming over typerend gedrag. Het cognitieve domein wordt in het rationalistische Westen gewoonlijk beschouwd als op, door en in zichzelf voldoende: helder, onderscheiden en leidend tot zekerheid. Het is een zelfstandig en afgebakend gebied.

Is het zinvol om de drie oriëntaties te integreren of in een hiërarchisch verband te plaatsen? Of domineert één oriëntatie? Waarom zoeken naar eenheid? Is dat een gevolg van de westerse monotheoretische - één theorie over alles - voorkeur? Wat is de uitslag van de drie twee bij twee combinaties van oriëntaties? Wat betekenen de combinaties voor de diagnosticus?

1. Eén theorie?

Intelligentie en cognitie verwijzen naar kennis over de wereld, dat wil zeggen over objecten, gebeurtenissen, verschijnselen, mensen en de samenleving. Mensen streven naar een gemeenschappelijk systeem of stelsel van richtlijnen voor hun kennen. Dit blijkt bijvoorbeeld uit de impliciete overtuiging dat intelligentie en cognitie *niet* kunnen veroorzaken dat we in een volstrekt onkenbare, chaotische wereld leven. En mensen willen 'alles' weten, begrijpen en doorgronden. Een idealistische notie van waarheid blijkt uit het Hegeliaanse adagium: *Das Wahre ist das Ganze*. Een soepele versie van de evolutietheorie staat een monistische interpretatie toe over 'al wat is'. Er is een voortdurende differentiatie van de materie en tussen en binnen species. Toevallig heeft dat geleid tot de homo sapiens, een wezen met bewustzijn en complex een taal- en communicatiesysteem. Als het bewustzijn alles is en het zichzelf kan doorgronden is er geen ondoordringbaarheid of mysterie. Kant en Sartre zijn het daar niet mee eens. Er is immers altijd een perspectief. Kennisvermeerdering verschuift de horizon. Daar komt zolang de mens bestaat en zal bestaan geen einde aan. Dat is de strekking van Merleau-Ponty's ontkenning van de *régard survolante*.

Natuur- en sterrenkundigen zijn bescheiden als het gaat over één omvattende theorie. Toen de bewonderende journalist Hagens tegen Robbert Dijkgraaf zei hoeveel zaken hij doorgrondde als wis-en natuurkundige en sterrenkundige, ontkende hij dat meteen: van 95% weet ik niets af (Buitenhof TV, 29-12-2013). Tegelijkertijd zijn natuurkundigen zelfbewust want ze kunnen bogen op technologische prestaties. De Newtoniaanse natuurwetten zijn op te vatten als boekhouden van een deel van het universum zegt Verlinde, een collega van Dijkgraaf. Door zich bescheiden op te stellen openen ze meteen een perspectief. Er wordt overigens regelmatig kritiek geleverd op de 'Theorie over al wat is'. Ze zou gaan van de astronomie tot de deeltjesfysica en ondertussen nemen ze de biologie - inclusief de homo sapiens - mee. In een boek *Wetenschappelijk Onkruid* (redacteur John Brockman, 2105) wordt 'De theorie van Alles' regelmatig aangevochten door onder meer West en Steinhardt. Dijkgraaf zegt meteen dat we meer te weten kunnen komen over het universum en de mens. Voor de mens zag Dijkgraaf wel iets in *data mining*, maar misschien is dat wel het ultieme boekhouden van de psychologie, heel respectabel, maar inzicht vloeit er niet noodzakelijk uit voort. Het levert een wirwar aan sporen op. Om dat te organiseren is een separate theoretische, wiskundige, psychologische, enzovoort inspanning nodig.

Als we dit transponeren naar kennis van het menselijk gedrag is het daar ook zo dat er 95% niet verklaard is, of ligt het anders? En is die kennis netjes vastgelegd in d- en r-waarden of in gedragswetten, analoog aan Newtons wetten of in probabilistische, statistische wetten? Is het ook boekhouden van ons gedrag? Ik heb psychometrie 'boekhouden van de ziel' genoemd (ter Laak, 1995). Kunnen de sociale wetenschappen ook zo'n perspectief bieden? Is er nog veel te winnen?

De overtuiging van leken en wetenschappers is dat intelligentie en cognitie niet volkomen blind zijn en dat we daarom niet in een chaotische, contingente, onbegrijpelijke wereld

leven. De slogan zou kunnen zijn: 'intelligentie is één'; ze moet zonder interne tegenspraken leiden tot een voor een deel begrijpelijke, verklaarbare wereld of 'intelligentie is niet'. Psychologen claimen iets te weten van intelligentie.

Claims van psychometrische, cognitieve ontwikkelings- en contexttheorieën
Psychometrische intelligentie gaat over individuele verschillen. Spearman's G bevatte de boodschap van een unificerende kracht die in staat stelt de wereld en de mensen te begrijpen. Mensen beschikken over *mental power*. Thurstone claimde het bestaan van verschillende geschiktheden om de complexiteit van de fysische wereld recht te doen, te 'spiegelen'. Intelligentie is niet één maar multipel want de wereld van objecten en verschijnselen doet zich aan ons voor in verschillende kwaliteiten: verbaal, numeriek, ruimtelijk, enzovoort. Guilford organiseerde het kenbare en het kennen in een complexe structuur: 120 of zelfs 150 intelligenties. Hij is er van overtuigd dat vele intelligenties nodig zijn om te overleven en succesvol te zijn in de fysische en sociale wereld. Intuïtief accepteren we dat verschillende intelligenties en vaardigheden nodig zijn om te overleven en dat deze verbonden zijn met specifieke kenmerken van de wereld. De wereld 'is' verbaal, ruimtelijk, numeriek, of staat ons toe hem zo te definiëren. Men kan de wereld ook als jungle, een chaos omschrijven, waarvoor alle intelligenties van Guilford zelfs niet genoeg zijn. We kunnen uiteindelijk immers maar een beperkt deel begrijpen en verklaren.

Een antwoord op veelheid is vorming van een hiërarchie. Dat is een westerse stap die we gemakkelijk maken. Alles krijgt zijn plaats (Vernon en Carroll in hoofdstuk 3). Psychometrische intelligentie plaatst het meten van individuele verschillen centraal. Dat is de koninklijke weg om kennis te vergaren over menselijke intelligentie. Zij vormt echter geen eenheid. De vraag naar het ene of het vele wordt overgelaten aan resultaten van factoranalyses op tests, items, taken en vragen en aan moderne psychometrische modellen. In de moderne testtheorie is de eenheid, de bouwsteen de latente trek en de IRF. Wellicht in navolging van het succes van de natuurkunde met atomen, neutronen, neutrino's, enzovoort wilde men de 'eenheden van de geest' beschrijven: latente trekken, IRFs, ECTs, het EEG en EVPs. De rol die deze spelen bij het oplossen van problemen, het beantwoorden van items en tests is onvoldoende uitgewerkt en misschien lukt dat nooit. En ze voorspellen criteria niet beter dan IQ tests.

Intelligentie en hersenen Een recente uitwerking van de relatie hersenen en intelligentie is van Kievit et al. (2012) en Kievit (2014). Hij plaatst het vraagstuk van de relatie in het kader waar het thuishoort: de eenheid van wetenschappen zoals beschreven door Oppenheimer & Putnam en Fodor. Hij schuwt het *mind-body* vraagstuk en zijn wijsgerige implicaties niet. Er wordt niet volstaan met het berekenen van correlaties tussen hersen- en gedragskenmerken die al in grote hoeveelheid beschikbaar en samengevat zijn (zie bijvoorbeeld het overzicht van Deary et al., 2010).

Ze zoeken naar een model of werkwijze die een gemeenschappelijke structuur voor beide blootlegt, creëert of uitvindt. Ze passen verschillende latente structurele

vergelijkingsmodellen op rechtstreeks gemeten biologische/neurologische parameters en G-verzadigde intelligentietests. Tachtig subjecten (29 mannen, 51 vrouwen; tussen 18 en 29 jaar oud namen deel: $M = 21.1$ jaar, $SD 2.55$ met een gemiddeld WAIS-III- IQ dat 1.0 SD hoger is dan gemiddeld: vooral studenten dus. Er is een sterke algemene WAIS factor maar doorgaans worden vier domeinindices vermeld: Werkgeheugen, Verbaal Begrip, Perceptuele Organisatie en Informatieverwerkingssnelheid. Daarnaast is bij elke deelnemer in acht gebieden, loci van de hersenen de grijze massa integriteit en dichtheid bepaald. Dit zijn rechtstreekse metingen aan het brein en die zijn niet mogelijk zonder complexe apparatuur en experts die de parameters vaststellen en hun betrouwbaarheid schatten (zie Kievit, 2014, pp. 80-81).

De auteur stelt vier modellen voor en vindt - als te verwachten - de meeste steun voor het meest complexe. Het eerste model waarin alle hersen-indicatoren de G reflecteren, viel af. In deze steekproef kunnen de neurologische metingen niet als metingen van G beschouwd worden. In het tweede model viel de veronderstelling af dat de neurologische metingen één factor (de neuro G) vormden. Er is in de steekproef geen steun voor een eendimensionele 'hersens-fitness factor' (Hebb, zie hiervoor, dacht van wel met als basis een goed metabolisme). In het derde model werden gecombineerde exploratieve en confirmatieve factor analyses gebruikt met twee latente neurologische factoren die correleren met G. De auteurs vinden dit model intuïtief plausibel maar krijgen onvoldoende steun uit de data. Ze kiezen voor een vierde model. De acht neurologische metingen voorspellen samen de G, terwijl ze op zich geen unidimensionele schaal vormen. De acht voorspellen de G tussen de -0.051 en +0.295 terwijl de G met een r van .442 (Perceptual Speed) en met een r van .675 (Verbal Comprehension) de vier indicatoren voorspelt.

Het ligt voor de hand dat het meest ingewikkelde model het dichtst bij de gegevens komt, als je je (voor)oordeel aanhoudt dat alles uiteindelijk ingewikkelder is dan je denkt of wil. Het neemt niet weg dat deze wijze van verbinden van moderne psychometrie (latente structurele modellen) en rechtstreekse bepaling van neurologische parameters een stap verder gaat dan de correlatieve hersen-gedrag onderzoeken.

Veel verklaringen: Bij de verklaring van de intelligentieverschillen zijn veel factoren betrokken. Daar is geen integratie of eenheid te verwachten, eerder een opsomming en soms strijd over wie gelijk heeft of erger nog, het voor het zeggen heeft. Van tijd tot tijd wordt nadruk gelegd op *nature* en dan weer op *nurture*. Doorgaans wordt de variantie gedeeld. De *nurture* vertegenwoordigers slagen er overigens minder goed in om de variantiebronnen te beschrijven. Ze sprokkelen ongeveer 20% bijeen. De *nature* onderzoekers daarentegen ongeveer 50%. De hypothesen over de oorzaken variëren van snelheid van zenuwgeleiding, een goed metabolisme in het centraal zenuwstelsel tot een psychometrische latente trek.

In de theorievorming over mechanismen is geen integratie te bespeuren. Het enig zekere is dat intelligentiescores - bescheiden tot gemiddeld en met grote variatie - met zo ongeveer alle kenmerken van biologische, neurologische en fysiologische aard samenhangen. Je

vraagt je soms af of er niet af en toe ook *niet* significante correlaties gevonden zijn. Worden die niet gepubliceerd omdat geen redacteur dat zou publiceren? Er moet immers iets uitkomen. Dit geldt ook voor effecten van kenmerken van de sociale context op intelligentie en cognitie.

Geschiktheden: Bij geschiktheden wordt gauw aan een aantal en niet aan één G gedacht. Er zijn verschillende opleidingen, functies en beroepen, waarvoor uiteenlopende vaardigheden nodig zijn. Prestaties zijn direct aan training en instructie gebonden: men kan zwak zijn in de stof van curriculum A, maar niet in de onderwerpen van curriculum B.

Cognitieve ontwikkelingspsychologie bevat de claim de ontwikkeling van het kennen compleet te beschrijven en te verklaren. Zij verloopt van reflexen tot formele operaties. Dit is het gemeenschappelijk pad dat ieder mens als epistemisch subject gaat. Piaget is niet geïnteresseerd in individuele verschillen en verwacht niet veel van het manipuleren van de context om cognitie te bevorderen of te versnellen. De feiten hebben zijn *nature* standpunt verzacht. Er is bijvoorbeeld vertraging in ongeletterde samenlevingen, de Franse kolonie Mauritanië van de jaren 60, bijvoorbeeld. De enige sociale context waar Piaget iets van verwachtte was de peergroep. Deze kan een cognitief conflict veroorzaken dat leidt tot accommodatie naar een hoger stadium. Dit lijkt behalve theoretisch ook ideologisch gemotiveerd. Hij was getuige van twee wereldoorlogen waarin volwassenen rigide waren en volledige gehoorzaamheid eisten of toonden met desastreuze gevolgen van onder meer genocide. Judith Harris (1995) legde eveneens nadruk op leeftijdgenoten als oorzaak van socialisatie en leren.

Sommige auteurs voegden aan de stadia de *postformele periode* toe en betwisten daarmee het Piagetiaanse eindstadium van de cognitieve ontwikkeling. Piaget zou er niet van onder de indruk geweest zijn. Het ging hem om de formele structuur, niet om inhoudelijke uitwerkingen met dialectiek en integratie op een hoger niveau. Er is bij Piaget geen verdeeldheid over de verklarende mechanismen van de ontwikkeling. Assimilatie en accommodatie leidden tot een stadiumgewijs verloop van biologische reflexen naar logische formele operaties. Er is wel slijtage van de theorie doordat niet alle veronderstellingen geldig bleken. Formele operaties zijn bijvoorbeeld gebonden aan specifieke gebieden. In de confrontatie met andere opvattingen over mechanismen van de ontwikkeling zijn nieuwe interpretaties voor dezelfde verschijnselen toegevoegd. De overgang van pre- naar concreet operationeel is zichtbaar gemaakt aan de hand van conservatietaken. Bij het vermoeden dat het een waarnemingsverschil is, kun je oogbewegingen van pre- en concreet operationele kinderen bestuderen. Kijken de concreet operationele proefpersonen bij conservatie van water inderdaad vaker dan de pre-operationele- naar hoogte én breedte van de glazen, en wisselen ze daarbij steeds af? Als je denkt dat het een leerproces is, ligt training voor de hand. Bevorderen leren en trainen om rekening houden met twee onafhankelijke kenmerken bijvoorbeeld het conserveren van massa, hoeveelheid? En generaliseert dat

naar andere taken? Er zijn studies die deze veronderstellingen in beperkte mate ondersteunen.

De claim van Piaget: een volledige cognitieve ontwikkelingstheorie is niet waargemaakt, maar de theorie leidt tot minder verdeeldheid en hiërarchisering dan bij individuele verschillen theorievorming rond psychometrische intelligentie.

Context theorieën is in de VS en het VK van meet af aan verbonden met het empirisme: 'Zijn is waargenomen worden'. Gestaltpsychologie is een andere, Duitse uitwerking van de waarnemer-omgeving relatie. Deze is minder prominent in de huidige psychologie. Het behaviorisme in zijn klassieke uitwerking (Watson) en in de versies van Rotter en Bandura verklaren gedrag vanuit stimuli. In deze abstracte vorm is er de claim van eenheid: een breed, eenvoudig, *parsimonious* en dwingend model $S \rightarrow R$. Er zijn bezwaren geuit tegen het naïef empirisme als verklaring voor intelligentie en cognitie. Mensen zijn geen onbeschreven blad. En de eenheid zou ten koste van het bereik gaan. Intelligentieverschillen en cognitieve ontwikkeling zijn niet volledig te verklaren met *behavioral repertoires* (Staats) en ook niet met de reinforcementgeschiedenis (Rotter) en door imitatie en leren van een model (Bandura).

Informatieverwerkingstheorie Deze legt de ontwikkeling neer in de centrale systemen van geheugen en processen van verwerking. Ontwikkeling is het sneller verwerken en het meer onderdelen van de stappen uit een proces foutloos afleggen. De benadering was als contexttheorie een kandidaat voor een omvattende verklaring van intelligentie. Algemene, abstracte kenner-omgeving uitwisselingsprocessen als encoderen, decoderen, inferenties trekken en korte en lange termijn geheugen hebben groter bereik dan hun behavioristische $S \rightarrow R$ voorgangers. Kunnen deze processen zo samenwerken dat er coherente, gedeelde kennis over objecten, verschijnselen en mensen tot stand komt? Zijn de processen hetzelfde voor verschillende domeinen, bijvoorbeeld het verwerken van plaatsen, tegenover visuele objecten, procedurele versus declaratieve kennis, taal, rekenen en ruimte? Bestaat er een empirisch getoonde vaste volgorde in de processen en producten bij het verwerven kennis? Zijn er individuele verschillen in de informatieverwerking van objecten en personen? Hebben we als gevolg daarvan verschillende cognities over dezelfde objecten en verschijnselen? Dit is waarschijnlijk het geval, gegeven de moeizame communicatie die nodig is om een gemeenschappelijke wereld te construeren. Het kost al veel tijd om een duidelijk onderwerp te bepalen voor een discussie of debat. Onderwijs, wetenschap, media, politiek en reclame moeten alle samenwerken anders ontstaat er geen gemeenschappelijke wereld. Dat ligt niet voor de hand en bijgevolg is er feitelijk een gefragmenteerde wereld. Anderson et al. (2004) hebben geprobeerd om een geïntegreerd informatieverwerkend model voor kennisverwerving en cognitie te maken. Het moet verklaren waarom we een coherente waarneming en cognitie hebben van een willekeurig verschijnsel. Het heet: *Adaptive Control of Thought-Rational* (ACT-R). Ze veronderstellen verschillende betrekkelijk zelfstandige modules die geïntegreerd worden om coherente cognities te produceren. Ze

noemen bijvoorbeeld motorische, doelgerichte en declaratieve geheugenmodules. Ze zijn verbonden met verschillende gebieden in de hersenen. De modules plaatsen klonters (*chunks*) van informatie in buffers die ontdekt en opgeroepen kunnen worden door een productiesysteem, dat zo een antwoord vindt in de informatiepatronen die de buffers 'bevolken'. Het cognitiesysteem wordt gaande gehouden door selectie en interne operatieregels en levert uiteindelijk een zinvolle cognitie op van een object, verschijnsel of persoon. Het model is een architectuur en architecturen maken indruk. Ze ademen rationaliteit want alles is er en op de juiste plaats. Maar er zijn altijd andere architecturen mogelijk die ook tot dezelfde coherente cognities kunnen leiden in de veronderstelling van de multiple realiseerbaarheid van elke gedraging. Of cognities coherent zijn is een vraagstuk op zich. Criteria voor zo'n bewering kunnen verschillen. En, is coherentie, een *Gestalt*, een gevolg van een dominante Gestaltwet of een coherentie die *achteraf* gemaakt, geconstrueerd wordt? Is coherentie een onontkoombaar een gevolg van oorzaak-gevolg patronen? In concreet onderzoek blijkt het moeilijk de informatieverwerkende processen te isoleren en de specifieke modules aan het werk te zien en vervolgens te diagnosticeren. Daarna komt de taak nog om na te gaan of ze geïntegreerd tot een cognitie van een object of persoon leiden.

Effecten van de context op intelligentie en cognitie kunnen bepaald worden, als die goed in kaart gebracht wordt. Naast de controle in het laboratorium zijn er natuurlijk contexten van gezin, peers, school, buurt en (sub)cultuur. In gedragsgenetische modellen neemt de omgeving een beperkt deel van de variantie voor zijn rekening. De variantie wordt verdeeld over twee bronnen: binnen en tussen gezinnen ($e^2 + c^2$), maar er zijn ook auteurs die het breder zien, bijvoorbeeld klimaat bij Kanazawa (2004, 2005, 2010). Er wordt doorgaans een aanzienlijke invloed verondersteld van de gemanipuleerde sociale context op individuele verschillen in intelligentie, geschiktheid en prestatie. Men wil de laatste immers verbeteren en men ziet daar steeds weer brood in. De studies zijn gefragmenteerd en volgen een S→R schema.

Eén intelligentie en cognitie theorie Pogingen om theorie over intelligentie en cognitie te integreren, zoals Mayer en McAdams dat gedaan hebben met typerende gedrag ontbreken bij intelligentie en cognitie. Het domein van kennen en intelligentie is complex en wordt omringd door moeilijk op één lijn te brengen mechanismen en invloeden. De claim van één geïntegreerde theorie voor intelligent en cognitief gedrag op basis psychometrische individuele verschillen theorievorming, cognitieve ontwikkelingspsychologie, het behaviorisme, sociaal leren theorie of op grond van een architectuur van informatieverwerking wordt niet waargemaakt.

Samenvatting en conclusie

Individuele verschillen in intelligentie leveren geen geïntegreerd beeld op. Is er één intelligentie of zijn er vele? Het voorlopig antwoord of uitweg is: een hiërarchie van

abstract, algemeen naar specifiek en bijzonder. Geschiktheden zijn van meet af aan gedifferentieerd naar opleidingen, functies en beroepen. Prestaties zijn vooral naar inhouden gespecificeerd en er is daar geen behoefte aan eenheid. Integratie komt ook niet voort uit mechanismen die verantwoordelijk geacht worden voor individuele verschillen. Ze variëren van snelheid van zenuwgeleiding tot de invloed van het gezin.

Piagets theorie heeft eenheid gebracht in het definiëren van ontwikkeling en in hoe ze verloopt en welke mechanismen verantwoordelijk zijn: assimilatie en accommodatie. Er is kritiek op de tijd waarop de stadia verschijnen en op het abrupte en vroege einde met de formele operaties. Het nut voor diagnostiek is beperkt en zijn thema: ontwikkeling het Westerse epistemisch subject is niet centraal in de psychologie.

De sociale (soms ook fysische) context en de gemanipuleerde omgeving, bijvoorbeeld scholing en instructie spelen een rol bij intelligentie en cognitie. In de gedragsgenetica worden de varianties met behulp van lineaire modellen verdeeld tussen genetische- en omgevingsvariantie. In de Piagetiaanse ontwikkelingstheorie is de rol van de sociale omgeving beperkt. Er is geen geïntegreerde theorie over de invloed van natuurlijke contexten. Er is eerder de overtuiging, dat alles 'wel helpt'. Er wordt in de theorie van een $S \rightarrow R$ schema uitgegaan terwijl omgeving en subject altijd interacteren. Er zijn veel en moeilijk te ordenen of te integreren mechanismen verantwoordelijk voor individuele verschillen en ontwikkeling van intelligentie en kennen.

Ik heb eerder gesuggereerd dat integratie niet zal lukken omdat de drie oriëntaties verwijzen naar niet tot elkaar te herleiden kenmerken van menselijk gedrag. Theorieën worden soms om uit te maken welke het minst 'belast' is door onvoldoende en strijdige informatie? Dit vergelijken is een gevolg van Poppers voorstel om cruciale experimenten uit te voeren die zouden kunnen beslissen tussen twee hypothesen of theorieën en een uitvloeisel van Kuhns paradigmashifts. De eerste geldt niet voor de oriëntaties in de psychologie en de tweede mag dan gelden voor natuurwetenschappelijke paradigmata (dat is ook niet onomstreden), maar in ieder geval niet voor de psychologie. Daar is eerder sprake van verschuiving van de aandacht en van verdieping. Oude theorieën worden niet afgedaan en vervangen door nieuwe. Als de aandacht ervoor verdwenen is betekent dat niet dat het een gevolg van het ontbreken van empirische steun is. De Gestaltpsychologie is bijna verdwenen uit de handboeken maar dat wil niet zeggen dat de Gestaltwetten onjuist zijn of te zeer belast zijn door het ontbreken van empirische steun.

2. Individuele verschillen x Ontwikkeling: Stabiliteit tegenover cognitieve ontwikkeling

De integratiepogingen bij intelligentie/cognitietheorie worden hier geanalyseerd aan de hand van drie combinaties van twee oriëntaties. Ze gaat *niet* over een integratie van wetenschappen zoals filosofen als Oppenheimer & Putnam, Fodor en Carnap voor ogen stond. Bij het vaststellen van *individuele verschillen* in intelligentie, geschiktheid en prestatie wordt de vraag over verandering door de tijd heen, ten eerste, opgevat als *stabiliteit en continuïteit*. Dit toont de dominantie van de individuele verschillen oriëntatie. Ten tweede

gaat diagnostiek over één cliënt. Generaliseren van steekproefparameters naar de cliënt is een probleem. Ten derde zijn er enkele pogingen tot integratie.

Homo- en heterotypische stabiliteit van intelligentie en cognitie Absolute stabiliteit wordt onder meer vastgesteld om na te gaan of er afname in IQ en Gf en Gc is met het ouder worden. De Gf neemt gemiddeld genomen af maar de Gc blijft gemiddeld gelijk en neemt bij enkelen zelfs iets toe. Waarneming, motorische vaardigheid, snelheid van informatieverwerking, leren en cognitie nemen af. De afname is groot, generaliseert over meer intelligentiefactoren en verschilt per individu (Salthouse, 2004a, 2004b). Birren en Schaie (2006) beweren dat die afname gecompenseerd kan worden door actief te blijven. Snelheid van informatieverwerking (responstijd) wordt als een niet-taakspecifiek kenmerk van de G beschouwd. Kail (1991) verzamelde 72 studies en concludeerde dat de responstijd met het toenemen van de leeftijd eerst lineair toeneemt maar daarna naar beneden gaat. Rangorde stabiliteit is een deel van de betrouwbaarheidsbepaling van meetinstrumenten. Je wilt beschikken over betrouwbare tests en dat betekent dat items die niet stabiel zijn door de tijd heen verwijderd worden. De waarden zijn meestal hoog: intelligentie is een van de rangorde - stabielste gedragskenmerken. Voorbeelden:

(1) Watkins en Smith (2013) onderzochten de lange termijn stabiliteit van de WISC. De test-hertest waarden werden bepaald in een steekproef van 344 studenten van twee schooldistricten die twee keer getest werden met het oog op plaatsing in het speciaal onderwijs. Het tijdsinterval was gemiddeld 2.84 jaar. De waarden voor de *Verbal Comprehension Index (VCI)*, de *Perceptual Reasoning Index (PRI)*, de *Working Memory Index (WMI)*, de *Processing Speed Index* en het *Full Scale IQ (FSIQ)* waren respectievelijk .72, .76, .66, .65 en .82. De vuistregel van Nunnally en Bernstein (1994) eist wat hogere waarden bij belangrijke beslissingen. De auteurs gebruikten niet de term factor, maar index. Ze erkennen daarmee dat 'onafhankelijke factoren' een te sterke uitdrukking zou zijn voor de combinaties van de Wechsler subtests die de vier indicatoren vormen. Ze wezen er bovendien op dat 25% van de studenten *full scale scores* toonden die 10 of meer IQ punten verschilden op de twee meetpunten en dat 29%, 39%, 37% en 44% VCI, PRI, WMI, en PSI scores toonden die 10 of meer punten verschilden. De steekproefwaarde is een gemiddelde en geldt niet voor iedere student. Gemiddeld verschilden de scores op de indices niet meer dan 2 punten: sommigen gaan immers vooruit en anderen achteruit.

(2) Deary (2014) heeft de rangorde stabiliteit (test-hertest met meer dan 60 jaar tussentijd) van het IQ van drie Schotse cohorten - twee van 1921 en een van 1936 - bepaald. De deelnemers werden op 11- jarige leeftijd en op 70 tot en 90 jarige leeftijd onderzocht met de WAIS III-UK of de Raven. De correlaties waren in de studies: $r = .67, .63$ (95% BI: .50 - .74), $.66, .51$ (95% CI .38 - .68) en .45.

Er is dus substantiële rangorde stabiliteit van de meest gebruikte intelligentietest. Het kan vertekenend zijn omdat niet 'betrouwbare' items al verwijderd zijn.

Structurele stabiliteit is geen onderzoeksthema bij *maximum performance*. Het is evenwel evenals bij de persoonsontwikkeling mogelijk om na te gaan of meer of minder factoren nodig zijn om de *positive manifolds* van tests door de tijd heen te beschrijven. Sternberg

suggereert dat verandering in factorstructuur een indruk kan geven van differentiatie en integratie van intelligentie door de tijd heen.

Ipsatieve stabiliteit wordt zelden vastgesteld. Het kan zinvol zijn want het is niet duidelijk of de rangorde van geschiktheden en prestaties behouden blijft door de tijd heen en of deze gevoelig is voor de context, bijvoorbeeld scholing.

Heterotypische stabiliteit komt aan de orde als je de relatie tussen voorlopers van het IQ en het latere IQ bestudeert bij jonge kinderen. De scores op de Bayley-ontwikkelingsschalen voorspellen het latere IQ. De correlaties zijn significant maar bescheiden. Onderzoekers verwachten hogere waarden en wijten het ontbreken daarvan aan de veranderende expressie van intelligentie, bijvoorbeeld van motorisch naar verbaal. Kavšek (2004) vermeldt een correlatie van $r = -.37$ tussen dishabituatie (ontregeld raken bij een nieuwe stimulus of situatie) van baby's en IQ als kind en adolescent.

Steekproef stabiliteit is geen individuele stabiliteit De typen stabiliteit worden bepaald in steekproeven. De correlaties zijn verre van perfect, ook niet na correctie voor attenuatie. Hetzelfde verschijnsel doet zich ook voor bij de consistentie, stabiliteit, continuïteit van de persoonlijkheidstrekken. Betrouwbaarheidscoëfficiënt kan verschillen tussen steekproeven en tussen individuele cliënten. Er is variatie op individueel niveau en die kan verschillen per leeftijdsperiodes. Leken en professionals nemen aan dat sommige periodes volatieler zijn dan andere, bijvoorbeeld de latentieperiode (basisschoolperiode) zou stabiel zijn dan de adolescentie. Een voorbeeld laat zien dat de interpretatie gevoelig is voor de tijdgeest:

Ramsden et al. (2011) vonden bij 33 gezonde, neurologisch normale adolescenten die ze twee keer onderzochten in 2004 en 2007-2008 ($M = 14.1$ jaar en $M = 17.7$ jaar) een test-hertest waarde van $r = .79$. Er was een substantiële variatie op individueel niveau. Die liep van minus 20 punten Verbaal IQ (VIQ) tot plus 23, en -18 tot + 17 Perfoormaal IQ (PIQ). Bij een betrouwbaarheidsinterval van 90% was er bij 39% een verandering in VIQ en bij 21% in PIQ. Deze subjecten vielen met andere woorden 'buiten de regressielijn'. De auteurs maten ook de grijze hersenmassa dichtheid op de twee tijdstippen. Omdat de veranderingen in VIQ correleerden met de veranderingen in de motorische spraakregio van de hersenen ($r = .251$ tot $.284$) en PIQ met veranderingen in het anterior cerebellum (kleine hersenen) ($r = .363$ tot $.437$) concludeerden ze dat de variabiliteit mede aan hersengroei gebonden kan worden.

Integratie van psychometrische intelligentie en cognitieve ontwikkeling Er zijn conceptuele en empirische inspanningen om twee tradities, de psychometrische en Piagetiaanse te combineren. Voorbeelden:

(1) De Griekse onderzoekers Andreas Demetriou en Anastasia Efklides (1987) hebben een *cognitieve intelligentie- en ontwikkelingstheorie* voorgesteld die gebaseerd is op Piaget-taken én op tests uit Thurstones PMA-traditie. Ze verzamelden longitudinale gegevens van kinderen met psychometrische tests en Piagetiaanse taken en formuleerden een multiple-factor theorie. Deze bevat de hypothese dat de menselijke geest verschillende factoren tot uitdrukking moet brengen bij het waarnemen en

leren kennen van de fysische en sociale wereld. Deze veroorzaken kennen in de vorm van getalsmatige, logische, verbale, ruimtelijke en sociale categorieën en dimensies. Dit is het gevolg van het 'actieve verstand' dat hun landgenoot Aristoteles al 2400 jaar eerder beschreven had. De categorieën moeten met confirmatieve factoranalyse terug te vinden zijn in de structuur van de gegevens (Demetriou et al., 1993). Daarnaast is er de hypothese van een geleidelijke, stadiumgewijze ontwikkeling en dat betekent dat er een natuurlijke volgorde te zien is. Ze voegen Piagetiaanse theorievorming toe omdat intelligentie 'genetisch' is, dat wil zeggen zich ontwikkelt van biologie (reflexen) tot logica (formele operaties) (Case et al., 2001; Demetriou & Efklides, 1987; Efklides & Platsidou, 1993; Demetriou & Raftopoulos, 2005). Deze uitgangspunten en de empirische gegevens leiden tot zes factoren die zich tot de volwassenheid ontwikkelen. De theorie van Demetriou et al. trok geen aandacht van Amerikaanse onderzoekers en er is daar dan ook nauwelijks onderzoek aan gedaan. Voor gebruik in de diagnostiek is de combinatie tijdrovend omdat hij veel taken individueel moet afnemen en als resultaat een complex profiel verkrijgt, waarvan de stabiliteit weinig onderzocht en de uitslag niet meteen te interpreteren is. Het model biedt in beginsel de mogelijkheid om een persoon te diagnosticeren aan de hand van een profiel dat bestaat uit zes factoren, met op elke factor het bereikte ontwikkelingsniveau. Het aantal stappen is groot en dat leidt ertoe dat enkele dicht bij elkaar liggen en moeilijk te onderscheiden zijn. De factoren en ontwikkelingsniveaus zijn:

I Kwantitatief-relatieve capaciteit verwijst naar de kwantificeerbare werkelijkheid (lengte, breedte en hoogte), naar de capaciteit om kenmerken van objecten en gebeurtenissen te tellen, naar inzicht in metrische systemen en naar de constructie van dimensies en relaties tussen reeksen getallen. Piagetiaanse conservatietaken, getalbegrip en principe van proportionaliteit laden op deze factor. De ontwikkeling verloopt in negen stappen tussen 3 en 22 jaar en gaat van pre-dimensioneel naar het hanteren van meer onafhankelijke dimensies tegelijkertijd.

II Kwalitatief-analytische capaciteit heeft betrekking op categoriseren en hiërarchiseren door verschillende criteria te gebruiken, die achtereenvolgens en tegelijkertijd aangewend worden om objecten, gebeurtenissen en mensen in te delen. Er zijn zeven stappen tussen 3 en 18 jaar en gaan van het willekeurig bijeenzetten van objecten tot het gebruik van een aantal onafhankelijke criteria tegelijkertijd.

III Ruimtelijk voorstellingsvermogen gaat over het ruimtelijk representeren van de omgeving door details weg te laten, door twee of meer voorstellingen in één geheel onder te brengen, door delen te herordenen en door voorstellingen mentaal te draaien. De zes ontwikkelingsstappen lopen via statische en onwrikbare voorstellingen tot het inzicht in mentale rotaties, bijvoorbeeld van geometrische figuren.

IV Oorzakelijke verbanden vaststellen gaat over combineren, hypothesen formuleren, experimenten bedenken en toetsen welke variabele of welke combinatie van variabelen de veroorzaker van een verschijnsel is. Zes stappen van 3 jaar tot het eind van de middelbare school lopen van denken zonder oorzaken ('het gebeurt gewoon') naar het bedenken en toetsen van oorzaken van verschijnselen.

V Verbale capaciteit verwijst naar het vormen en ontdekken van betekenisrelaties. Er worden tien stadia tussen 3 jaar en 18 jaar onderscheiden. Ze lopen van eenvoudige uitspraken tot het redeneren volgens formeel logische regels en het onderscheiden van inhoudelijke betekenissen volgens een systeem.

VI Metacognitieve reflectie gaat over het in staat zijn tot nadenken over de niet zo duidelijke wereld van gevoelens, ervaringen en gedachten. Er moet vastgesteld worden welke taken en gebeurtenissen nieuw zijn, welke elementen belangrijk zijn om een probleem te analyseren en op te lossen. De ontwikkeling gaat door tot in de volwassenheid en heeft ook betrekking op zichzelf realistisch zien, dat wil zeggen inzicht in eigen sterke en zwakke kanten.

(2) De Canadees Robert Case en de Griekse onderzoekers (2001) hebben elkaar gevonden en een deel van het model getoetst bij 120 zeven- tot elfjarigen. Ze namen de Wechsler-intelligentietest voor kinderen en een reeks Piaget taken af, bijvoorbeeld papier vouwen op een specifieke manier; een fles met water die scheef gehouden wordt en waarbij het kind moet tekenen hoe de stand van het water in de fles eruitziet; letters scannen; tekenen van iets met als doel te laten zien welk perspectief een ander inneemt, WISC-blokken om patronen te vormen en de balanstak. Ze gaan ervanuit dat de kinderen hun ruimtelijk, kwantitatief, verbaal/propositioneel, kwalitatief/analytisch (categorisch) en causaal denken mobiliseren bij het uitvoeren van deze taken. Case et al. (2001) analyseerden de prestaties als individuele verschillen met behulp van factoranalyses. De exploratieve factoranalyse leverde vijf factoren op:

I Ruimtelijk

II Kwantitatief

III Sociaal-verbaal

IV Logisch-analytisch

V Causaal denken

Confirmatieve factoranalyse (CFA) bevestigde deze vijf. Er was ook een G-factor aan te wijzen omdat de factoren niet onafhankelijk waren. Ze correleerden ook weer niet zo hoog dat de vijf niet onderscheiden konden worden. Ze interpreteerden de resultaten met behulp van de opvatting van Case over het gebruik van verschillende conceptuele structuren en met het idee van Demetriou et al. over gespecialiseerde capaciteitsdomeinen. Ze maakten ook gebruik van een hiërarchie in de geest van Vernon en Carroll: een model met een G-factor en specifieke groepsfactoren daaronder.

De complexiteit van deze metingen, de tijdrovende afname en de onzekerheid over de scheiding tussen de stadia op elke factor zorgen ervoor dat de diagnosticus niet gemakkelijk van dit model gebruik zal maken. Het biedt wel een mogelijkheid tot verrijking van de interpretatie van klassieke intelligentietaken en Piaget proeven.

Er zijn kortom pogingen om psychometrische intelligentie en ontwikkeling te combineren. Het is een hybride die een genuanceerde beschrijving van intelligentie bevat. Het zal echter niet gemakkelijk zijn weg naar de diagnostische praktijk zal vinden, ik denk zelfs nooit ofte nimmer. Af en toe verschijnen er essays die toch de hoop op zo'n samenwerking tonen. De Ribaupierre (2015) uit de Piagetaanse school noemt het een *droom*. Dit type studies is complex want het vereist

(a) het bestuderen van meer uiteenlopende taken op één tijdstip

(b) herhaald verrichten van taken om binnen-taak-variabiliteit vast te stellen

(c) het bestuderen van samples om gevolgen (covariaten) van demografische en persoonsvariabelen te schatten

(d) longitudinaal onderzoek met oog voor intra-individuele veranderingen

(e) nieuwe methodologie en statistiek. Die is er overigens al, bijvoorbeeld het werk van Bergman en de vele statistische programma's om al die lagen (multi-level analyse) tegelijk te onderzoeken.

Het zal wel bij een droom blijven en je kunt je afvragen wat dit nu oplevert. Waarschijnlijk weer gesegmenteerd nieuws en bekends over individuele verschillen en ontwikkeling.

Samenvatting en conclusie

Absolute verschillen worden vooral vastgesteld om te bepalen of kinderen, adolescenten en volwassenen winst maken door de tijd heen en om na te gaan of er met het ouder worden afname is. De uitslag: de Gc blijft gemiddeld gelijk maar de Gf neemt af, evenals motorische vaardigheden en responstijd, een index van de G-factor. Absolute verschillen op jonge leeftijd vallen weg. Ze zijn niet zichtbaar door de leeftijdsgewijze normering van IQ- en prestatietests. Test-hertest coëfficiënten van intelligentie-, geschiktheids- en prestatietests geven informatie over rangorde stabiliteit. Deze is doorgaans hoog, ook al omdat items verwijderd worden die de betrouwbaarheid 'drukken'. Structurele en ipsatieve stabiliteit zijn nauwelijks onderzocht. Heterotypische stabiliteit is bescheiden bij vergelijking van scores op de Bayley-test met het IQ op de WISC. Dishabitatie (ontregeld raken bij een nieuwe stimulus of situatie) bij baby's correleert gemiddeld met het IQ als kind en adolescent.

Demetriou et al. combineerden Piaget taken en items uit multiple individuele verschillen intelligentietests. Ze legden ze voor aan kinderen, adolescenten en volwassenen. Ze trokken zes factoren en specificerden bij iedere factor een aantal ontwikkelingsstappen tussen 3 en 18-22 jaar. Dat aantal stappen is groot met het risico dat ze moeilijk te onderscheiden zijn en overlappen. De factoren zijn: I kwantitatief-relacioneel, II kwalitatief-analytisch, III ruimtelijk voorstellen, IV causaal-relacioneel, V metacognitief-reflectief en VI verbaal. Er is steun uit onderzoek met de WISC en Piagetiaanse taken door Case et al.: er zijn vijf factoren getrokken: I ruimtelijk, II kwantitatief, III sociaal-verbaal, IV logisch-analytisch en V causaal denken. Al lijkt deze integratie geslaagd, het is en blijft een hybride. Dat houdt in: iets sterker dan de twee afzonderlijk, net als een muilezel, maar niet vruchtbaar voor verder onderzoek en vooral de praktijk: de muilezel kan zich niet voortplanten. De diagnosticus zal deze hybride niet gebruiken als beschrijving van de cliënt: te complex en tijdrovend. Er is mij geen voorbeelden bekend van gebruik en toepassing van dit cognitief ontwikkelingsprofiel bij een cliënt. In essays wordt de verbinding gepropageerd, maar zelden gematerialiseerd in empirisch onderzoek.

3. Individuele verschillen x sociale context: Aanleg tegenover omgeving

Deze combinatie is aanleiding geweest tot verhitte debatten, denk aan Jensen & Rushtons en Hernnstein & Murray's *nature* standpunt, aan Neisser en Gottfredsons pogingen een balans te vinden tussen *nature* en *nurture* en aan de overzichtsstudies van Nisbett et al. (2012), Deary (2012) en Deary et al. (2010), zie boven. Enkele thema's worden besproken:

Ten eerste, komt de verdeling van de IQ varianties over genen en omgeving aan bod en de lastige opgave ze te ontwarren. Ten tweede worden gevolgen van historische tijd en historische veranderingen (verschillen tussen generaties) gerelateerd aan IQ: het Flynn effect. Ten derde blijkt SES, een kenmerk van de sociale context waaraan een persoon niet (meteen, zomaar of als wilsbesluit) aan kan ontsnappen. Het is niet zijn vrije keuze maar is wel de 'habitat' waarin hij zich aantreft: 'U zit niet in uw SES; uw SES zit in U'. Er is sociale mobiliteit maar die is beperkt: geen enkele krantenjongen wordt miljonair ook niet in de VS. De correlatie tussen SES en IQ is gemiddeld en robuust dat wil zeggen wordt telkens opnieuw gevonden. Ten vierde, scholing is een schoolvoorbeeld van een ontworpen context. Mate en niveau van scholing hangen samen met intelligentie, geschiktheid en prestatie. Onderzoekers proberen effectieve ingrediënten in scholing te omschrijven. Ten vijfde is er aandacht voor een thema dat leraren bezighoudt: de relatie tussen klassengrootte en schoolprestaties. Ten zesde, wordt de relatie tussen speciaal ontworpen programma's (interventies) en maximum performance besproken.

Ontwarren van varianties van genen en omgeving Het meeste *nature-nurture* onderzoek wordt gedaan met G-verzadigde IQ tests. De varianties worden opgeteld zoals bij onafhankelijke orthogonale factoren of als hoofdeffecten in de variantieanalyse. Bij het interpreteren van verschillen in intelligentie, geschiktheid en wat minder bij prestaties, wordt aanleg - mate van genetische overeenkomst - gezet tegenover sociale context - verschillen tussen en binnen gezinnen. Omdat de genetische variantie meer verklaart in de totale IQ variantie in de populatie (ongeveer 50% en een flinke variatie van wel 20%) dan de varianties tussen en binnen gezinnen (ongeveer 20%) wordt de aanleg belangrijker genoemd. Afhankelijk van de aangehangen ideologie worden interventies nutteloos genoemd of aanbevolen om de verschillen te verminderen. De 50% schatting berust op studies met één- en twee-eiige tweelingen, broers en zussen en niet-verwante personen. Gedragsgenetische studies gaan over varianties van kenmerken in populaties en niet over individuen. Het is zinloos om te zeggen: uw IQ is voor 20% bepaald door het gedrag van uw ouders en uw sociale omgeving en voor 50% door de met hen gedeelde genen. Het model schat de bijdragen van de bronnen in de populatie apart en telt ze op.

Genen drukken zich *altijd* uit in een omgeving. Je kunt niet verdedigen dat één of een combinatie van genen een gedrag 'veroorzaakt'. De stap van genen naar kleur van de ogen en de lichaamslengte is al moeilijk te volgen, laat staan die van combinaties van genen naar het correct beantwoorden van vragen uit een intelligentietest. Bij mentale retardatie zijn, als eerder vermeld, 262 genen betrokken en bij lengte zijn ze al rond de 700 aangeland die zijn ook weer bij andere kenmerken betrokken. We komen er dan ook niet achter hoe de genetische uitrusting nu precies werkt om een concreet probleem op te lossen, een taak uit te voeren of een item te beantwoorden.

De schatting van de variantiecomponenten berust op lineair veronderstelde samenhangen, niet op oorzaak-gevolgrelaties. Genen veroorzaken niet rechtstreeks gedrag en verklaren dus ook niet het verschijnen van intelligente handelingen. Genetische modellen zijn lineair

en dat kan vreemd uitwerken: Stel er zijn twee factoren A en Z. Ze zijn noodzakelijk en voldoende om een uitkomst tot stand te brengen. Als A niet varieert in een bepaalde context dan verklaart Z alle variantie. Neem redactiesommen, leerlingen moeten het vraagstuk lezen (A: alle leerlingen lezen perfect) en hun kennis van vierkantsvergelijkingen gebruiken (Z). We willen inderdaad Z weten maar A blijft noodzakelijk ook al verklaart Z *alle* variantie in deze steekproef. Of, neem het voorbeeld dat we twee armen en twee benen hebben. We hebben die bijna allemaal, dus lage variantie en dat kan dan ook niet veel variantie verklaren of binden. Dat is vreemd want er zijn zeker genen betrokken bij onze 'twee-armigheid' en 'twee-benigheid'.

Gedragsgenetici hebben hun werk goed gedaan. De omgevingskundigen nog niet: welke omgevingsvariantie draagt bij aan variantie in IQ, ontwikkeling, cognitie en geschiktheid? Er zijn wel veel correlaties maar niet van het niveau van de h^2 : de overerfbaarheidscoëfficiënt. Ze zijn groot in aantal, divers, laten lage correlaties zien en overlappen. Dat levert geen helder beeld op. Ze worden in complexe modellen ondergebracht, bijvoorbeeld padanalyses en structurele vergelijkingen. Al zijn de plaatjes indrukwekkend, ze helpen nauwelijks om helder te maken wat nu de meestal zwakke (in)directe invloeden van al die omgevingsvariabelen op intelligentie en cognitie precies zijn. Het is wel waardevol dat plausibele modellen vergeleken worden en de beste - de meest passende - overblijft. Hierbij geldt het Fodor principe dat het te beschrijven/verklaren gedrag multiple realiseerbaar is. Er is niet één oplossing. Het is geaccepteerd dat aanleg en omgeving interacteren en dat een deel van de variantie vastligt en een deel niet. Dat ongebonden deel van de variantie wordt niet helemaal opgesoupeerd door verschillen tussen en binnen gezinnen. Er blijft iets 'onverklaard' dat mogelijk ooit nog verklaard wordt, maar wellicht moeten we aanvaarden dat *toeval* bestaat en dat kunnen we per definitie niet verklaren (Rutter, 2007).

De aanleg-omgeving discussie heeft dus niet tot integratie geleid, wel tot een zo nu en dan oploeiende confrontatie en voorstellen om de samenhang tussen veel kenmerken van de omgeving en IQ af te beelden met ingewikkelde structurele vergelijkingstechnieken.

De discussie en resultaten van *nature versus nurture* zijn van beperkte betekenis voor de diagnosticus. Ze zijn relevant omdat de cliënt er een opvatting over heeft en zijn handelen erdoor kan laten leiden. 'Het is aanleg', 'ik heb het ook', zeggen ouders regelmatig bij een probleem met hun kind. De diagnosticus weet dat dit verandering niet uitsluit. Daarnaast moet hij wellicht toevalsfactoren erkennen en die heeft niemand in de hand, evenals de '*weemoedigheid*', die komt als het avond is en die '*niemand kan verklaren*' uit het gedicht 'Het huwelijk' van Elsschot.

Flynn Effect Het idee van vooruitgang in IQ, SAT, geschiktheden en SES door de tijd heen krijgt steun in het Flynn effect. James Flynn (Dunedin, Nieuw Zeeland) rapporteerde al voor 1980 over de winst in IQ scores over generaties. Als te verwachten was hij niet de eerste. Lynn (2013) wijst erop dat deze studies al vanaf 1936 gedaan zijn door de Amerikaan Runquist. Het zou dus ook het Runquist-effect kunnen heten. Deze auteur vond in 1936 al generatieverschillen in intelligentie: de mediaan van de scores op de Minnesota College

Aptitude Tests nam tussen 1929 en 1932 toe: van 33.7 naar 42.3. Dit is een korte periode en niet te vergelijken met een generatie.

Het Flynn effect is vaak gemeld en onderzocht. Er worden hieronder voorbeelden gegeven en er is naar tegenvoorbeelden gezocht. Het is intuïtief wellicht moeilijk in te zien of te aanvaarden dat elke generatie zo'n 10 IQ punten vooruitgaat. Of zoals de generaties vanaf 1910-1930 zeiden als ze zagen dat zoveel kleinkinderen HBO- en universitaire studies volgden: 'We krijgen gebrek aan domme mensen'.

(1) Flynn rapporteerde in 1987 resultaten uit veertien landen onder meer uit Europa en de VS, Japan, Australië en Nieuw Zeeland. Hij vond *massive gains* in SAT- en IQ-scores (Raven Progressive Matrices, Wechsler, Otis - een test voor rekruten -, verbale- en wiskundetests) over een periode van dertig jaar. De toename liep van 1.60 punten in Noorwegen tot 20.3 punten in Japan (1 SD = 15 punten). In één generatie was er winst tussen de 5 en 25 punten. Gemiddeld genomen wint men per generatie één SD (15 IQ punten). In 1998 rapporteerde hij vervolgens verschillen tussen domeinen: cultuurarme tests (G-fluid, Raven) laten winst van rond de 20 punten zien, per formaal IQ tussen de 10 en 20 en verbaal IQ < 10 punten. Tests die dicht bij onderwerpen die op school geleerd worden staan, bijvoorbeeld rekenkundig redeneren, algemene informatie en woordenschat laten weinig of geen winst zien. De vooruitgang is volop aanwezig bij kinderen. De oorzaken moeten met andere woorden vooral op die leeftijd hun invloed uitoefenen. Later is dit tegengesproken: Trahan et al. (2014) vinden *geen* steun voor leeftijd als moderator variabele. Ook capaciteitsniveau blijkt geen moderator: er deed zich geen toe- of afname in het aantal hoog- en laagbegaafden voor in twee generaties.

(2) Wai en Putallaz (2011) bekeken 1.7 miljoen scores van 11-13 jarigen op de SAT en twee andere prestatietests: ACT en EXPLORE van 1981 tot 2010. Ze vroegen zich af of het Flynn effect ook te vinden was bij de 5% hoogste scoorders: de hoogbegaafden. Het effect was aanwezig en liet zien dat er over de hele intelligentie normaalcurve een toename was en een toename met een constante snelheid. Het effect werd bij meisjes en jongens gevonden en was vooral zichtbaar bij de rekensubtests van de SAT, ACT en EXPLORE. Het Flynn effect wordt kennelijk ook gevonden bij *outliers*. Ze vonden verder dat de m/v verhouding bij *outliers* weliswaar afnam maar het kleine verschil nog wel een reden zou kunnen zijn voor het feit dat vrouwen in bepaalde hoge functies ondervertegenwoordigd zijn. Hoewel politiek incorrect, kunnen motieven, belangstelling en ambitie factoren zijn. Misschien nog incorrecter: het verschil tussen mannen en vrouwen kan gebaseerd zijn op het vereiste apenrotsgedrag en daar zijn mannen beter in dan vrouwen, evenals in het graaien naar bonussen en hoge salarissen. Wai en Putallaz's gegevens (zie p. 338, Tabel 10) laten de *jaarlijkse* veranderingen zien: de veranderingen bij meisjes zijn groter op de SAT-samengestelde scores (0,07), SAT-Rekenen (0,04) en EXPLORE-Engels (0,09). Meisjes behielden hun scores op de SAT-taal maar jongens gingen achteruit (-0,10). Jongens gingen ten opzichte van meisjes meer achteruit op ACT-Wetenschap. De resultaten betreffen een beperkte groep: de 5% hoogste scoorders en de gemiddelde verschillen zijn klein want er zijn er bijna evenveel die vooruit als die achteruitgaan.

(3) Het Flynn effect is ook elders gevonden. De Chinese adaptatie van de WISC liet een winst van 4.53 (*full scale*) zien tussen steekproeven van 1984 en 2006. Dat is 2.06 IQ punt per 10 jaar. Het verbale IQ won 4.27 en het performale IQ 4.08 (Liu et al., 2012). In Saudi-Arabië vonden Batterjee et al. (2013) verschillen tussen 1977 en 2010 op Raven's Progressive Matrices. De gemiddelde score

nam substantieel toe: $d = 0,78$, dat is 11.7 IQ punten in een generatie: ongeveer 30 jaar). Per 10 jaar was er een vooruitgang van 3.55 IQ punten.

(4) Een recente en weer wat rigoureuze meta-studie van Trahan et al. (2014) omvat 285 studies ($N = 14.031$). Ze verzamelden sinds 1951 scores op twee intelligentietests. Dat leverde een gemiddelde winst van 2.31 IQ punten per decade op (Betrouwbaarheidsinterval: BI, 95% 1.99 - 2.64). Een vergelijking met Stanford-Binet en Wechsler IQ schalen laat een winst per 10 jaar van 2.93 (BI, 95% 2.30 - 3.50) bij 53 vergelijkingen ($N = 3.951$). De meta-studie laat ook zien dat er geen afname (of toename) door de tijd heen is en dat leeftijd en capaciteitsniveau (in tegenstelling tot de Wai & Putallaz's, 2011 bevinding) geen moderator variabelen zijn.

Het gemiddelde beeld is dat IQ scores op bekende tests in 10 jaar tussen de 2.50 en 3.50 IQ punten toenemen in westerse samenlevingen. Het effect is ook buiten deze landen te vinden. De waarden kunnen daar soms wat hoger zijn en meer variëren door de toename van toegang tot het onderwijs.

Het Flynn effect wordt dus in empirische studies in verschillende culturen en herhaald aangetroffen. Verderop worden tegenvoorbeelden vermeld. Als oorzaak voor de vooruitgang wijst Flynn (1998, 2007) op de omgeving, vooral onderwijs op jonge leeftijd en cognitieve stimulering. Er is ook commentaar. Er zouden geen echte IQ tests maar schoolprestaties vergeleken zijn. Dit is slechts voor een beperkt deel terecht. Flynn (2007) noemt in een samenvatting de volgende oorzaken voor de IQ stijging:

Onbekende omgevingsfactoren

Ervaring met gestandaardiseerde tests

Voeding

SES

Verstedelijking

Bestrijden en uitroeien van kinderziekten

Historische rampen als de WO II

Verbetering van scholing

Onderwijsprogramma's op de televisie

Opleiding in het algemeen

Hij gaat er niet van uit dat het effect berust op genetische, biologische en psychologische veranderingen. Deze nemen meer tijd in beslag dan één generatie.

Sommige auteurs delen zijn *interpretatie* niet en pleiten voor een genetische basis. Voorbeelden:

(1) Rushton en Jensen (2010) verdedigen de genetische basis van de G-factor. Ze vermelden gegevens waaruit blijkt dat het WAIS-IQ-verschil tussen zwarten en blanken *binnen* de VS stabiel is gebleven tussen 1954 en 2008. Beide auteurs hebben van meet af aan in het IQ debat de genetische bepaaldheid verdedigd. De IQ verschillen waren in 1954 15 punten (1,0 SD) bij 17-jarigen, 18 punten in 1965 en 19 in 2008. Deze bevinding laat zien dat ofschoon er vooruitgang is dat de verschillen tussen groepen niet hoeven af te nemen. Het is vergelijkbaar met economische groei waarbij de *verschillen* tussen arm en rijk gelijk blijven en soms zelfs toenemen terwijl allen *vooruitgaan*. De verschillen tussen SES groepen liggen niet vast: ze worden de laatste tijd weer groter: de 85 rijksten

bezitten 80% van de aarde. Dit wordt wel het Mattheus effect genoemd, een van de vier schrijvers van het Nieuwe Testament (Nederlandse vertaling, 2004). Deze vermeldt dat Christus, de hoofdpersoon van het Nieuwe Testament, zei dat hij er niet altijd zou zijn maar de armen wel. Deze uitspraak deed hij naar aanleiding van een bezwaar van een apostel (een fervent aanhanger van zijn leer: een 'groupie') tegen het gebruik van kostbare olie om zijn voeten te balsemen. De Rushton en Jensen studie ontkracht het Flynn effect overigens niet want er kan over generaties vooruitgang zijn bij zowel blanken als zwarten.

(2) Mingroni (2007) gaat evenals Rushton en Jensen tegen de contextverklaring van het Flynn effect in. Hij voert *heterosis*' aan, een genetisch effect dat tot stand komt doordat exemplaren van verschillende genetische populaties mengen. Dit kan zich nog vanaf 1950 hebben voorgedaan in Europa en al eerder in de VS want kleine geïsoleerde gemeenschappen verdwenen en veel jonge mensen trokken naar de steden. De trend kan zich ook nu nog voordoen: veel jonge mensen gaan studeren en mengen met andere studerenden. Hij geldt met andere woorden niet alleen voor ontwikkelingslanden. Hij doet zich ook in Nederland voor want sommige provincies ontvolken en vergrijzen. Mobiliteit is de oorzaak van genetische veranderingen en die 'veroorzaken' op hun beurt veranderingen in intelligentie, geschiktheid en prestatie. Woodley (2011) kritiseerde dit en vermeldt de proportie. Als heterosis een rol zou spelen dan zou de bijdrage aan de IQ stijging sterk afgezwakt worden door de grotere winst op basis van de ontwikkeling van de mensen en van hun levensgeschiedenissen. De simulaties van Mingroni over 50 jaar met credits voor de verhoogde demografische mobiliteit kunnen volgens hem niet meer opleveren dan 3 IQ punten (0,20 SD) per generatie. Dat is veel minder dan Flynn vermeldt.

De IQ winst per generatie is volgens deze auteurs dus een gevolg van veel factoren, genetisch en omgeving: vroeg naar school, stimulering thuis, culturele veranderingen, individuele levensgeschiedenissen en de interacties tussen deze klassen van factoren. Roivainen (2012) geeft een argument voor een directe causale relatie van GDP en onderwijs naar IQ, dat wil zeggen van de economische en culturele omgeving naar IQ winst. Hij vermeldt onderzoek uit het Oost-Duitsland van voor de *Wende* op 9 november 1989. Tussen 1990 en 2006 bleken de scores van de rekruten uit de DDR met 0.5 IQ punt per annum gestegen. De winst was met een r van .89 gecorreleerd met het bruto nationaal product en met een correlatie van .78 met de verbetering van het onderwijs.

Het Flynn effect lijkt algemeen aanvaard. Er is dus vooruitgang, al zijn de mechanismen niet precies aan te geven en wordt er van alles bijgehaald worden om het effect te verklaren. Er zijn *tegenvoorbeelden*:

(1) Een omgekeerd Flynn effect doet zich voor bij Piaget taken. Shayer et al. (2007) gebruikten in 1975/1976 verzamelde gegevens van deze taken en vergeleken die met gegevens uit 2003. Er blijkt een terugval voor jongens ($d = -1,04$) en voor meisjes ($d = -0,55$). Dit zijn gemiddelde tot grote effecten. Het Flynn effect geldt met andere woorden *niet* voor Piagetiaanse taken. Ze hadden er geen goede verklaring voor en vermeldden als redenen: computerspelletjes; televisiekijken; afname van buiten spelen en van discussiëren met leeftijdgenoten; nadruk op lezen, schrijven en rekenen: de *back to basics* beweging in het basisonderwijs.

(2) In Pune, India is een differentiële aanlegtest (IAM: *Intelligence-Aptitude-Measurement*) ontwikkeld om ouders en leerlingen te helpen bij de keuze voor vervolgonderwijs. Meer dan 2200

leerlingen vanaf 12 jaar hebben deze gecomputeriseerde test gemaakt. De laatste vijf jaar is er een *afname* geconstateerd op de subtests voor rekenen en wiskunde (persoonlijk communicatie Dr. S. Watve, directeur van het Jhana Prabodhini Instituut in Pune dat de tests afneemt: 02-09-2013). De resultaten kunnen niet meteen raak geïnterpreteerd worden. De steekproeven kunnen gewijzigd zijn: er is een verschuiving naar hogere opleidingen in Maharashtra (India), in onderwijsprogramma's, kwaliteit van leraren en behuizing van de scholen. Watve wees echter op dezelfde factoren als Shayer et al. Ze kende die studie niet.

(3) Gignac (2014) ging na in hoeverre er een Flynn effect te bespeuren viel in geheugen subtests (WAIS): *Digit span forward* en *Digit span backward*: het nazeggen van een reeks willekeurige getallen in dezelfde of omgekeerde volgorde. Hij kon gegevens traceren over 85 jaar (1923-2008) bij 7.007 en 6.841 deelnemers. De werkgeheugen scores namen niet toe door de tijd heen. Het Flynn effect geldt hiervoor niet. De waarden bleven gelijk: dezelfde volgorde 6.56 ± 2.39 en de omgekeerde 4.88 ± 2.58 . Niettemin correleren deze geheugenprestaties hoog met Gf en die vertoont het effect wel.

(4) Woodley en Meisenberg (2013) vonden een verlies van -1,35 punt per decade in vier Nederlandse studies met de Differentiële Aanleg Test (DAT), de General Aptitude Battery (GATB) en de Amsterdamse Kinder Intelligentie Test (AKIT). De geboortecohorten waren van 1950-1990. Dit wijkt af van Flynn's samenvatting van 2009 dat de Gf tussen de 6,57 en 7,14 punten (+/- 0.5 SD) toeneemt per generatie. De auteurs suggereren dat geschiktheidstest minder gevoelig zijn voor vooruitgang dan de Raven. Het is ook mogelijk dat het effect terugliep door de verandering van de steekproef: latere studies hadden (meer) immigranten in de steekproef. Bij Finse rekruten werd zowel vooruitgang (4.0 punten op tests voor vormen, woorden en getallen) tussen 1988 en 2007 gevonden, als achteruitgang voor perceptuele, verbale en numerieke tests van 1997 tot 2009. Dit komt ook voor in het VK en in andere Scandinavische landen. Afname wordt meestal toegeschreven aan immigratie. In 2010 waren 13% van de Noren afkomstig uit landen buiten Europa, 11% van de Denen, 11% van de Engelsen en ongeveer 20% van de Nederlanders. Het laatste getal lijkt hoog, maar zie het artikel van de auteurs, het gaat om westerse en niet-westerse immigranten samen, ongeveer 3.0 miljoen van de bevolking in 2010. Dit verschilt van Finland: 4.8% Finnen was in 2010 immigrant. Er is nog een interpretatie toegevoegd. De verschillende vruchtbaarheid van bevolkingsgroepen, i.e., de negatieve correlatie tussen het aantal kinderen en intelligentie: lage IQ ouders hebben meer kinderen. Zo is bijvoorbeeld de Iranese bevolking goed opgeleid in vergelijking tot de omringende landen omdat er stilzweigend een twee-kind plan in werking zou zijn getreden. Feitelijk blijkt het gemiddeld aantal kinderen per vrouw in Iran zeer laag te zijn 1,3. De negatieve 'vruchtbaarheid-IQ' correlatie wordt geschat op 1.0 IQ punt per decade (Dutton & Lynn, 2013).

Het Flynn effect doet zich dus weer niet altijd en ook niet bij ieder cognitief vermogen of intelligente activiteit op dezelfde wijze en in dezelfde omvang voor. Niettemin wordt een populair en positief verschijnsel als het Flynn effect gemakkelijk gegeneraliseerd naar andere steekproeven en vaardigheden. Voorbeelden:

(1) Lynn (2009) stelde een toename in DQ (Developmental Quotient) vast met de *Bayley Scales of Infant Development* en de *Griffith Scales*. De DQ waarden namen 3,7 punt per decade toe en dat is vergelijkbaar met de IQ winst. De schalen zijn gemaakt voor jonge kinderen van 3 maanden tot 3 jaar. Lynn noemt als redenen voor de winst de effectiviteit en verbeteringen in de prenatale en

vroege postnatale zorg. Er is geen samenhang met verbetering van onderwijs en met ervaring met testen.

(2) Williams (er zijn veel Williamsen: R. L., 2013) verrichtte een overzichtsstudie over het Flynn effect in een nummer van *'Intelligence'*: <http://doi.org/10.1016/j.intell.2013.04.10> (p. 1): Na het positieve nieuws van voor 1980 is duidelijk geworden dat de winst niet beperkt is tot een kleine groep of tot bepaalde naties. De manifestatie en omvang van de winst waren afhankelijk van plaats en tijd. En voor ieder positief resultaat was er wel een tegenvoorbeeld te vinden. De winst verscheen op G- en abstracte tests als de Raven maar ook op Gc: uitgekristalliseerde intelligentie. De winst is soms wel en soms niet gelijk verdeeld over de IQ range. Er zijn gegevens die suggereren dat de winst groter is voor de linkerhelft van de IQ normaalverdeling (regressie naar het gemiddelde?) maar andere gegevens laten gelijke winst op elk niveau zien. Twee studies hierboven suggereerden immers *gelijke* vooruitgang bij de hoogste 5% door de generaties heen. De oorzaken of redenen zijn divers: fysische en sociale omgeving, genetische factoren, afnemende fertiliteit en de onderzoeksmethode.

Het Flynn effect is na een veelbelovende start ook weer niet het robuuste en eenduidige verschijnsel geworden dat sommigen er van verwachtten of hoopten.

SES en prestaties Een veel onderzochte ontmoeting tussen IQ, geschiktheid, prestatie en context is die met de Sociaal Economische Status (SES). Het is een complexe variabele en als zodanig met veel verschijnselen gecorreleerd, zoals schoolbezoek, schoolkeuze, opvoedingsstijl en inkomen, stad of platteland en toegang tot gezondheidszorg. De relatie SES-IQ is in veel studies bepaald en geïnterpreteerd. Voorbeelden:

(1) Kraus et al. (2012) leggen uit dat sociaal economische klasse *ook* verwijst naar de *perceptie* van de eigen rangorde in de samenleving: een gepercipieerde 'pikorde'. Een waargenomen lage rang schept een omgeving in lagere-klasse-personen een *contextualist tendency*: een focus op externe, niet te controleren sociale krachten en het idee dat anderen je leven beïnvloeden. Stijging in rangorde gaat gepaard met gebruikmaken van de persoonlijke vrijheid en het creëert solipsistische sociale cognities met nadruk op eigen interne toestanden, doelen, motivaties en emoties. Uit dit kader leidden de auteurs hypothesen af en lieten zien dat de sociale-klasse-gebonden tendenties van invloed zijn op het zelfbeeld, de waarneming van de sociale context en van relaties met anderen.

(2) Tucker-Drop et al. (2014) nemen als het te voorspellen kenmerk de wetenschappelijke interesse van 400.000 middelbare scholieren in 57 landen. Ze proberen dat te prediceren uit hun SES en bruto nationaal product (BNP). Een naar verhouding hoog BNP en SES voorspelden als te verwachten interesse in *science* en een nadruk op studeren en presteren. De complexe architectuur van SES maakt een dekkende meting moeilijk maar staat niettemin toe dat veel van zijn kenmerken gecorreleerd zijn met IQ, geschiktheid en schoolprestaties, zoals: schoolbezoek, kans om te spelen, beschikbaarheid van speelmateriaal dat op taken op school lijkt, onderwijs televisie, verfijnde typen instructie, speciaal onderwijs en toename van instructietijd, leerlingen 'testwijs' maken, gedragsverandering bij ouders: meer vrijheid om te exploreren, opvoedingsstijl van autoritair naar autoritair, interesse in de prestaties van hun kinderen, woonplaats: stad of platteland, toestaan van vriendjes, beschikbaarheid van gezondheidszorg voor zwangere vrouwen en jonge kinderen, solipsistische cognitieve tendenties bij hoog SES subjecten tegenover de focus op moeilijk te

controleren sociale omstandigheden bij laag SES subjecten en interesse in de wetenschap van hoog SES middelbare scholieren uit landen met een naar verhouding hoog BNP.

SES hangt dus samen met prestaties van leerlingen en studenten (zie bijvoorbeeld Sackett et al., 2009). SES was en is aanleiding om interventies in te zetten bedoeld om ouders en leraren te ondersteunen. Er doet zich bij interventies het Mattheus-effect voor (een term van de socioloog Merton, 1968), bijvoorbeeld het behoud van winst van een interventie is bij moeilijk opvoedbare kinderen moeilijker voor de lagere SES ouders dan voor ouders uit de midden SES (Leijten et al., 2013). Von Stumm en Plomin (2015) toonden bovendien aan dat het SES verschil in IQ op tweejarige leeftijd 6 IQ punten is, maar op zestienjarige leeftijd bijna verdrievoudigd is. De relatie SES-IQ kent een intercept verschil: lage SES kinderen beginnen lager en een slope verschil: het wordt groter, verdrievoudigd tussen twee en zestien jaar. De auteurs onderzochten dit bij 14.853 kinderen en adolescenten uit de *Twins Early Development Study* waarbij negen keer het IQ bepaald werd tussen 2 en 16 jaar. Ze stelden het intercept vast en pasten de slope lineair en kwadratisch. De lineaire relatie voldeed. Het effect dat de minst behoeftigen het meest krijgen dankt als boven gezegd zijn naam aan de apostel van Christus die zei (Mattheus 13:12): *Hij die bezit, zal meer en overvloedig krijgen, maar hij die niets bezit, zal beroofd worden van wat hij heeft*. Dat liegt er niet om en lijkt een vroege voorbode van opvattingen en feitenmateriaal in *Das Kapital* van Marx en van Piketty.

Onderzoek laat vaak zien dat het effect van interventies verschilt voor SES: de hogere klassen winnen meer dan de lagere klassen. In de VS leggen ongeveer acht miljoen kinderen jaarlijks *college admission tests* af. De prestaties zijn gemiddeld gecorreleerd met SES. Sackett et al. (2009) hebben een samenvatting gegeven van de relaties tussen SES en schoolprestaties gemeten met SAT's en Grade Point Average, GPA bij 167.816 Amerikaanse leerlingen. SES en SATs correleren in een grote representatieve steekproef: $r = .42$; De SAT-scores van de toegelaten studenten correleren $r = .35$ met toekomstige prestaties. Als gecontroleerd wordt voor *restriction of range* dan is de waarde $r = .47$; De tegenwerping dat de SATs en GPAs slechts SES meten is dus niet terecht. Als SES eruit gehaald, uitgepartialiseerd wordt, blijft er een correlatie van $r = .44$ over; mét SES: $r = .47$. Er zijn wel veel andere factoren maar die zijn vaak gecorreleerd met SES.

Murray (2012) bekend van zijn controversiële boek met Herrnstein (1994) over het IQ verschil tussen blank en zwart in de VS en de nutteloosheid van interventies heeft geschreven over het ontstaan van een *nieuwe* tweedeling in de samenleving. Deze vindt plaats in de blanke bevolking tussen de 30 en 49 jaar. Er komt een nieuwe upper-class, ongeveer 5% met de hoogste posities in het bedrijfsleven en met gewaardeerde goed betaalde beroepen. Hij wijst onder meer medisch specialisten, advocaten en een enkele academicus (laag salaris, maar tot voor kort met enig prestige). Daarnaast komt er een nieuwe onderklasse van personen met een high school diploma die werken in eenvoudige niet-technische functies. De auteur noemt ze Belmont tegenover Fishtown, gebaseerd op gemeenschappen in Philadelphia in de VS. Deze dichotomie correleert met IQ. Murray

noemt als een oorzaak voor de scheiding: migratie van de leiders van Fishtown naar de middenklasse suburbs. Zijn conclusie is gebaseerd op het opleidingsniveau en de cognitieve complexiteit van beroepen en de Philadelphia demografische statistieken van 1960 tot 2010.

Scholing en prestaties Scholing is de ontmoeting van een gemanipuleerde context met intelligentie/cognitie. Het is echter niet gemakkelijk te zeggen wat er precies gemanipuleerd wordt met scholing. Het is in ieder geval een ingewikkelde reeks interacterende kenmerken. De correlaties tussen scholing en IQ zijn: $r = .50$ of iets hoger (Ceci, 1991). Boven is de bijna perfecte correlatie tussen IQ en prestaties op *math*, *reading* en *science* in 108 landen vermeld (Lynn & Meisenberg, 2010). Deze waarde is niet representatief en zo hoog vanwege de ongelijke toegang tot onderwijs in Afrika versus Europa en de VS. Door meer scholing kom je in hogere IQ-regio's en neemt het uitbijter effect af.

Scholing en een hoger IQ bevorderen opwaartse sociale mobiliteit. Scholing is gericht op kennen en op het verstand. De school is tegelijkertijd een speciale sociale context waar kinderen en ouders zich niet aan kunnen onttrekken. Ceci (1991) suggereert de volgende factoren, effectieve ingrediënten of mechanismen die met scholing samenhangen en het IQ beïnvloeden:

Combinatie van directe en indirecte instructie.

Verwerven van manieren van leren en kennen en weten wat prestaties op gestandaardiseerde tests waard zijn voor de toekomst.

Duur van de scholing.

Kwaliteit van curriculum en leraren.

Gericht inzetten van perceptuele vaardigheden: leren van figuur-achtergrond onderscheid en hoofd- en bijzaken, abstracte figuren leren, geheugentraining, leren coderen en het onthouden en verbinden van informatie.

Leren van feiten en regels.

Leren abstract te denken, dat wil zeggen iets waarnemen en begrijpen los van eigen directe en persoonlijke ervaring en belang.

Leren van een formele taal, selecteren van belangrijke informatie, denken in hiërarchische categorieën.

Motivatie om te presteren.

Klassengrootte en prestaties Een ander voorbeeld van een te manipuleren sociale contextvariabele die met scholing te maken heeft, is klassengrootte (NICD-Netwerk, 2004). Deze verschilt over landen en door de tijd heen. Voorbeelden:

(1) In een steekproef van 651 klassen in de VS met 10 tot 30 kinderen blijkt dat voor de eersteklassers (groep 6) de variatie wat betreft klassengrootte samenhangt met leesvaardigheden. Kinderen uit klassen met minder dan 20 leerlingen scoren hoger op lezen dan leerlingen uit klassen

met meer dan 20 leerlingen. Het verschil op de SATs is overigens bescheiden. De sociale interacties verschillen. In kleinere klassen is de instructiekwiteit hoger, is er meer emotionele ondersteuning, maar de kinderen zijn minder geëngageerd. Leraren van kleinere klassen noemen hun leerlingen sociaal vaardiger en minder agressief dan hun collega's in de grotere klassen. Leerlingen in grotere klassen doen gemiddeld meer groepsactiviteiten en de kinderen zijn meer betrokken.

(2) De Australische/Nieuw Zeelandse onderwijsonderzoeker John Hattie heeft in zijn boek van 2009: *'Visible learning: a synthesis of over 800 meta-analyses relating to achievement'* meer dan 50.000 studies uit 816 meta-studies samengevat. Eén van zijn vele conclusies: klassengrootte is een *minor factor* voor het prestatieniveau. Hij noemt als top tien variabelen, de *major factors*: Verwachtingen van leerlingen, afwezigheid van storende leerlingen, stimuleren van gepast gedrag in de klas, kwaliteit en motivatie van de leraar, verbeteren van taalbegrip, eerdere prestaties van leerlingen, relatie tussen leraar en leerling, feedback aan de leerling over zijn prestaties en gedrag. Een samenvatting van zijn aanbevelingen staat op het internet: *tgl.pdf: Visible learning, Tomorrow's Schools, the Mindsets that make the difference in education.*

Klassengrootte en prestatie hangen wel samen maar de samenhang wordt in ten opzichte van andere voor de hand liggende factoren als verwachtingen van leerlingen en kwaliteit van de leraren overschat (Hattie, 2009).

Hoge verwachtingen van interventies Naast natuurlijk aanwezige omgevingsverschillen als SES is er de gemanipuleerde omgeving, vooral scholing. Daar wordt veel van verwacht en dat is niet onterecht, zoals indirect blijkt uit de IQ verschillen tussen landen die verschillen in toegang tot goed onderwijs. Binnen geïndustrialiseerde moderne samenleving met goed onderwijs wordt energie gestoken in interventies en programma's om prestaties te verhogen. Het werk van Hattie (2009) vat dat samen. Een studie van Ou (2005) is een voorbeeld van een interventie op jonge leeftijd, want daar wordt het meeste succes verwacht. Het programma was bestemd voor kinderen in de sociaal zwakke buurten van Chicago en werd uitgevoerd door het *Chicago-Child-Parent Center*. Het vroege interventie programma had effect op de gezinnen, school en de kenmerken van de leerlingen, ze gingen bijvoorbeeld vaker naar school. Op 22-jarige leeftijd werd opnieuw gemeten en er werd een effect van de interventie geclaimd: vaker werk, minder criminaliteit, enzovoort.

Je wordt al met al door gepubliceerd onderzoek verleid om te concluderen dat er door de tijd heen grote *vooruitgang* is in IQ, geschiktheden en SATs (Cito-toetsen) en dat er een opwaartse beweging is wat betreft SES. De oorzaken legt men bij de sociale context: scholing, dagopvang voor baby's en kleuters, meer speelgoed dat met school te maken heeft, onderwijstelevisie, verbeterde instructiemethoden, meer tijd voor instructie, testwijs worden van leerlingen, verschuiving naar een autoritaire naar een autoritatieve opvoedingsstijl, meer vriendjes in huis, minder kinderen per gezin en verbeterde gezondheidszorg voor kinderen en hun ouders. Er is echter ook *achteruitgang* bijvoorbeeld met Piagetiaanse taken en die wordt voor een deel verklaard met *dezelfde* factoren die de vooruitgang mogelijk maken.

Voor de relaties tussen sociale context en IQ, geschiktheden en prestaties is geen bevredigend definitief overkoepelend raamwerk bedacht. Een $S \rightarrow R$ schema is de basis en alle invloeden zijn 'goed' als ze maar prestaties verhogen. De nadruk op- en toeneming van cognitieve prestaties maken een samenleving niet noodzakelijk gelijk of sterker en weerbaarder. De Gini Index (zie google) laat zien dat moderne samenlevingen de tendens tonen om de ongelijkheid te vergroten. Dat kan tot sociale instabiliteit leiden: een explosief mengsel van lage opleiding, geen werk, wonen in achterstandsbuurten, kleine criminaliteit, radicalisering, met de rug naar de samenleving staan en ga zo maar door. Meta-studies laten significante effecten van interventies zien en er zijn significante correlaties tussen contextkenmerken en prestaties. Ze zijn echter bescheiden tot gemiddeld (Cohen) en lange termijn effecten zijn niet vaak bepaald.

Door de tijd heen hebben *vooronderstellingen* een rol gespeeld in de discussie over natuurlijke en gemanipuleerde contexten in relatie tot intelligentie, geschiktheid en prestatie. In het westers rationalistisch denken hebben verstand, intelligentie een belangrijke positie verworven. Het wordt belangrijker gevonden dan lichaamskracht. Dat 'goede (en rationele, lijnrechte) verstand' is opgevat als de toegang tot kennis én tot een goede positie in de maatschappij. Lewis Terman selecteerde op de Stanford Universiteit kinderen met de hoogste IQs in de overtuiging dat ze later succesvol zouden worden (Chapman et al., 1991).

Het behaviorisme en de positieve houding van bewoners van de VS leiden verder tot de veronderstelling dat intelligentie, geschiktheid en prestaties beïnvloedbaar zijn in de gewenste richting. Daar komt nog de veronderstelling van het vooruitgangdenken bij: we worden intelligenter, rijker. Vooruitgang wordt met scholing verbonden, en scholing maakt sociale mobiliteit mogelijk. Een nieuwe middenklasse met een gewaardeerde positie en status ontstaat in (ontwikkelings)landen meestal als gevolg van scholing. Intelligentie is een voorspeller van het succes op school. Daarnaast doet de 'geschiedenis zijn werk': het IQ en prestaties moeten door de tijd heen wel stijgen, want alles wordt beter en daar is het allerwegen omarmde Flynn effect een voorbeeld van. Empirisch onderzoek nuanceert deze vooronderstellingen en stelt de verwachtingen doorgaans naar beneden bij.

De diagnosticus kent de betekenis van de sociale context voor IQ, school- en werkprestaties. Hij is op de hoogte van het belang dat cliënten toekennen intelligentie en prestaties. De sociale context kan een bijdrage leveren aan vooruitgang. De diagnosticus moet uitzoeken wat in de concrete sociale context van de cliënt bijdraagt. Uit studies komt immers een beeld naar voren dat 'alles een beetje' helpt: veel verschillende kenmerken van de context tonen bescheiden d- en r-waarden. Het Flynn effect veroorzaakt een praktisch probleem bij testafname want normen moeten regelmatig bijgesteld worden en cliënten moeten hoger scoren voor een zelfde criteriumprestatie.

Samenvatting en conclusie

Bij de verklaring van individuele verschillen in intelligentie en geschiktheid levert de gedragsgenetica een lineair model dat de variantie van genen en omgeving schat en optelt.

De variantie in genen, vastgesteld aan de hand van de verwantschapsgraad, dekt ongeveer 50% van de variantie van het IQ in de populatie. De rest is omgeving-gebonden. Onderzoekers kunnen slechts een deel daarvan traceren op basis van verschillen binnen en tussen gezinnen. Genen interacteren altijd met de omgeving, zonder omgeving komt immers geen gen (genotype) tot uitdrukking in het gedrag (fenotype). Eén gen of een compositie van genen werkt niet in vacuo. Dat betekent dat het zoeken naar het gen dat IQ, persoonlijkheid of stoornissen veroorzaakt onvruchtbaar is. Er zijn er steeds veel en van die vele, zijn er ook weer bij andere gedragskenmerken betrokken: het is een fijne en complexe bedrading of een wirwar. Ook al zijn er soms in de media sensationele berichten (uit de MRI scan kan men afleiden of men CEO kan worden) realistische genetici aarzelen en ontkennen dat. Joseph en Ratner (2010) zijn een project gestart met de veelzeggende titel '*Gene Myths Project*' en Kendler (2005) concludeert na jaren onderzoek dat er geen sterke en directe causale relatie is tussen genen en psychiatrische stoornissen.

In de studie van context, intelligentie en prestatie is er historisch aandacht voor intelligentie als een vermogen van het verstand, het belang van IQ voor vooruitgang in de maatschappij en de beïnvloedbaarheid van IQ en prestatie door scholing en opvoeding. Scholing is gemiddeld tot hoog gecorreleerd met IQ, SES en SAT. De correlatie hangt ook af van de variantie in scholing: als men er ontwikkelingslanden bij betreft is de correlatie zeer hoog, maar in moderne samenlevingen gemiddeld. Omdat men wil beïnvloeden is gezocht naar effectieve bestanddelen in de context voor de toename in intelligentie. De relatie wordt bemiddeld door uiteenlopende ingrediënten, van perceptuele vaardigheden tot het leren van het formele schoolregime.

Het Flynn effect bestaat uit de vaststelling dat er zich per generatie (25-30 jaar) ongeveer één SD vooruitgang in IQ voordoet en wat minder in SATs. Als redenen daarvoor wordt gewezen op verstedelijking, toename van testwijsheid en scholing. Het effect doet zich ook voor bij ontwikkelingstests en wordt toegeschreven aan betere voeding voor en na de geboorte. Er zijn tegenvoorbeelden, bijvoorbeeld bij Piaget taken is er achteruitgang en dat wordt aan ongeveer dezelfde factoren toegeschreven als de vooruitgang. Het is mogelijk dat het takenverschil in variantie-analytische zin interacteert met bepaalde contextkenmerken, bijvoorbeeld de sterke gerichtheid op leren lezen, schrijven en rekenen, tijdens een *back to basics* trend in het basisonderwijs. Het Flynn effect vermindert bestaande verschillen tussen groepen niet. Sommigen interpreteren dit als een uitdrukking van de genetische bepaaldheid van intelligentieverschillen, anderen als de verandering in de genenpool zelf (heterosis) en weer anderen wijzen op de culturele omgeving, die zulke verschillen in stand houdt.

Het *nature-nurture* debat is volgens sommigen achterhaald: een relikwie uit het westers rationalistisch, dualistisch denken. Sommige auteurs hangen een op de biologie gebaseerd denken over het ontstaan van gedragskenmerken door creatieve emergentie aan, dat wil zeggen allerlei interacties en toevalsprocessen leiden tot cognitieve prestaties. Recent onderzoek over effecten van de sociale context laat echter zien dat een $S \rightarrow R$ schema de

basis is voor onderzoek en het debat niet gestopt is. Eenvoudige dichotomieën bevallen ons, want ze zijn helder: het klare en distincte van Descartes.

4. Ontwikkeling x sociale context: Organismische groei tegenover leren

De Piagetiaanse psychologie beschrijft een aantal kwalitatief verschillende stadia, die eenieder doorloopt en tot een gelijk eindniveau brengt. De processen zijn assimilatie en accommodatie. Het laatste doet zich voor bij een conflict waardoor het nodig is om een bestaande structuur aan te passen, zodat nieuwe informatie opgenomen en verwerkt wordt en nieuwe taken opgelost kunnen worden. Een stadium is een tijdelijk rustige periode waarin assimilatie plaatsvindt. Een bron voor het verstoren van het evenwicht vormen leeftijdsgenoten die een ontwikkelingsstadium verder zijn. Ze confronteren een persoon met een ander perspectief op een taak. In de natuurlijke sociale context lossen kinderen immers samen problemen en geschillen op. Dit standpunt heeft onderzoek uitgelokt met kinderen van een verschillend niveau die samen bijvoorbeeld een conservatietaak oplossen. Er was enige versnelling; volgens Piaget niet noemenswaard. Als de cognitieve structuur gewijzigd is moet transfer optreden naar soortgelijke taken. En proefpersonen mogen niet verstoord of onzeker worden als iemand, zelfs een autoriteit, de oude, onjuiste oplossing als de goede presenteert. Aan deze voorwaarden voldoet de verandering meestal niet want daarvoor is zij niet stabiel genoeg. Vooral het idee dat de overgang van de pre- naar de concreet operationele fase versneld kan worden, heeft onderzoek uitgelokt.

Versnellen van de cognitieve ontwikkeling Zou training helpen om sneller van de pre-operationele naar de concreet operationele fase te gaan? Piaget dacht van niet; Amerikaanse onderzoekers van wel. Hun doel was niet zozeer trainingseffecten te bepalen als wel een andere interpretatie voor een stadiumovergang te geven. Ze gebruikten bijvoorbeeld de leer- en informatieverwerkingstheorie. Bij de laatste worden taken in stappen verdeeld en het leren en oefenen van elk van die stappen zou versnelling van de overgang betekenen. Het vraagstuk is niet opgelost in de zin dat het cruciale, beslissende informatieverwerkende experiment is gedaan waardoor conservatie beter te verklaren is. Wel leerde deze aanpak veel over wat een subject doet bij zo'n taak: kijkt het bij de conservatietaak met water naar de lengte en de breedte? Vergelijkt het beide bij elk glas? Maakt het gebruik van de mogelijkheid om de inhoud van de twee glazen in twee andere, gelijke glazen te gieten? Zijn er individuele verschillen? Training op de onderdelen van de taak bevorderde op soortgelijke taken de snelheid van overgaan naar het volgende stadium enigszins. Het kostte tijd voor de antwoorden stabiel werden, ook als de leer-/trainingsfase afgerond was. De interesse voor dit type onderzoek is verdwenen.

Integratie op abstract niveau De organismische en sociale contextbenadering zijn op abstract wijsgerig niveau geïntegreerd door Ennis en Overton (2006). Dit heeft geen gevolg voor onderzoek en verheldert niet veel, want op hoog abstract niveau is alles uiteindelijk één geheel. De mechanistische (nurture) en organismische (nature) opvattingen zijn een

tijdje geconfronteerd waarbij de informatieverwerkingsaanpak de overwinning opeist. Niettemin is er niet aan te ontkomen dat niet alles op elke leeftijd geleerd kan worden, al worden de taken nog zo precies uiteengerafeld in kleine stapjes en nog zo grondig aangeleerd. Integratie op abstract niveau is een filosofische oefening zonder gevolg voor het begrip van intelligentie, cognitie en prestatie. Na 1990 is er minder onderzoek naar de effecten van leren, training en informatieverwerking in stapjes.

De diagnosticus kan de discussie over de sociale context versus ontwikkeling van intelligentie, cognitie prestatie niet naast zich neer leggen. De sociale context van medeleerlingen, leeftijdgenoten, studenten en van collega's op het werk en de materiële context, zoals curricula en werkverdeling en -planning moeten passen bij het cognitief niveau, intelligentie en geschiktheden van de cliënt. Voor het vaststellen van dat niveau is de Piagetiaanse indeling te ruw, zeker voor studenten en volwassenen. Het uiteenleggen van belangrijke taken in noodzakelijke informatiecomponenten zou dan eerder passend zijn, maar dat is tijdrovend en voor veel taken niet uitgewerkt. Het is begrijpelijk dat de diagnosticus terugvalt op de bekende geschikthedentests.

Samenvatting en conclusie

Organismische ontwikkeling veronderstelt een natuurlijke opbouw en een vast verloop. Er is nauwelijks specifieke training nodig want waarnemende- werkendeweg, spelenderwijs doet het kind ervaring op. Verandering wordt beperkt bewerkstelligd door leeftijdgenoten van verschillende ontwikkelingsniveaus die aan dezelfde taken werken. De informatieverwerkingsbenadering legde Piagetiaanse taken in stappen uiteen. Het trainen van deze stappen bevorderde de overgang naar een volgend stadium enigszins, maar het geleerde was niet stabiel (terugval) en werd niet overgedragen naar andere domeinen (beperkte transfer). De confrontatie is voorbij en de integratie op abstract filosofisch niveau heeft geen gevolgen voor onderzoek en diagnostiek.

Voor de diagnosticus telt het ervaringsfeit dat niet alle taken op alle leeftijden en door alle mensen geleerd kunnen worden. Kennis van de samenstelling van de natuurlijke context van leerlingen, studenten en werknemers en van de context van school en werk is voor hem van belang. Deze moet op de een of andere wijze passen bij de ontwikkeling, ervaring en intelligentie van de cliënt. Voor het verwerven van deze kennis is hij op zijn ervaring en kennis en die van zijn collega's aangewezen, want het onderzoek is gefragmenteerd en de resultaten laten beperkte effecten zien.

5. Reflectie en evaluatie

Integratie is het tegenovergestelde van chaos of leegte. In het Grieks en Latijn zijn chaos en leegte woorden voor hetzelfde. Het is moeilijk te aanvaarden dat we leven in een volstrekt toevallige, contingente wereld met willekeurige niet op elkaar betrokken elementen en mensen. Een recente en kale opvatting is dat we leven in reeksen van verbonden informatiesystemen. De strikte cartesiaanse scheiding telt daar niet want 'al wat is' is een systeem en als zodanig ook weer onderdeel van andere systemen. 'Alles in allen' zoals

Multatuli als amateurfilosoof zei. De Amerikaanse dichter, vanaf 1908 levend in Europa, Ezra Pound leed onder de chaos in zijn Canto CXVI: *I can't make it cohere* (Terrell, 1980, 1984). Misschien zocht hij daarom steun bij regimes die later desastreus bleken. Een contingente uitspraak is een epistemische en metafysische bewering die waar kan zijn maar ook niet waar en die niet noodzakelijk waar of vals is. Het klinkt door in de veelgebruikte quote: *it depends*, het is maar hoe je het bekijkt en daar heb je niet veel aan als je wil beschrijven, voorspellen, controleren en beslissen.

Velen beschouwen intelligentie en cognitie als de mentale kracht die ons leert over stabiliteit en invarianties in de fysische (natuurwetten) en in de sociale wereld. Darwin had het ingenieuze idee dat uit amorf materiaal structuren te voorschijn konden komen die niet volgens een strikt plan maar willekeurig leiden tot de soorten die we kennen. Hij gaf later toe dat hij natuurlijke selectie misschien moest aanvullen met wat mensen zelf leren en doorgeven. Contingentie is zo geen vreemde of vijand en kan leiden tot wat wij waarnemen en bedenken in en met behulp van geordende structuren waar ze ook vandaan komen.

Zo beschouwd komt Spearman's bijdrage tegemoet aan een wens. Thurstone maakte er met goede argumenten zeven van en Guilford veel meer maar ze bleven netjes geordend langs de drie ribben in een rechthoekige doos. Het latere werk met psychometrische intelligentie bestaat uit ordeningen door hogere-orde en confirmatieve factoranalyses. Deze weergave van de stand van zaken houdt de betekenis en interpretatie van intelligentie en het meten ervan binnen de perken. Ook informatieverwerking leidde tot het uiteenrafelen in deelprocessen die een plan 'volgen', bijvoorbeeld in modules, zoals Anderson en collega's ze noemen. Deze ordenen de chaotische omgeving van objecten en personen en hebben een materiële basis in verschillende delen van de hersenen. We worden weer eens gered door onze hersenen met keurige modules.

Analyse van oorzaken voor de variantie in algemene (G) of meervoudige intelligenties (bijvoorbeeld PMA) aan de hand van gedragsgenetica, biologische, neurologische, EVPs en EEGs, geeft enig inzicht in de erbij betrokken processen. Het is vooralsnog niet waarschijnlijk dat deze kennis bijdraagt aan het verbeteren van het diagnostisch handwerk van beschrijven, prediceren, controleren en beslissen van het (on)gewenste gedrag van de cliënt. De G, meervoudige factoren en latente trekken kunnen daarnaast gelezen worden als emergente structuren in plaats van als genetische en/of omgeving gestuurde karakteristieken van intelligentie en cognitie. In het intelligentie/cognitie debat doet zich het verschijnsel voor dat de ene theorie de constructen van een andere theorie beschouwt als een uitwerking van de eigen theorie. Dit gebeurde bij het begrijpen van ontwikkeling als homo- en heterotypische stabiliteit. De 'immigrant', i.c. ontwikkeling is een variatie op de bekende cultuur (individuele verschillen). Er is geen 'acculturatie'. En als er zich kwalitatieve ontwikkelingsverandering voordoet kan dat in een psychometrisch model gepast worden. Er blijft natuurlijk intra-individuele variatie, zelfs in IQ en dit type variatie is niet constant in elke levensperiode.

De poging tot integratie van de meervoudige intelligentie en cognitieve ontwikkeling door Demetriou en Efklides leidt tot het behoud van beide: stadia blijven en ze worden per factor

uitgewerkt. Het leidt tot een complex model. De auteurs laten 'acculturatie' zien en vatten hun model als een *Gestalt* op waarvan de som meer is dan de samenstellende theorieën. Het model kreeg weinig navolgers en is nu geïsoleerd. In 2013 vatten Demetriou et al. hun frame nog eens samen en presenteren het als hun *Differential-Developmental-Theory-of-Mind* en ook al wordt het model door de informatietheoreticus Robert Kail (2013) een *powerful integrative theory* genoemd, zal ze waarschijnlijk niet door anderen uitgewerkt en gebruikt worden. Voor de diagnosticus leidt het model tot een profiel opgespannen door zes dimensies met daarop een wisselend aantal Piagetiaanse ontwikkelingsstappen. In beginsel krijgt de diagnosticus een genuanceerd beeld van factoren en hun ontwikkeling. De factoren moeten structureel stabiel blijken over leeftijd en dat is niet onderzocht voor de leeftijden waarop het model past. De ontwikkelingsstappen moeten goed te onderscheiden en cumulatief zijn. Daarover is geen onderzoek bekend. Het aantal is bovendien zo groot dat het moeilijk zal zijn om ze te onderscheiden. De testafname en -scoring zijn complex en tijdrovend. Het is dan ook niet te verwachten dat een diagnosticus dit systeem voor individuele verschillen en ontwikkeling zal gebruiken.

Individuele verschillen, variatie in intelligentie, geschiktheid en prestatie gecombineerd met omgevingsvariatie heeft in de gedragsgenetica geleid tot een half-half verdeling van de varianties over de twee bronnen. De gedragsgenetici hebben de genetische variantie goed in kaart gebracht en geschat. Omgevingstheoretici niet. De sociale context (omgevingsvariantie) blijft in gedragsgenetische studies vrijwel beperkt tot verschillen tussen en binnen gezinnen. Genen kunnen niet zonder een omgeving want het genotype moet zich uitdrukken in een fenotype. Niettemin gaat de discussie minder over die interactie dan over het schatten van het aandeel van de twee variantiebronnen: *nature en nurture*. Daar verbinden we conclusies aan over welke bron er het meest toe doet. Hoe kan een specifiek gen of een combinatie van genen en de erbij betrokken biochemische processen leiden tot bijvoorbeeld een te observeren gedragskenmerk, een psychiatrische stoornis, of de observatie dat iemand bovengemiddeld intelligent is? Gezond verstand doet inzien dat er geen definitieve en unieke verbinding bestaat tussen een gen of genen combinatie en bijvoorbeeld inzicht in experimenteren of een persoonlijkheidsstoornis: Ten eerste dragen honderden genen bij tot variantie in IQ, cognitie, cognitieve stoornissen en zelfs lichaamslengte. Ten tweede wordt de ontwikkeling van een hoog of laag IQ of een stoornis poligenetisch bepaald en spelen genetisch bepaalde risico- en beschermende omgevingsfactoren mee: G x E. Ten derde kan de bijdrage van een gen of genencombinatie niet geïsoleerd worden want er is altijd een werkzame omgeving. De afzonderlijk variantieschattingen zijn mogelijk door gebruik van een lineair model dat factoren schat en optelt. Het is geen representatie van complexe interacties (Moore, 2003). Bevindingen uit de gedragsgenetica gaan bovendien niet over een cliënt maar over populaties. Je kunt niet zeggen: 50% van uw IQ variantie is bepaald door uw genetisch make-up. Uw opvoeding droeg maar 15% bij; spreek je ouders er eens op aan. Bovendien is er geen zicht op hoe de factoren elkaar beïnvloeden.

Er zijn veel omgevingskenmerken die met intelligentie, cognitie, geschiktheid en prestatie samenhangen. Er is gefragmenteerde kennis over de relatie van IQ en cognitie met kenmerken van scholing, SES, aantal leerlingen in de klas, opvoedingsstijlen, motivatie en bekwaamheid van leraren en de kwaliteit van het curriculum. Deze kennis is moeilijk te benutten in het concrete diagnostisch onderzoek van de cliënt want het aantal factoren is (te) groot en de effecten van elke factor afzonderlijk zijn beperkt. Niettemin wordt er veel verwacht van de mogelijkheden om intelligentie, cognitie en persoonlijkheid te verbeteren. Er zijn zelfhulpboeken om intelligentie en persoonlijkheid te verbeteren. Meestal tref je open deuren aan zoals, verhoog uw zelfdiscipline, wees geduldig en emotioneel stabiel. De zelfverbeteringsmarkt had in 2008 in de VS een budget van 10 miljard dollar en hield 30.000 *personal coaches* van de straat en aan het werk (Marketdata Enterprise Inc., 2010). Zo'n markt is er ook voor managen, een activiteit die tot grotere productie zou leiden. Het begon bij Frederick Taylor in het begin van de 20^{ste} eeuw. Hij stond aan de wieg van het Taylorisme dat arbeidsproductiviteit wilde bevorderen door atomistisch en minutieus alle bewegingen van arbeiders aan machines te analyseren en voorstellen te doen om ze in minder tijd te laten verlopen (1911). Denk aan de huidige minuten per cliënt per handeling in de thuiszorg. We treffen een soortgelijke overtuiging aan bij kenmerken van leiderschap (Covey, 2005). In deze lektuur een wordt een stroom aan methoden gepresenteerd die productiviteit zou bevorderen. De methoden wisselen zo snel dat het kortstondige modes zijn. Ik noem een paar van die ronkende beschrijvingen, door sommigen 'jeukwoorden' genoemd.

360-graden feedback, balanced score cards, strategic scenario analyses, brand key performances, analyses of competition, employers satisfaction research, evidence-based management, holistic leadership, lean six sigma en ga zo maar door.

Het Engels maakt het nog indrukwekkender. De stroom is echter incoherent en empirisch onderzoek over effectiviteit is karig. Managers zetten hun grote talenten kennelijk karig in voor empirisch onderzoek.

Heatherton en Weinberger (1994) vormden de redactie van een boek dat onder de vleugels van de APA (de Amerikaanse psychologen vereniging) werd uitgegeven met als titel: *Can Personality Change*. De APA wilde eens op een rij hebben wat er nu verantwoord te zeggen viel over persoonsverandering. Het boek bevat vooral gegevens over stabiliteit. Er is aandacht voor methoden, dat is altijd veilig: zeggen hoe het moet: hoe longitudinale gegevens te analyseren. Bepaling van constantie en groeicurve-analyses worden uitgewerkt, maar er is niet veel over verander- of ontwikkelingsprocessen. Er worden twee vermeld: *quantum change* dat lijkt op de abrupte overgang van een toestand naar de andere uit de chaostheorie en niet-lineaire-dynamica en *crystallization of discontent*, druk op de ketel of 'de kruik gaat zolang te water tot die breekt': een plotselinge overgang of uitbarsting.

Bij intelligentie en cognitie ligt dit anders. Er is en blijft de verwachting dat ze maakbaar zijn, dat wil zeggen, te verbeteren zijn. De correlaties en effectgrootten zijn bescheiden tot gemiddeld en Flynn bracht het goede nieuws dat intelligentie per generatie substantieel

stijgt. Er zijn echter tegenvoorbeelden en het effect is misschien niet zo robuust als gedacht en gehoopt. De houding ten aanzien van wat de omgeving kan bijdragen aan intelligentie en cognitie verschilt en gaat op en neer tussen 'bijna alles ligt genetisch vast' tot 'we kunnen alles tot tand brengen, als we het maar willen'. De diagnosticus neemt een realistisch standpunt in. Er is sprake van omgevingsinvloeden maar ook die zijn niet zonder meer maakbaar. Hij ontkomt er niet aan om bij hulpverlening en advies aan de cliënt binnen smalle marges te opereren.

Naast de aanpak door middel van verrijgings- en compensatieprogramma's om intelligentie en cognitie te verbeteren is er voor leerstoornissen een specifiekere benadering gekomen. Deze wil afwijkingen in executieve functies opsporen, bijvoorbeeld de planning van handelingen, het achtereenvolgens uitvoeren en controleren van die handelingen. De stappen vertonen samenhang met IQ en met de BF, onder meer de planning van de factor Gewetensvolheid.

De integratie van de ontwikkelingsbenadering volgens Piaget en de aard en werking van de natuurlijke en vormgegeven context is niet gelukt. De abstract filosofische integratie van Ennis en Overton heeft geen gevolgen voor de diagnostiek. Ontwikkeling kan iets versneld worden maar niet alles kan op elke leeftijd geleerd en begrepen worden. De confrontatie van deze oriëntaties lijkt voorlopig van de onderzoek agenda te zijn afgevoerd.

De professional moet ook hier zelf uitzoeken of intelligentie en ontwikkelingsniveau van belang zijn voor de vraag of het probleem van de cliënt. Hij ziet in dat de twee bij twee integraties van oriëntaties zelden slagen want het is of wordt (a) een compromis waar elke oriëntatie zijn identiteit behoudt, bijvoorbeeld Demetriou et al. met als gevolg een complex systeem (b) een 'vijandige' overname: individuele verschillen en ontwikkeling vloeien samen in studie van de stabiliteit van IQ en cognitie of (c) er wordt weinig aandacht aan geschonken: ontwikkeling en sociale context.

Ik heb het standpunt ingenomen dat de drie oriëntaties en bronnen niet tot elkaar te herleiden zijn of in een hiërarchie ondergebracht kunnen worden. Als dat het geval is, is het gerechtvaardigd dat de diagnosticus eclecticisch te werk gaat. De oorspronkelijke betekenis van eclecticisch past hier. Het verwijst oorspronkelijk naar een bouwstijl waarin elementen van vroegere architectuurstijlen worden verenigd. De bouwstijlen zijn daarbij metafoor voor de drie oriëntaties. De boodschap is: 'Integreer niet wat niet geïntegreerd kan worden' en 'Begin met psychologische mechanismen' en 'Vorm een rijke probleemdefinitie met behulp van het kennisbestand uit de drie oriëntaties en de drie bronnen'.

VI Operationalisatie, meting en instrumentatie van intelligentie, cognitie, geschiktheid en prestatie

Operationaliseren heeft betrekking op het empirisch specificeren of realiseren van intelligentie- en cognitieconstructen met behulp van taken en vragen: items. Het meten en toetsen van intelligentieverschillen wordt geassocieerd met statistische modellen, item respons functies, data reductie analyses, bijvoorbeeld multivariate technieken en variantie analyse met experimentele en correlatieve ontwerpen voor onderzoek naar samenhangen en verschillen. Ook kun je instrumenten beschouwen als een vrucht van operationaliseren en meten. Dat zijn intelligentie- en geschiktheidstests en prestatietoetsen. Betrouwbaarheids- en validiteitsanalyses bevatten procedures om respectievelijk te bepalen of er betrouwbaar gemeten en voorspeld is. De producten van meten zijn categorieën, bijvoorbeeld ontwikkelingsstadia, observatie- en interviewtopics, rangorde- en intervalschalen, bijvoorbeeld eendimensionele schalen, latente trekken, theoretische attributen en aantallen vectoren of dimensies, bijvoorbeeld multiple intelligenties. Een theoriegeleide constructie van intelligentie en geschiktheidstests maken daarop afgestemde operationalisaties en metingen mogelijk. Bij prestatietoetsen overheerst de moderne testtheorie terwijl de inhoud bestaat uit domeinen van bijvoorbeeld rekenen, taal en wereldoriëntatie en middelbare schoolvakken en vaardigheden.

Meer dan bij studie van de persoonlijkheid zijn theorie, operationalisatie en meten van intelligentie- en cognitiever verschillen op elkaar afgestemd. Het dominante onderzoeksontwerp is het correlatieve, gevolgd door de (quasi)experimentele. Cross-sectionele en longitudinale ontwerpen worden bij de studie van ontwikkeling van intelligentie en cognitie gebruikt. Factoranalyse, correlatie, multiple regressie en t- en F- toetsen zijn resp. statistische analyses en toetsingen. Er is in beginsel een goed gevulde gereedschapskist om (theorie geleid) instrumenten te maken. Na de niet zo succesvolle pogingen van Galton en McKeen Cattell, had Binet meer geluk. Hij bedacht en ontleende aan Binet eenvoudige opdrachten en omschreef het intelligentieniveau als de verhouding tussen chronologische en mentale leeftijd. Spearman gebruikte bestaande taken en items en paste die in zijn model: de G- en S-factoren. De Amerikaanse onderzoekers gebruikten bestaande en ontwierpen taken en items voor multiple intelligentie- en geschiktheidstests: PMA, SOI, en Gf, Gc en uitwerkingen van de twee laatste die dicht bij de PMA komen. De Amerikaanse onderzoekers benoemden de factoren. Thurstone zei dat de G boven de intelligenties zweefde. Spearman beweerde dat niet. Hij aarzelde over de inhoud: Was het wel intelligentie? Zo werd het een abstracte tweedeling van algemeen en specifiek en dat onderscheid past op ongeveer alles. Denk aan spreekbeurten de middelbare school met als aanhef: 'Ik wil het hebben over X (onderwerp) in het algemeen en over Y (een perspectief of deelonderwerp) in het bijzonder'.

Hoe worden intelligentie-, geschiktheid- en prestatiever verschillen concreet gemaakt, dat wil zeggen geoperationaliseerd en hoe gemeten en tot welke instrumenten heeft dat geleid? Wat is de kwaliteit van die instrumenten. Hoe is ontwikkeling geoperationaliseerd en

gemeten? Zijn er meetinstrumenten voor intellectuele en cognitieve ontwikkeling? Hoe wordt de sociale context geoperationaliseerd en gemeten? Zijn er bruikbare instrumenten?

1. Individuele verschillen

Operationalisatie verwijst naar vinden van rake, treffende voorbeelden in de vorm van taken en items die individuele verschillen in intelligentie, geschiktheden en prestaties zichtbaar maken. Meten wordt gerealiseerd met behulp van psychometrische modellen en statistische methoden, zoals IRT, factoranalyse en schaling. Het operationaliseren en meten van ontwikkeling wordt uitgewerkt aan de hand van het vaststellen van stadia bijvoorbeeld door een ontwikkelingsschaal te beschrijven en toetsen. De sociale context wordt geoperationaliseerd door contexten te categoriseren en kenmerken van gezinnen, scholen, buurten en culturen te schalen aan de hand van die kenmerken.

Ten eerste worden operationalisatie en meten van individuele intelligentie- en cognitiever verschillen uitgewerkt bij psychometrische intelligentie: IQ en geschiktheden. De interpretatie van de relatie tussen de items en een alternatieve representatie van de *positive manifold* van items en tests worden uiteengezet. Ten tweede wordt de operationalisatie van prestaties besproken en ten derde worden enkele instrumenten vermeld.

Psychometrische intelligentie en geschiktheid Theorie over individuele verschillen in intelligentie en meten zijn bij Spearman op elkaar afgestemd. Zijn theorie houdt in dat individuele verschillen in ware scores op elke test voor mentaal vermogen toe te schrijven zijn aan twee factoren: de G die gemeenschappelijk is aan alle tests en de S die specifiek is voor groepjes tests. Intelligentie wordt vanuit formele kenmerken van taken gedefinieerd en verwijst tegelijkertijd naar een inhoudelijk kenmerk van cognitief gedrag. Bij Spearman gaat het erom een dominante factor te vinden. De centroïd of zwaartepunt methode van factortrekking (Harman, 1967, p. 171-185) is daarvoor geschikt. Hij is zo niet alleen de vader van de G en S maar ook van een van de grondleggers van het factoranalyseren van de correlatiematrix van items en tests.

Binet was niet bekend met factoranalyse. Hij was vindingrijk in het maken en lenen van vragen en taken die intelligent handelen - een hogere mentale functie - vereist om ze te beantwoorden en uit te voeren. Zijn achtergrond blijkt uit het gebruik van syllogismen dat het toepassen van logische regels vereist. Om te laten zien dat de twee grondleggers van de intelligentiemeting empirisch niet zo ver uit elkaar liggen bij het gebruik van taken wordt gewezen op de relatie tussen syllogismen en de G. Syllogismen bestaan uit twee premissen en een conclusie: alle a's zijn b's, geen a's zijn b, en sommige a's zijn niet b. Het bekendste voorbeeld is van Aristoteles: alle mensen zijn sterfelijk; alle Grieken zijn mensen, dan zijn alle Grieken sterfelijk. Spearman en Binet hebben uiteraard geen kennis van hedendaags onderzoek maar het klassieke syllogisme verbindt de Binet en Spearman in later onderzoek. Voorbeelden:

(1) Shikishima et al. (2011) gebruikten een vijftal syllogismen dat aan 487 paren tweelingen tussen de 13 en 18 jaar werd voorgelegd. Zij vroegen ook de ouders om de syllogismen te beantwoorden. De resultaten leken op kenmerken van de G: het gemiddeld niveau nam toe met leeftijd, de scores

van kinderen en ouders konden significant voorspeld worden uit hun SES en de gedeelde omgevingsvariantie nam af bij de overgang van adolescentie naar volwassenheid. Er was een significante relatie met schoolprestaties. Binet maakte zijn test om precies die te voorspellen. De auteurs zien een relatie tussen syllogistisch redeneren en de G. Binets Franse rationalistische achtergrond en zijn praktische houding leiden tot taken die achteraf correleerden met de G-factor van Spearman. Het syllogisme is een taak waarbij de onderzoeker naar argumenten voor het antwoord kan vragen. Waarom is je antwoord juist?

(2) Khemlani en Johnson-Laird (2012) hebben langdurig het oplossen van syllogismen als een cognitief proces onderzocht. De taken zijn als de items in de IRT en hun denken lijkt op aandacht voor zowel heuristische als formele modellen. De auteurs sommen twaalf minitheorieën op om syllogistisch redeneren te beschrijven. Ze reduceren ze tot drie hoofdtypen:

- (a) heuristische theorieën die intuïtieve regels bevatten
- (b) rationalistisch gecontroleerd redeneren dat logische regels toepast
- (c) set theoretische diagrammen die met gelijkenissen rekening houden.

Khemlani en Johnson-Laird zien gebreken in alle minitheorieën want ze volgen het proces zelf niet op de voet en ze geven daarom geen zicht op hoe een inferentietaak precies uitgevoerd wordt.

Deze studies laten zien dat twee losstaande stromingen over intelligentie empirisch samenhang vertonen. Er is immers een sterke eerste factor waar vele taken op laden.

Thurstone roteerde factoren naar een *simple structure* oplossing. Hij was ervan overtuigd dat er verschillende factoren nodig waren om de prestaties op de 56 tests uit zijn onderzoek te beschrijven. Intelligentieverschillen zijn weer te geven met een aantal onafhankelijke, niet-significant correlerende factoren. *Simple structure* houdt in dat elke afzonderlijke test hoog laadt op één factor en laag op de overige. Dat maakt de interpretatie gemakkelijk want we kunnen dan bijvoorbeeld beweren dat deze test nagenoeg alleen ruimtelijke intelligentie meet en geen of bijna geen of weinig verbale, rekenkundige, enzovoort vermogens. Het SPSS programma drukt bij varimax-rotatie met Kaisers criterium een matrix met correlaties tussen de factoren af. Vaak blijkt dat de factoren samenhangen. Bij de *Big Five* leidde dit tot het AB5C model van Hofstee en De Raad. Als je dit op de zeven PMA factoren zou toepassen zijn er 21: $(n \times (n - 1) / 2)$ combinaties van twee factoren. Deze zouden evenals bij de Big Five tot een genuanceerde interpretatie van intelligentie kunnen leiden. Dat is niet uitgewerkt. Waarschijnlijk is het aantal groot is zijn 21 combinaties te complex voor gebruik in de praktijk. Niettemin zou zo'n exercitie een subtiele interpretatie van intelligentie en cognitie kunnen opleveren. Ik verwacht het niet vanwege het gemak van een profiel dat opgespannen wordt door zeven afzonderlijke scores.

Guilford's SOI bevat drie ribben met daarop kwalitatief verschillende producten, inhouden en operaties. Dat resulteerde in 120 en later 150 factoren. Hij past de analysetechniek aan om zoveel mogelijk ongecorreleerde factoren te vinden. Een van die technieken is de Procrustes methode. Dat is een scheve rotatie en dwingt tests in de factoren die het model voorschrijft. In een zo'n analyse bleek dat 24% van 7082 correlaties *niet* significant was. Als

de correlaties in een frequentieverdeling geplaatst worden leidt dat tot een normaalverdeling met een gemiddelde van $r = .45$. Deze waarden zijn gecorrigeerd voor de steekproeftrekking (meer of minder representatief) en *restriction of range*. Dit resultaat laat zien dat het moeilijk is om zoveel onafhankelijke factoren te onderscheiden in de *positive manifold* van tests. Wat conceptueel onafhankelijk te formuleren is hoeft niet onafhankelijk te zijn in het gedrag van een steekproef proefpersonen.

Vernon en Carroll hebben hiërarchische modellen voorgesteld en pasten hun analyses aan met eerste-, tweede-, n-de orde factoranalyses. Je kunt immers de factorscores van de eerste-orde oplossing gebruiken als invoer voor een tweede-orde analyse. Het aantal factoren wordt dan minder. Zo vonden ze steun voor de G-factor en daaronder de primaire en ten slotte de specifieke factoren. In de *Berlin Intelligenz Structur* (BIS) van Beauducel en Kersting (2002) is gebruikgemaakt van Guttman's facetanalyse en van Horn en Cattell's Gf en Gc om de intelligentie-, geschiktheid- en informatieverwerkingstests en -opdrachten te structureren. De Gf is aanwezig in de informatie *Processing Capacity* en Geheugen en de Gc in *Fluency* en Kennis. De tests zijn afgenomen bij 9520 studenten. Dat is weliswaar een homogene groep maar met voldoende variantie om de volgende resultaten te verkrijgen:

De G-factor correleert .59 met Gf én Gc. Ze delen 35% van de variantie maar er blijft genoeg over voor een afzonderlijke G.

Gf correleert .82 met Geheugen, .48 met *Processing Speed* en .54 met *Processing Capacity*.

Gc correleert .45 met *Processing Speed*, .16 met *Processing Capacity*, .63 met *Fluency* en .21 met Kennis.

Processing Speed maakt geen verschil tussen Gf en Gc maar *Processing Capacity* wel. *Fluency* laadt hoofdzakelijk op de Gc factor. De vijf factoren correleren matig tot gemiddeld met de overige subtests. De BIS kan verschil maken tussen de factoren. Het is echter niet te verwachten dat ze factoranalytisch onafhankelijk zijn en niet significant en rond 0 correleren.

Een metafoer om intelligentie en geschiktheid te beschrijven is door de individuele benadering te beschrijven als water dat stroomt van hoge bergen met kracht (G: *mental power*) naar beneden. Ze maakt en volgt verschillende geulen (Gf, Gc, PMA) en deze worden behoorlijk diep en verzamelen water van kleinere paadjes en zijn diep genoeg om als zelfstandige riviertjes verder te gaan. Een redelijk groot aantal trekt zijn eigen smalle spoor (specifieke en SOI-factoren). Het idee van onafhankelijke factoren (correlaties rond 0) wordt niet ondersteund maar er is voldoende ruimte voor onderscheid. Deze stellen in staat tot een genuanceerde beschrijving van intelligentie die past bij selectie voor verschillende beroepen, functies en gedragscriteria.

Interpretatie van correlaties van intelligentie- en geschiktheidstests Individuele verschillen in intelligentie en geschiktheden worden geoperationaliseerd met items en taken. Deze worden gekozen omdat ze passen in een structuur van capaciteiten. Daarbij wordt

factoranalyse gebruikt om die structuur te toetsen. Het idee van onafhankelijke factoren wordt niet geheel empirisch ondersteund maar je kunt zeggen dat één of een aantal factoren domineert, ook al binden, dekken, verklaren ze niet alle variantie. Voorgaande is een *top-down* redenering die in de praktijk van de testontwikkeling niet vaak voorkomt. Operationalisatie van intelligentie en geschiktheid gebeurt meestal niet theorie geleid. Tests zijn nodig voor praktische doelen en daarbij wordt een *bottom-up* werkwijze gevolgd, denk aan selectie van piloten, specialisten, chauffeurs en leerlingen voor basis/speciaal onderwijs en hbo en universitaire opleidingen.

Het basismateriaal voor analyse van intelligentie is de correlatiematrix van testcores. Deze staat meer interpretaties toe: G en S, PMA, SOI en de hiërarchische lagen van Vernon en Carroll. Er zijn geen criteria om een interpretatie te verwerpen. Als dat wel zo zou zijn dan was dat al lang gebeurd. Wel zijn er nieuwe interpretaties mogelijk. Van der Maas et al. (2006) voegen een interpretatie toe. De *positive manifold* is niet veroorzaakt door een G-factor maar door *dynamische ontwikkelingsprocessen*, dat wil zeggen, positieve, vruchtbare interacties tussen cognitieve processen in de loop van de ontwikkeling. Dit zou ook de hiërarchische natuur verklaren want gedurende de ontwikkeling is er opeenstapeling en opeenvolging van (wijzigingen in) waarnemen, ervaren en vaardigheden verwerven. De erkenning van een opbouw gedurende de ontwikkeling is een aanvulling op de individuele verschillen benadering. Het blijft een interpretatie en die volgt niet dwingend uit de correlatiematrix. Het is niet mogelijk om tussen het voorstel van Van der Maas et al. en andere modellen te kiezen zeggen de statistici Bartholomew et al. (2009) dan ook. In de *framing* van dit tekst wordt een ontwikkelingsinterpretatie verwelkomd. Die schenkt immers aandacht aan het dynamische karakter van intelligentie zoals dat gedurende de ontogenetische ontwikkeling wordt opgebouwd.

Elementaire processen Bij het zoeken naar de oorzaak van individuele verschillen werd gedacht dat je – precies als in de natuurkunde - de kleinste eenheden, de basis ingrediënten moest identificeren: Elementaire Cognitieve Taken, ReactieTijden en EEGs en EVPs. Ze worden in de toekomst mogelijk aangevuld met hersenscans en DNA-kenmerken. Het idee is dat als we die kennen de rest vanzelf volgt. IQ, cognitie, intelligent handelen zijn samenstellingen van deze eenheden. Dat is nog niet gelukt en als ze ‘gevonden’ worden zullen ze de bekende IQ tests niet verdringen. De correlaties van de eenheden en IQ variëren immers van bescheiden tot gemiddeld. Ze voorspellen bovendien criteria niet beter. Het interpreteren en bewerken van de output van EEGs, EVP en MRI hersenscans is een vak op zich. Hoe moet je uit de aangebrachte kleurverschillen, vormen, patronen in een hersengebied interpreteren? Als iemand een complexe activiteit uitvoert hoe kunnen we dan zien dat het verschilt van een eenvoudige taak?

Het vraagstuk van de kleinste eenheden komt elders voor. Deeltjesfysici zijn er naar op zoek en evolutiebiologen zoeken naar de eenheid van selectie: de samenleving, de groep, de soort, het individu, het gen. Daarbij krijgen de genen de meeste aandacht. Deze overleven en ‘gebruiken’ daar mensen voor die steeds in koppels de helft van hun genen doorgeven

aan een nieuw individu. Dit is Dawkins' (1976) verhaal over *The selfish gene*. Het heet ook wel het 50% - 50% probleem in de evolutiebiologie want wat is het evolutionaire voordeel van deze verspilling? Of is het een verkeerde vraag: misschien is verspilling een adaptief verschijnsel: psychologische gezien is het uitdrukkelijk nodig niet alles te behouden van wat waargenomen, geleerd, gevoeld en bedacht wordt. Een menselijke kenner is geen container waarin alles verzameld wordt en waar hij een deksel op doet. Het kiezen en vinden van de ware, geschikte, goede eenheid is dus geen vraagstuk dat alleen in de psychologie voorkomt.

Bij de analyse van intelligentie is onderzoek naar processen van *informatieverwerking* een aanvulling. Het gaat om het beschrijven van het aantal stappen dat noodzakelijk is voor het oplossen van een vraagstuk, bijvoorbeeld het doen van een adequate waarneming of het trekken van een juiste conclusie. Het is niet gemakkelijk gebleken die processen te isoleren en te meten. Bovendien moet aangetoond worden of de stappen zich empirisch feitelijk, en in de beschreven volgorde voordoen. Daarmee kent deze benadering dezelfde problemen als het maken en toetsen van ontwikkelingsschalen. Is het bijvoorbeeld slechts een 'logische' volgorde of is er natuurlijke volgorde die empirisch blijkt? En zijn de processen wel voldoende precies beschreven om bruikbaar te zijn voor de diagnose van de intelligentie van een cliënt? De benadering pretendeert operationalisaties en metingen te leveren voor fundamentele cognitieve processen. Sommige van de processen zijn onderdeel van de bekende tests geworden (WAIS en WISC). Sternbergs poging om een test te maken gebaseerd op informatieverwerking bleek een lastige opgave en leverde geen winst voor het voorspellen van criteria. Onderzoek naar informatieverwerking gaat door maar biedt vooralsnog geen betere of bruikbare tests.

De operationalisaties van uitvoerende processen (*Executive Functions: EFs*) spelen al een rol bij analyse en voorspellen van basisvaardigheden als rekenen en taal en leerstoornissen. Deze functies worden opgevat als fundamentele cognitieve processen die denken en handelen controleren. Men verbindt ze graag aan neurale substraten in de prefrontale cortex en ze worden gebruikt om gebrek aan controle te verklaren als gevolg van hersenletsel. Men onderscheidt drie functies:

EF (1) *updating* ligt dicht bij het werkgeheugen en verwijst naar het monitoren van inkomende informatie en de revisie van de werkgeheugen-inhoud door oude informatie te vervangen door nieuwe en relevante te kiezen met het oog op een taak.

EF (2) *shifting* is het proces van switchen tussen taken en mentale sets. Als condities veranderen zijn andere regels en oplossingen geschikt. Je moet een eerdere mentale set los kunnen laten als die irrelevant is geworden.

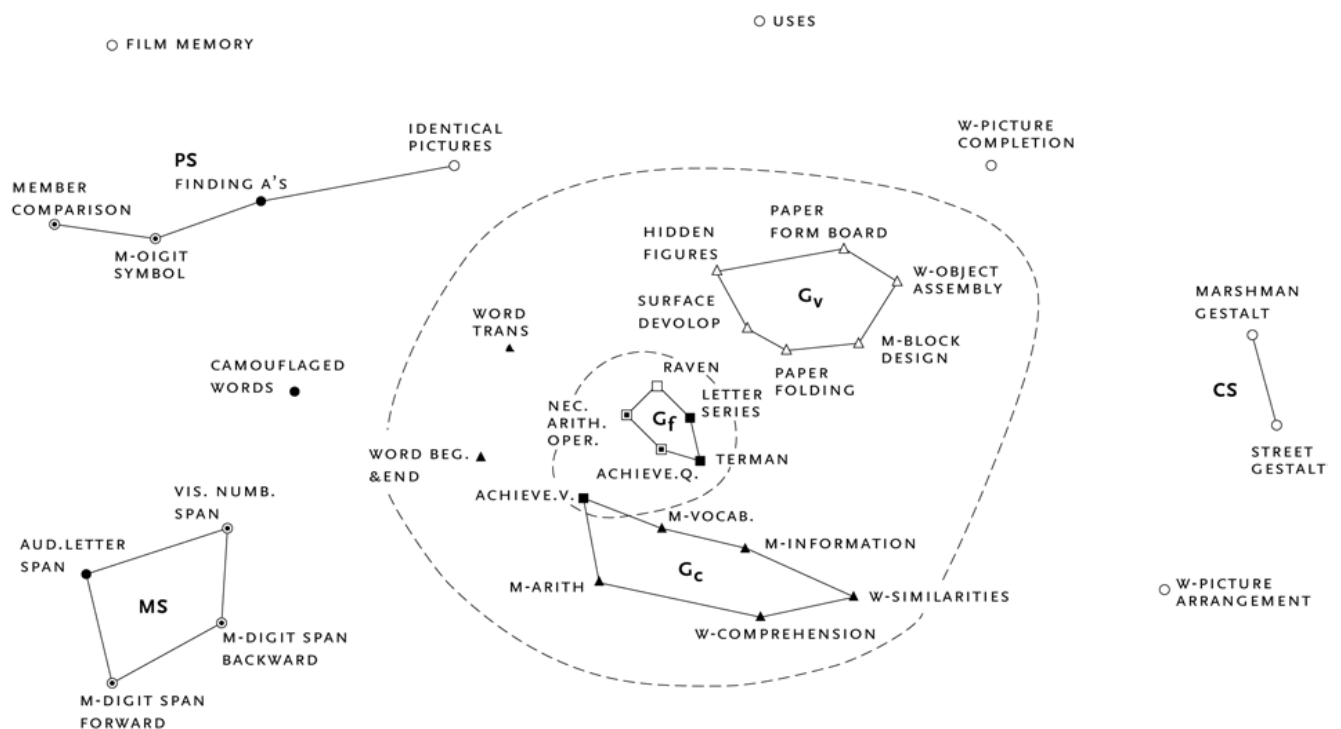
EF (3) *inhibition* zorgt voor het onderdrukken van dominante maar irrelevante respons tendenties, zoals bij het uitvoeren van de Stroop woord-kleur taak nodig is.

Een alternatieve representatie De *positive manifolds* berusten op ruwe testcores. Guttman bedacht een andere maat voor verwantschap tussen testcores. Deze is door Maraun (1997)

al gebruikt om de BF in een plat vlak af te beelden. De verwantschap wordt uitgedrukt als een afstand (Schlesinger & Gutmann, 1969):

$d_{ij} = \sqrt{1 - r_{ij}}$, waarbij d_{ij} de afstand is tussen test i en test j . De rij's staan voor de $n(n - 1) / 2$ correlaties tussen de n -tests.

De afstand tussen test i en j wordt weergegeven met een getal dat het resultaat is van de wortel uit $1 -$ de correlatie tussen test i en test j . Als test i en j hoog correleren liggen ze dicht bij elkaar in het platte vlak. In dat vlak worden alle paren $[n(n - 1)/2]$ van correlaties door afstanden weergegeven. Meestal levert dat na een multidimensionele schaalanalyse een reeks min of meer cirkelvormige verzamelingen van tests op (Figuur 1).



Figuur 1: Een aantal cognitieve tests is afgebeeld met de correlaties tussen de tests als index voor hun afstanden. Deze afbeelding berust op gegevens van 241 middelbare school leerlingen. De factoren zijn dik gedrukt: G-fluid, G-culture, G-visual en Memory Speed, Perceptual Speed en Cognitive Speed (Marshalek et al., 1983): ‘The complexity continuum in the radix and hierarchical models of intelligence’. In: *Intelligence*, 7, 107-127. De Gf in het midden is de Raven: een G-verzadigde test, de Culture en Visual G liggen in de eerste cirkel en de andere drie: MS, PS en CS (speed factoren) aan de rand.

De algemene G-tests liggen in het midden dicht bij elkaar terwijl tests met dezelfde inhoud (verbaal, numeriek, ruimtelijk) in sectoren van de cirkel terechtkomen. De complexe tests tonen de hoogste gemiddelde correlatie met de andere. Wegwaaiend uit het midden liggen tests die minder algemeen zijn en gemiddeld lagere correlaties laten zien. Nabijheid tot het midden wijst ook op algemeenheid (de G): de Raven (een cultuurarme, hoog G-ladende test) ligt precies in het midden. Guttman heeft dit model al in de jaren zestig en zeventig van de

20^{ste} eeuw ontwikkeld. Er wordt om onduidelijke redenen weinig gebruik van gemaakt. Guttmans werk was in de VS in niet populair. Meestal worden relaties tussen items en (sub)tests weergegeven met factoren of schalen ontstaan uit factoranalyses op correlatiematrixen en met structurele vergelijkings- en schalingstechnieken en Rasch-modellen.

Alternatieve theorieën hebben tot instrumenten geleid voor competenties, emotionele en praktische intelligentie. Deze worden af en toe gebruikt. De constructvaliditeit van de instrumenten vraagt nader onderzoek en de extra variantie die ze in criteriumgedrag verklaren, ontbreekt of is gering.

Metten van prestatie Operationalisatie van school- en werkprestaties wordt aan inhoudelijke specialisten overgelaten. De items worden in eerste aanleg niet geschreven om een theoretisch veronderstelde structuur te toetsen. Het gaat om inhoud en die wordt bepaald door curriculumontwerpers en deskundigen die onderwijsdoelstellingen voor leerlingen of *targets* voor werknemers formuleren. Het onderscheid tussen niveaus van doelstellingen heeft soms het karakter van een ontwikkelingsschaal. Bloom (1956) heeft een opbouw van onderwijsdoelstellingen beschreven die gaat van eenvoudig 'weet hebben van' tot 'zelfstandig beoordelen en evalueren van kwaliteit van kennis volgens interne en externe zelfbedachte criteria'. Empirische studie naar de ordening van gedragingen bij deze doelen is overigens niet verricht. De ordening wordt voorondersteld en is dwingend, is 'logisch'. Dat kan tot gevolg hebben dat empirisch onderzoek niet nodig wordt geacht.

Daarnaast kan taakanalyse van bijvoorbeeld natuurkunde- en taalvraagstukken tot het inzicht leiden dat er verschillende intelligentiefactoren of geschiktheden betrokken zijn bij het oplossen van de vraagstukken. Dit komt aan de orde in de informatieverwerkingsanalyse van dit type opdrachten maar speelt geen rol in het operationaliseren en meten van school- en werkprestaties.

De constructie van prestatietoetsen staat sinds de oprichting van het Cito onder het regime van de IRT. Dat betekent dat de items bepaalde kenmerken moeten vertonen. Ze moeten passen in een model. In de klassieke testtheorie worden eveneens itemkenmerken bepaald:

p-waarden: de moeilijkheidsgraad van de items, dat wil zeggen het percentage deelnemers dat een item goed beantwoordt: niet te veel en niet te weinig anders heeft het item nauwelijks variantie en dus weinig covariantie: correlatie met andere items.

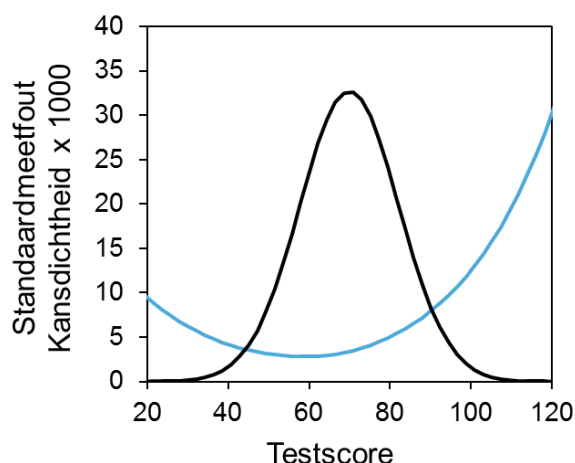
d-waarden: de differentiatie die items aanbrengen tussen goede en slechte presteerders: als een item even goed door goede als slechte leerlingen gemaakt wordt is het item niet bruikbaar (deze niet verwarren met de Cohens d-waarden: de index voor de waarde van verschillen tussen gemiddelden in een experiment).

item-test correlaties: als een item niet met de andere correleert, slaat dat item kennelijk op iets anders als op de prestatie op onderwijs doel Y.

betrouwbaarheid van de test: meestal Cronbachs alfa (maar waarom eigenlijk?), de herhaalbaarheid door de tijd heen: stabiliteit en de integratie: generaliseerbaarheid.

De moderne testtheorie wordt gebruikt door de gespecialiseerde instellingen die toetsen ontwerpen voor alle soorten onderwijs. In de VS is dat de *Educational Testing Service* in Princeton en in Nederland het Cito in Arnhem. Deze instituten zijn voor het grootste deel bevolkt door inhoudelijke toetsontwerpers, de 'itembakkers' die voor een klus ingehuurd worden en door een vaste afdeling Psychometrie die op het Cito met zelfontwikkelde varianten van het OPLM-model werkt. De geteste heeft een positie op een latente schaal en die bepaalt de kans dat hij een item goed beantwoordt. Het item heeft een p-waarde (moeilijkheidsgraad) en discriminatiewaarde (het vermogen om verschil te maken tussen personen met een bekende plaats op de latente schaal). De constructie van de latente schaal gebeurt op basis van waarden in een steekproef na iteratieve procedures die je bijna niet met de hand kunt uitvoeren of narekenen. Als een item onderscheid kan maken tussen personen die dicht bij elkaar liggen op de latente schaal dan discrimineert het goed. Het is moeilijker om items te schrijven die verschil maken tussen proefpersonen met een IQ van 125 en van 130 dan met een IQ van 100 en van 130 en van 100 en van 105. Dit onderscheidend vermogen is te zien aan de steilheid van de item-responsfunctie. Als bijvoorbeeld een persoon met een latente schaalscore van 5, 55% kans heeft om een item goed te beantwoorden en een persoon met een score van 4, 80%, dan onderscheidt dat item beter dan één waarbij een persoon met 5, 55% kans heeft om het item goed te beantwoorden en één met een score van 4, 60%.

Met behulp van IRT kun je de lokale betrouwbaarheid schatten. Aan de uiteinden is een test vaak minder nauwkeurig dan in het midden en daar vlak onder en boven. Stel dat de populatieverdeling voor een bepaalde vaardigheidsscore er zo uit ziet als in Figuur 2 dan kun je op basis van IRT de lokale meetfout schatten (voor elke waarde afzonderlijk dus). In de figuur is dat de blauwe curve.



Figuur 2 drukt uit dat deze test het nauwkeurigst meet in de lagere en midden vaardigheidsregio. Hoogbegaafden zijn volgens deze curve minder nauwkeuring te meten. Dit is niet puur hypothetisch. Het is moeilijker goede items te schrijven voor de uiteinden van de schaal.

Eerder heb ik vermeld dat de predictieve validiteitcoëfficiënten van klassiek en modern testtheoretisch geconstrueerde tests nauwelijks verschillen. De IRT is gebaseerd op een meetmodel en de KTT niet. De eerste leidt tot unidimensionele schalen. Er is geen discussie en onderzoek over het representeren van complexe prestatiedomeinen. Als daar al aan gedacht wordt, is de oplossing meestal gebruik van *verschillende* tests voor ieder domein één. Elk domein moet voldoen aan IRT vereisten. Het aantal domeinen heeft betrekking op het aantal onderscheiden vaardigheden waarvoor itemschrijvers afzonderlijke tests maken. De IRT is dan ook niet behulpzaam of constructief voor het vormen van een beeld dat bijvoorbeeld een remedial teacher wil hebben over welke vaardigheden aanwezig zijn en welke ontbreken bij een leerling.

Ondanks uiteenlopende opvattingen en bijbehorende tests voor intelligentie en geschiktheden en het ontbreken van consensus over de structuur, is de keuze van de diagnosticus overzichtelijk. Het gaat om het meten van de G-factor of van een reeks Primaire factoren of nog Specifiekere factoren. Zijn keuze hangt af van hoe hij en zijn cliënt de vraag formuleren. Er is geen juiste of dwingende interpretatie van de correlaties tussen intelligentie(sub)tests. Een algemene test als de Raven en wat minder de Wechsler geven een beeld van de G. Deze factor correleert met bijna alle prestaties. Voor een specifieke vraag, zoals beroepskeuze of toelating tot een gespecialiseerde opleiding, is een gedifferentieerde test zinvol, bijvoorbeeld de DAT: Differentiële Aanleg Test. Geschiktheden en Primaire capaciteiten zijn niet geheel onafhankelijk. Prestaties in onderwijs en beroep kunnen zorgvuldig met op IRT gebaseerde toetsen vastgesteld worden maar zo'n score geeft geen informatie over wát ontbreekt als een prestatie onvoldoende is.

Instrumentatie In het ideale geval is er theorie, een verzameling ontwerpen en een aantal statistische modellen die passen bij de theorie. Daaruit vloeien theoriegeleide, betrouwbare en valide meetinstrumenten voort. Deze drie-eenheid informeert je bovendien over het voorspellen van interessante (criterium)gedragingen. De werkelijkheid is anders. De meest gebruikte instrumenten zijn IQ tests zoals de G-verzadigde Raven en de Wechsler schalen. De handleidingen bevatten informatie over juist gebruik, betrouwbaarheid en validiteit. De tests zijn bewerkt in veel landen en dat maakt vergelijking mogelijk, hoewel daar haken en ogen aanzitten. Betekenen items hetzelfde in verschillende landen? De instrumenten worden steeds aangepast en opnieuw onderzocht. Voorbeelden:

(1) Flynn en Rossi-Case (2010) toonden dat de sekseverschillen op de *Raven* bij 14- tot 18-jarigen klein zijn. Op de leeftijd tussen 18 en 24 jaar scoren mannen iets hoger. De boodschap is steeds dat vrouwen hoger scoren op verbale taken en mannen op reken- en ruimtelijke taken. De gegevens lieten in vijf hoogontwikkelde, geïndustrialiseerde landen zien dat er gemiddeld geen sekseverschillen waren bij 15 tot 18-jarigen. Er waren wel meer mannelijke schoolverlaters.

(2) De *WAIS* wordt steeds 'ge-updatet' en onderzocht, ook in Nederland (Barelds et al., 2013). De handleiding van de Amerikaanse volwassenen versie is in 2008 gepubliceerd. Men kan het G-karakter van de *WAIS* benadrukken want er is een correlatie van rond de .70 tussen de twee hoofdfactoren: Verbale en Performale intelligentie. Niettemin is het ook mogelijk om vier eerste

orde factoren te gebruiken. Deze worden geïnterpreteerd vanuit een domein- en informatieverwerkingskader: I *Verbal Comprehension* II *Perceptual Reasoning* III *Working Memory* en IV *Processing Speed* (Wechsler, 2003, 2008).

(3) Watkins (2010) vond een eerste sterke WAIS factor die 48% van de variantie voor zijn rekening nam van de totale gemeenschappelijke variantie van 75%. Een scheve factoroplossing helpt echter om de prestaties genuanceerder te interpreteren. Dit is ook het geval bij de Big Five: het AB5C model. Canivez en Watkins'analyse (2010) liet opnieuw een sterke eerste factor zien in een andere steekproef en noemen de Wechsler een goede schatter van de G van Spearman.

Sekseverschillen in IQ blijven populair. Ze zijn echter gering. Er is meestal een sterke eerste factor bij intelligentietests die wel de G wordt gelezen, maar differentiatie is mogelijk. Bij de WAIS en WISC wordt dat uitgelegd als processen van informatieverwerking.

Geschiktheidstests steunen voor hun interpretatie op Thurstones PMA. In de VS zijn veel geschiktheidsinstrumenten beschikbaar en een groot deel daarvan is voor Nederland bewerkt. Het meten van school en werkprestaties staat voor een deel onder auspiciën van de staat (SATs van het ETS in Princeton en in Nederland de Cito-toetsen) en van grote bedrijven.

In Nederland zijn veel prestatietests beschikbaar. Er zijn meer dan 500 instrumenten beschreven en geëvalueerd. De meeste hebben betrekking op het vaststellen van individuele verschillen in persoonlijkheid, intelligentie, geschiktheden, prestaties (Cito-toetsen) en sociale contexten (gezin, school). Diagnostici hebben baat bij *maximum performance* tests die

eenvoudig af te nemen en te scoren zijn.

relevante criteria voorspellen.

gebaseerd zijn op een theorie (kader, conceptueel schema) van intelligentie, geschiktheid en prestatie.

niet te specifiek en niet te algemeen van inhoud zijn.

duidelijk zijn voor de tester en de geteste.

vrij zijn van culturele vertekening, van itempartijdigheid.

empirisch getoetst zijn: de inferenties over het construct, de inhoud en de voorspelling van criteria zijn gebaseerd op onderzoek.

Hieronder worden enkele in Nederland beschikbare *maximum performance* tests vermeld. De recentst beschikbare informatie over de instrumenten is te verkrijgen door de 'Cotan Aanvullingen' op de documentatie van Evers et al. (2000) te raadplegen. Ze zijn als men zich abonneert online te raadplegen: www.cotandocumentatie.nl.

Intelligentietests De *Binet-Simon* is de eerste intelligentietest in Nederland. Hij is bewerkt door Herderschêe en heeft nu antiquarische waarde. De Amerikaanse (Stanford-Binet) en de Duitse bewerkingen (Binet-Simon-Boberitag) zijn een tijdlang gebruikt. Deze zijn nu geschiedenis, onder meer omdat onderzoek heeft laten zien dat de items in de Nederlandse

bewerking niet de vereiste opklimming in moeilijkheidsgraad vertonen. Normen ontbreken en Betrouwbaarheid en Validiteit zijn onvoldoende zijn (Visser et al., 1982). Er zijn nu geschiktere instrumenten zoals de WISC-R, bewerkt door een projectgroep bestaande uit Van Haassen et al. (1986) en de RAKIT (Gereviseerde Amsterdamse Kinder Intelligentie Test van 4-12 jaar (Bleichrodt et al., 1984, 1987) en de recente bewerking (zie Cotan Aanvullingen).

De *Wechsler Intelligence Scale for Children (WISC-II)* is bestemd voor 6- tot 16-jarigen en is rond 1985 voor Nederland opnieuw bewerkt. Normen, Betrouwbaarheid en Begripsvaliditeit worden in de testdocumentatie (zie Evers et al., 2000, p. 475; en p. 1138-1160) als voldoende beoordeeld. Er was tot dan nauwelijks onderzoek naar Criteriumvaliditeit verricht. Daarna is er even een derde versie geweest van het opgeheven Dienstencentrum van het Nederlands Instituut voor Psychologen (2002). Over deze versie is discussie ontstaan en de beoordeling in maart 2003 viel ongunstig uit, zodat de verkoop gestopt is. Bij testontwikkeling is haastige spoed zelden goed. Productontwikkeling is arbeidsintensief. De beoordeling van de 'snelle' WISC-III NL (Kort et al., 2002) luidde dat op goede Uitgangspunten en de Kwaliteit van het Testmateriaal na, alle onderdelen onvoldoende waren (Cotan Aanvulling 2004/1). Er is een nieuwe poging ondernomen tot een instrument dat wel voldoet. Het instrument is beschikbaar sinds 2010 (zie Cotan Aanvullingen).

De *Gereviseerde Amsterdamse Kinder Intelligentie Test (RAKIT)* bevat twaalf subtests: figuur herkennen, exclusie, geheugenspan, woordbetekenis, doelhoven, analogieën, kwantiteit, schijven (pennenbord met drie lagen van negen schijven), namen leren, verborgen figuren, ideeënproductie en de opdracht een verhaal te vertellen naar aanleiding van een plaat met een scène: de vertelplaat. De test kan bij kinderen tussen de 4 en 12 jaar afgenomen worden. Er is veel onderzoek met dit instrument verricht en het voldoet met de kwalificatie 'goed' aan de criteria met betrekking tot theoretische achtergrond, Vorm en Kwaliteit van materiaal en de Handleiding. Aan de psychometrische criteria met betrekking tot Normen, Betrouwbaarheid en Begrips- en Criteriumvaliditeit is voldaan (Evers et al., 1992, pp. 532-538; Evers et al., 2000, pp. 472-473; 1129-1135). De beoordeling is van 1990. Deze test is bewerkt en geëxporteerd naar andere landen (Indonesië en India). Er is een grondige bewerking van de RAKIT in aantocht of inmiddels beschikbaar.

De *Intelligentietest Eindtoets Basisonderwijs* (Van Boxtel et al., 2011) van het Cito is een jonge aanwinst in het testen van intelligentie bij een specifieke steekproef: groep 7 en 8 van het basisonderwijs. De test is naast de Cito-eindtoets basisonderwijs geplaatst om na te gaan of leerlingen presteren in overeenstemming met hun IQ. Er zijn onderpresteerders, bijvoorbeeld door gebrek aan motivatie of gezinsomstandigheden en overpresteerders die ondersteund en gestimuleerd zijn. De correlatie tussen de twee is hoog rond de $r = .70$, maar er zijn altijd leerlingen die buiten de regressielijn vallen. De test wordt steeds geëquivalleerd, dat wil zeggen erop gecontroleerd of ze hetzelfde meet in opeenvolgende

schooljaren. Het instrument is afgeleid van de Intelligentie Schoolvorderingen Interesse (ISI) test van Van Boxtel et al. (1982) en test bevat Rekenen I en II, Zuiver schrijven I en II, Begrijpende lezen, Synoniemen, Tegenstellingen, Soortbegrip woorden, Geknipte Figuren, Draaien en Soortbegrip Figuren. Er zijn drie verbale en drie ruimtelijk-figurale subtests. Het doel is om de keuze voor vervolgonderwijs te ondersteunen. De test kan door de groepsleraar afgenomen worden. De opbrengst is een IQ score met $M = 100$ en $SD = 15$. Zes subtests leveren twee factoren op (Synoniemen, Tegenstellingen, Soortbegrip Woorden: Verbale Intelligentie (V) en Geknipte Figuren, Draaien en Soortbegrip Figuren: Ruimtelijke Intelligentie (R)). De varimax rotatie laat zien dat er geen nul-ladingen zijn op de andere factor: hetzij de V of de R factor; bijvoorbeeld Soortbegrip Figuren heeft een lading van .57 op de V en .49 op de R factor. Normering is gebaseerd op een representatieve steekproef, verdeeld over de vier regio's in Nederland (N, O, W en Z) en *gewogen* op grond van de vertegenwoordiging in de populatie, bijvoorbeeld een gewicht van 1.00 voor 'het gemiddelde Nederlandse kind', 1.25 voor 'arbeiderskinderen', 1.40 voor schippers- en internaat kinderen en 1.90 voor kinderen uit gezinnen waarvan minstens een ouder van niet-Nederlandse afkomst is. Het Cito-IRT model is gebruikt om de itemkenmerken te bepalen. De interne consistentie van de eerste 10 items voor de zes subtests varieert na correctie voor testlengte tussen .64 en .90. De *greatest lower bound* (GBL) wordt opgevat als een betere schatting en is iets hoger: vanaf .78. Betrouwbaarheid en Validiteit zijn goed: er zijn gedetailleerde gegevens die dit ondersteunen. Deze test is wat betreft psychometrische aanpak en rapportage een voorbeeld voor het handwerk van testconstructie.

De *Nederlandse Intelligentietest voor Onderwijsniveau* (NIO, 2004) is een recent bewerkt instrument voor algemene intelligentie voor groep 8 van het basis- en klas 1, 2 en 3 van het voortgezet onderwijs. Het is de opvolger van twee eerdere Groningse instrumenten (GIVO en GALOm zie Visser et al., 1982). De test kan individueel en groepsgewijs worden afgenomen en handmatig of met de computer gescoord worden. De drie eerste componenten (varimax rotatie) verklaarden 61%, 13% en 10% variantie. Er is dus een sterke eerste factor. De twee-factoroplossing laat het verschil zien tussen drie verbale en drie symbolische tests. Het onderscheid blijft over de leerjaren behouden. De test krijgt de kwalificatie goed voor alle rubrieken (Cotanbeoordeling, 2010). Er is veel onderzoek mee verricht. Er zijn substantiële correlaties met de WISC en schoolprestaties gevonden. Bij het onderzoek zijn verschillende populaties betrokken (Van Dijk, 2008a, 2008b; Tellegen & Van Dijk, 2009).

Raven Progressive Matrices (Raven, 1938; coloured versie: CPM, 1947) is een G-verzadigde 'cultuurarme' test. Hij test kan gebruikt worden vanaf 6 jaar en groepsgewijs en individueel worden afgenomen. Er zijn vijf reeksen van twaalf in moeilijkheidsgraad opklimmende opgaven die elk bestaan uit één of meer figuren. De proefpersoon moet ontdekken met welk deel elk figuur aangevuld moet worden door de systematiek in de reeks afbeeldingen

te ontdekken. Het goede antwoord moet gekozen worden uit zes of acht afbeeldingen. De totaalscore wordt verkregen door het aantal goede antwoorden op te tellen. Het instrument is in het verleden in Nederland bij keuring voor militaire dienst gebruikt. Hoewel er veel Nederlandse gegevens over zijn, ontbreekt informatie over Normen. De overige rubrieken worden voldoende/goed beoordeeld (Cotanbeoordeling, 1991) (Evers et al., 1992; Evers et al., 2000). Er is na 2000 geen Nederlands onderzoek gerapporteerd. Daarvoor wel, voorbeelden: Van Bon (1986) heeft normen voor de CPM gemaakt die nu verouderd zijn. De eerste factor verklaarde - toenemend met leeftijd - tussen de 17% en 28% variantie. Dat is beperkt, gegeven het idee van een sterke G-factor. Verouden et al. (1987) gebruikte de CPM bij 148 autochtone en 68 in Nederland geboren en 299 het buitenland geboren allochtonen leerlingen. De buitenlandse allochtonen scoorden op de CPM gemiddeld 0.25 lager; voor Verbale Analogieën was het verschil 0.60 en voor Figuur Sorteren ook 0.25. Diesfeldt en Vink (1989) gebruikten de CPM bij 65 gezonde ouderen (69-92 jaar) en vonden een relatie tussen leeftijd en scores van $r = -.27$ ($p = .03$). De CPM is in 1991 beoordeeld. De Normen zijn onvoldoende, de overige rubrieken: goed/voldoende.

De *Wechsler Adult Intelligence Scale (WAIS)* is een veel gebruikt instrument. De ontwerper had zich tot taak gesteld een algemeen en handzaam instrument te maken dat veel informatie opleverde in korte tijd. Er is al een Nederlandse Belgische bewerking van de WAIS in 1970 (Stinissen et al., 1970). Personen tussen de 15 en 64 jaar kunnen ermee onderzocht worden. Er zijn zes verbale en vijf performale subtests. Deze worden vermeld, omdat soortgelijke onderdelen in veel andere intelligentietests terugkeren. Het zijn de verbale (1-6) en performale (1-5) subtests van de Wechsler-schalen:

- 1 Informatie: kennis van algemene dagelijkse zaken; wat algemene ontwikkeling genoemd wordt.
 - 2 Begrijpen: begrip van alledaagse levenssituaties.
 - 3 Rekenen: redactiesommen.
 - 4 Overeenkomsten: er worden twee woorden genoemd en de overeenkomst moet genoemd worden.
 - 5 Cijferreeksen nazeggen.
 - 6 Woordenschat: woordbetekenissen geven.
-
- 1 Substitutie: in een beperkte tijd moeten bij zoveel mogelijk tekens cijfers gezet worden.
 - 2 Onvolledige tekeningen: ontbrekende kleine onderdelen in een tekening moeten ontdekt worden.
 - 3 Blokpatronen: kubusvormige blokjes, rood en wit en gecombineerd rood/wit dienen om patronen mee te leggen die op een voorbeeld op kleinere schaal weergegeven zijn.
 - 4 Plaatjes ordenen: afbeeldingen die passen in een verhaal, moeten in de juiste volgorde gelegd worden.
 - 5 Figuur leggen: een puzzel met een verschillend aantal stukken moet gemaakt worden tot één figuur, bijvoorbeeld een hand of een vliegtuig.

De normeringssteekproef van 1970 is gedateerd. De test is aanvullend genormeerd in 1976 voor 55-64-jarigen (Mas, 1976). In 2000 is een nieuwe bewerking verschenen. Deze berust op 745 deelnemers, 70% uit Nederland en 30% uit Vlaanderen. Er zijn zeven leeftijdsgroepen onderscheiden: 16-20, 21-25, 26-35, 35-50, 51-65, 66-75 en 76-85 jaar. De gemiddelde leeftijd is 43,34 jaar, SD: 21,18 jaar. Deze WAIS bevat enkele nieuwe subtests. Slechts 55% van de oude items bleef behouden. Het is een 'logische opvolger' (Uterwijk, 2001) van de uitgave van 1970. Vooral op de normering van de test, de omvang en de representativiteit van de steekproef, is kritiek geuit (Tellegen; 2002; Evers et al., 2000, pp. 1117-1124). In 2005 is de derde versie gepubliceerd. Sommige items zijn gemoderniseerd en er zijn nieuwe subtests ontworpen. Het instrument leidt tot één totaal IQ en tot twee scores voor performale en verbale intelligentie. Daarnaast zijn vier indexen geconstrueerd. De test is gehernormeerd en opnieuw gevalideerd (Uterwijk et al., 2005). Alle onderdelen worden als goed/voldoende gekwalificeerd (Cotanaanvulling 2007/2).

De jongste versie: *WAIS-IV-NL* is vanwege de rechten onder de naam van D. Wechsler (2012) verschenen onder de naam *Wechsler Adult Intelligence Scale - fourth Edition*: Nederlandstalige bewerking (Amsterdam: Pearson Assessment BV). Alle subtests zijn op onderdelen gewijzigd ten opzichte van de III versie (Barelds et al., 2013, p. 15). Het zijn: plaatjes ordenen, blokpatronen, overeenkomsten, cijferreeksen, matrix redeneren, woordenschat, rekenen, symbool zoeken, figuur leggen, informatie, symboolsubstitutie, begrijpen en onvolledige tekeningen. De scoring wordt vergemakkelijkt door voorbeelden. De normeringssteekproef is met zorg samengesteld en bevat 1500 deelnemers. De split-half subtestscores betrouwbaarheid varieert tussen de .77 en .93 en de vier indexscores tussen .88 en .97. Cronbach alfa's variëren van .68 tot .91. De test-hertest waarden lopen van .83 tot .96. Een confirmatieve factoranalyse ondersteunde het onderscheid tussen de vier indexen:

I: VBI (Verbaal Begrip Index)

II: PRI (Perceptueel Redeneren Index)

III: WGI (Werkgeheugen Index)

IV: VSI (Verwerkings Snelheid Index)

Het oordeel van de Cotan (2012) is 'goed' voor Materiaal, Handleiding en Betrouwbaarheid. Voor Normen en Begripsvaliditeit luidt het oordeel: voldoende. Onderzoek naar criteriumvaliditeit is beperkt en wordt voornamelijk als onvoldoende beoordeeld.

De *Groninger Intelligentie Test (GIT)* is een intelligentietest voor volwassenen van Nederlandse oorsprong (Luteijn et al., 1983). De test bevat tien subtests (woordenlijst, legkaart, vaaropdracht, sorteren, figuur ontdekken, cijferen, draaikaart, woordmatrijs en woord opnoemen I en II) en levert een IQ op. De *GIT-2* is een vernieuwde versie (Luteijn & Barelds, 2004). Uitgangspunten, Materiaal, Handleiding en Begripsvaliditeit zijn 'goed'. Normen, Betrouwbaarheid en Criteriumvaliditeit zijn voldoende (Cotanbeoordeling, 2006;

Cotan Aanvulling, 2007/1). Er is een schriftelijke verkorte vorm verschenen (Kooreman & Luteijn, 1987). De afname duurt dan 25 minuten en is onder meer gemaakt voor arbeids- en uitkeringsgerechtigden van die periode. De 2004 bewerking van Luteijn en Barelds heeft met feedback van gebruikers rekening gehouden. Een regel van het beoordelingssysteem is, een normering veroudert na 15 jaar. Een nieuwe normering is tijdrovend (stel maar eens een representatieve steekproef voor Nederland samen) en kostbaar. Daar komt bij dat een generatie (20-30 jaar) na het verschijnen van een test, en bij elke nieuwe versie zich een discussie ontspint over de normering. Het IQ stijgt in een generatie met ongeveer één SD volgens Flynn. Dat houdt in dat de oude normering gunstiger uitpakt dan de nieuwe. Zwanepol et al. (2002) hebben bijvoorbeeld gevonden dat de scores op de WAIS gemiddeld tien punten hoger zijn ten opzichte van zijn voorganger.

Geschiedstests: Er zijn Nederlandse bewerkingen van Amerikaanse geschiedstests in gebruik: de Differentiële Aanleg Test (DAT) en de *General Aptitude Battery* (GATB). Daarnaast hebben grote bedrijven en overheidsinstellingen eigen batterijen gemaakt. De DAT is voor Nederland bewerkt. De handleidingen dateren respectievelijk uit 1984 en 1991 en zijn van de hand van Evers en Lucassen. In 2007 zijn een nieuwe handleiding en bewerking beschikbaar gekomen. Veel leerlingen uit het voortgezet onderwijs hebben deze test gemaakt met het oog op hun vakkenpakket en beroepskeuze. De test is geschikt voor leerlingen vanaf 13 jaar. Er zijn negen subtests: denken met figuren, met woorden en getallen, en woordenlijst, ruimtelijk inzicht, taalgebruik, snelheid en nauwkeurigheid, woordbeeld en praktisch inzicht. Het onderzoek levert een schat aan gegevens op. Er zijn veel analyses op dit grote databestand mogelijk. Er zijn onder meer itemanalyses en factoranalyses verricht en er zijn correlaties met schoolcijfers en toetsen voor wiskunde en Nederlands berekend. Vermeldenswaard is dat bij deze test onderzoek is naar onder- of overpredictie bij allochtone leerlingen ten opzichte van autochtone leerlingen (Te Nijenhuis & Van de Vijver, 2001). De *bias* van de subtests met betrekking tot predictieve validiteit voor allochtone leerlingen ten opzichte van autochtone leerlingen blijkt gering. Een item of test is *gebiased*, partijdig als twee personen die dezelfde capaciteiten bezitten maar uit verschillende groepen stammen (sekse, ras, etnische groep) niet dezelfde kans op succes hebben op een item of test. Er zijn slechts enkele items met een *bias* voor het geslacht (denk aan opgaven voor technisch inzicht met voorbeelden van fabrieksmachines en auto's), maar dat had nauwelijks of geen betekenis voor de validiteit van de test. In de handleiding van 2007 is het onderzoek tot dan toe samengevat. De DAT kent een versie A voor HAVO/VWO (2005). Deze is in 2007 beoordeeld: Uitgangspunten, Materiaal en Handleiding worden als goed gewaardeerd. De Normen, Betrouwbaarheid en Begripsvaliditeit zijn voldoende. Criteriumvaliditeit is onvoldoende. De gerapporteerde correlaties zijn te laag en er is meer onderzoek nodig. Er is een versie B voor het VMBO (2005). De beoordeling is dezelfde als die van versie A. In 2007 is er een *Human Resource Management* (HRM) versie gemaakt. Deze krijgt een zelfde beoordeling als A en B met uitzondering van de Normering:

onvoldoende. De structuur het instrument is verbonden met het drie-lagen-model van Carroll (1993).

De *General Aptitude Battery (GATB, versie B 1971-1994)* is een tweede voor Nederland bewerkte geschiktheidstest. Er is een Stichting GATB Research. De laatste normering is van 2006 is het instrument is in 2009 beoordeeld. Uitgangspunten, Materiaal, Handleiding, Betrouwbaarheid en Criteriumvaliditeit krijgen het oordeel 'goed en Normen en Betrouwbaarheid voldoende. Personen vanaf 15 jaar kunnen ermee worden onderzocht. Het gaat om de meting van capaciteits- of geschiktheidsfactoren met het oog op beroepskeuze en selectie. De test bevat twaalf onderdelen. Acht onderdelen kunnen groepsgewijs worden afgenomen en vier individueel. De eerste acht subtests meten de factoren: Intelligentie (de G-factor), Verbaal vermogen (V), Numeriek vermogen (N), Ruimtelijk inzicht (S, space), Vormwaarneming (P, perception), Inzicht in administratief werk (Q) en Motorische coördinatie (K). Er wordt een profiel gemaakt dat de sterke en zwakke kanten van de sollicitant laat zien. Door herhaalde afnames van de test op verschillende tijdstippen bij vergelijkbare groepen is informatie verkregen over de prestaties van verschillende cohorten uit allerlei beroepsopleidingen. Van de Flier en Boomsma-Suerink (1990, 1994) ontdekten dat het niveau van het lager technisch onderwijs in 1990 verlaagd is ten opzichte van dat van 1975. De test is gebruikt om uit te zoeken of de computerafname van de test en de klassieke afname verschillen opleveren. Meestal treft men kleine verschillen aan. Zo blijkt dat bij computerafname meer items en vragen beantwoord worden dan bij de traditionele papier-en-potloodafname. Na analyse blijkt de structuur van de subtests enigszins te verschillen want de factorladingen van de subtests zijn niet gelijk bij de twee afnamen. Het effect van de computerafname is vooral zichtbaar bij de eenvoudige cognitieve tests.

Te Nijenhuis (1997) heeft onderzocht of er bij de GATB sprake is van bias ten opzichte van allochtonen in vergelijking met autochtonen. Dit berust op gegevens van 1322 eerste generatie allochtone en 806 autochtone sollicitanten bij de Nederlandse Spoorwegen. De leeftijd varieerde van 24 tot 31 jaar. De allochtonen verbleven gemiddeld 11,2 jaar, SD 6,9 jaar in Nederland. Enkele resultaten: de allochtone groep scoorde iets meer dan één standaarddeviatie lager dan de Algemene Nederlandse Schoolnorm. Op alle subtests scoorde de autochtone groep ongeveer één SD hoger. Er is verschil tussen Noord-Afrikanen en Turken en tussen Surinamers en Antillianen. Het is niet verwonderlijk dat het grootste verschil zich voordoet bij taal. Er blijven wel degelijk verschillen in gemiddelden nadat bepaald is of de items partijdig zijn: DIF. Ter vergelijking: Jensen (1998) en Rushton en Jensen (2010) vonden dat in de VS geboren zwarten ongeveer één SD IQ punten lager scoorden dan de blanke bevolking. Dat is vergelijkbaar met Nederlanders met een Noord-Afrikaanse en Turkse achtergrond. Dit kan wijzigen door de tijd heen. Het is zaak onderzoek te blijven doen en de recentste gegevens te gebruiken.

Prestatietests worden ontwikkeld door onderwijskundige experts. Deze zijn vertrouwd met de inhoud van de vakken en met onderwijsdoelstellingen. Zij kunnen vragen beantwoorden

als: 'Wat is het vereiste niveau van kennis van de Engelse grammatica in havo 4? Of 'Welke rekenvaardigheden moet een leerling aan het eind van de basisschool bezitten?' 'Hoeveel kennis van wiskunde moet een psychologiestudent hebben?' De modellen worden geleverd door moderne testtheoretici. Hun deskundigheid betreft niet de inhoud maar zorgt wel voor betrouwbare en meet-technisch deugdelijke instrumenten. De instrumentatie voor prestaties in het onderwijs is vergaand uitgewerkt in de VS en Europa. Nederland loopt hierin voorop. Bij prestatietoetsing ontbreekt volgens critici de psychologische basis. Er bijvoorbeeld ontbreekt informatie over benodigde kennis, begrip, toepassing, analyse, synthese en evaluatie van de schoolvakken. Kennis is eenvoudiger dan evaluatie maar in de tests overwegen kennisvragen. Dit geldt waarschijnlijk ook voor mc tentamens in het HBO en op de universiteit. Er zijn critici die erop wijzen dat prestatietoetsen vooral de volgende prestatie goed voorspellen en weinig gericht zijn op de snel veranderende eisen van functies en beroepen.

Er zijn in Nederland veel prestatietoetsen voor het onderwijs ontwikkeld. De bekendste is de Eindtoets Basisonderwijs van het Cito waarvan vanaf 1970 jaarlijks een nieuwe versie verschijnt. Deze heeft andere eerder gebruikte tests uit de 'markt gedrukt', zoals het schoolvorderingendeel van de Intelligentie Schoolvorderingen Interesse Test (de ISI: Van Boxtel et al., 1982). De Eindtoets wordt voor iedere school aanbevolen en mogelijk verplicht gesteld vanaf 2014 (plannen van staatssecretaris Dekker). De toets meet schoolvorderingen en neemt zes uur in beslag en is gespreid over twee ochtenden. Het is een meerkeuzetoets over: rekenen (getallen, hoofdrekenen, bewerkingen, meten, procenten, verhoudingen en vraagstukken); taal (correct taalgebruik (geen spelling), spellen van werkwoorden, spellen van niet-werkwoorden, begrijpelijk, passend taalgebruik en inhoud); informatieverwerking (hanteren van informatiebronnen, kaartlezen, lezen van tabellen en grafieken, lezen van teksten: reproductie en lezen van teksten: conclusies). De toets wordt bij grote groepen afgenomen en is ontworpen door deskundigen op inhoudelijk, onderwijskundig en psychometrisch gebied. De toets wordt op alle onderdelen als 'goed' gekwalificeerd, behalve op Begripsvaliditeit, dat met als 'voldoende' wordt gewaardeerd (Evers et al., 1992, 2000). De laatste beoordeling is van 2011 en het oordeel is gelijk gebleven. In 2010 zijn Materiaal en Normen herbeoordeeld. Er is een computerversie beschikbaar.

Periodiek peilingsonderzoek betreft eveneens toetsen van prestaties. Een gedateerd maar een voor het begrip duidelijk voorbeeld is van Wesdorp (1985). Het kan bovendien vergeleken worden met de PISA-onderzoeken (*Programme for International Student Assessment*) naar taalvaardigheid van 15-jarigen meer dan 25 jaar later. Wesdorps onderzoek is een voorstudie die betrekking heeft op functionele taalvaardigheid van leerlingen van de achtste groep in het basisonderwijs. Het onderzoek is verricht tussen september en december 1984. Er zijn zestien leestaken afgenomen. Dit type studies levert stof tot discussie over de kwaliteit van het onderwijs en over het peil van leerlingen. In de voorstudie wordt 7% van de leerlingen gekwalificeerd als een zeer slechte lezer. Deze 12-jarigen kunnen zich buiten de school niet redden met hun leesvaardigheid en we kunnen ze

beschouwen als functioneel analfabeet (Wesdorp, 1985, p. 79). Bovendien is er een 'twijfelgroep' van 14%. De schrijftaken laten 9% zeer slechte schrijvers zien. Bovendien is volgens de criteria 44% gekwalificeerd als twijfelachtig schrijfvaardig. Bij spreken worden minder problemen aangetroffen. Toch blijken sommige taken moeilijk, bijvoorbeeld het telefonisch meedelen aan de politie dat er op een bepaalde plaats een ongeluk heeft plaatsgevonden en het uitleggen aan een medeleerling hoe een spin een web weeft. Er zijn 1,5% zeer slechte sprekers en 11,5% twijfelachtige sprekers. Taal werd door de leerlingen een saai vak gevonden. Zelfstandig lezen komt weinig voor en schrijven nog minder. Strips worden naar verhouding nog het meest gelezen. Er wordt op de scholen gemiddeld ongeveer acht uur per week aan taal besteed. Het PISA rapporteert driejaarlijks onderzoek naar rekenen, lezen en (natuur)wetenschap bij de 15-jarigen in Nederland en vergelijkt de resultaten met 64 andere Westerse geïndustrialiseerde landen en Japan, Thailand en Singapore. In 2010 is geletterdheid vergeleken: 18% van de jongens scoort minimaal op leesvaardigheid en 11% van de meisjes. Leesvaardigheid omvat: 'zoeken en vinden', 'integreren en interpreteren' en 'reflecteren en evalueren'. Het Cito voerde dit onderzoek samen met een Australisch team uit. De waarden in de studie van Wesdorp van ruim 25 jaar daarvoor waren iets maar niet veel lager.

Het vergelijken van tests in verschillende landen vereist dat de tests hetzelfde meten in de landen. DIF is een aanpak om dat na te gaan. Voorbeelden:

(1) Kankaraš en Moors (2014) zijn dat nagegaan en troffen *non equivalence* (ongelijkheid, itempartijdigheid) aan bij de items van drie schalen: rekenen, *science* en lezen. De omvang van DIF is gemiddeld genomen bescheiden. De scores wijzigen voor rekenen en natuurwetenschap (*science*) en het meest voor lezen. De verandering is het grootst in de Zuidoost Aziatische landen.

(2) French et al. (2014) onderzochten DIF met het twee parameter logistisch IRT model van de 71 items tellende *Cornell Critical Thinking Test* (CCTT) bij studenten uit Noord Korea en de VS. De CCTT meet het proces waarbij een persoon laat zien of hij een situatie of probleem van verschillende invalshoeken kan bekijken. Het stereotype is dat Westerlingen dichotoom, binair denken in tegenstelling tot Aziaten. Hun denken is fluïde. Zijn de items partijdig in de twee culturen? Kan de vergelijking zonder meer gemaakt worden? De auteurs vonden bij 14% van de items DIF. Dit percentage wordt vaker aangetroffen bij tests bewerkt voor verschillende talen. Een uitzondering is Galic et al. (2014) die 36% DIF bij VS tegenover Kroatische studenten vonden in een *conditional reasoning* test. French et al. vonden bij één van de vier subtests (inductie) 26% DIF. De resultaten voor deductie, evaluatie en observatie, zoals geloofwaardigheid van beweringen, identificeren van verborgen aannames en betekenis waren aanleiding voor de auteurs om te zeggen dat het meeviel met DIF voor deze kritisch denken test.

De conclusie dat het meevalt met DIF van instrumenten in verschillende landen en groepen treft je regelmatig aan. Dit neemt niet weg dat er gelet moet blijven worden op (on)vergelijkbaarheid van tests in de verschillende landen/culturen. Telkens is een behoorlijk percentage van de items partijdig.

Criteria: betrouwbaarheid Uitspraken over prestaties kunnen alleen gedaan worden als er criteria zijn. Er zijn criteriumtoetsen: *criterion referenced tests*. Ze stellen vast of een persoon een afgesproken norm of criterium haalt dat door deskundigen in een vakgebied vastgesteld wordt. Daarnaast zijn er *norm referenced tests*. Daarbij weet je alleen wat een antwoord en een testscore voorstellen als je ze vergelijkt met de normeringsgroep. Er is nog een derde manier van normeren die minder bekend is maar aan invloed wint: *item referenced* normeren (Embretson & Reise, 2000). Daarin worden de nieuwe regels voor IRT toetsen geformuleerd. Rule 5 van *The New Rules* luidt bijvoorbeeld: *Test scores have meaning when they are compared for distance from items* (p. 15). De items worden aan de hand van de item-responsfunctie gewogen op basis van hun bijdrage aan de totaalscore. Een item dat goed discrimineert tussen goede en slechte leerlingen krijgt daardoor meer gewicht dan een item dat het onderscheid niet zo goed kan maken. Dit past in het modern testtheoretisch idee dat de waarde van een persoon op de latente trek bepaald kan worden, afgezien van de items en afgezien van de subgroepen personen, bijvoorbeeld van groepen op basis van leeftijd en sekse en autochtone en allochtone groepen.

Een in Nederland niet zo bekend verschijnsel is deelname aan openbare staatsexamens voor overheidsfuncties. Ofschoon kandidaten voor het diplomatenklasje van het Ministerie van Buitenlandse Zaken (BuZa) wel psychologisch onderzocht worden moeten ze daarnaast voor enige dames en heren verschijnen die hen aan de tand voelen over hun kennis van de Engelse en Franse taal en zij geven een oordeel of ze geschikt zijn voor deze keurige functies. Hij schijnt te helpen dat je uit een hoogleraarsgezin komt of kind bent van een politicus, maar wellicht is dat achterklap. Verder zoekt de overheid naar *high potentials* voor openbare functies. Selectie met behulp van tests en interviews wordt aan commerciële testbureaus overgelaten. Voor de Europese Unie kunnen aanstaande ambtenaren aan een openbare test deelnemen. Slechts enkelen met de hoogste scores worden geselecteerd. Het zijn aantrekkelijke functies en het aanbod is groot. Voor onze Zuiderburen zijn de tests en toetsen voor functies in buitenlandse dienst van meer gewicht. In de VS kun je bijvoorbeeld zes keer per jaar toetsen maken om je als leraar te kwalificeren. De uitslag daarvan is door onderzoekers uiteraard gecorreleerd met het functioneren als leraar. Dat verband bleek gering (Walsh & Betz, 1990, p. 246-247).

Een overzicht van alle in Nederland gebruikte tests en vragenlijsten is te vinden in Evers et al. (2000) en in de Aanvullingen die enkele keren per jaar bij uitgeverij Boom verschijnen. Daarin staat een oordeel over de psychometrische kwaliteiten van elke test, vragenlijsten of andere diagnostische procedure.

Elementaire cognitieve taken, EEG, EVP, reactietijden en informatieverwerkingsprocessen worden niet als tests of diagnosticum opgevat. Het zou een interessante oefening zijn om het beoordelingschema toe te passen op deze operationalisaties van intellectuele en cognitieve processen en constructen.

Samenvatting en conclusie

Psychometrici gebruiken bestaande items om hun modellen voor individuele verschillen in intelligentie, geschiktheden en prestaties te toetsen. Het SOI-model van Guilford is een uitzondering en was aanleiding tot het ontwerpen van nieuwe tests. Andere diagnostische procedures krijgen weinig aandacht omdat ze subjectief en/of inefficiënt zouden zijn. Informatieverwerkingsmetingen worden nauwelijks gebruikt door diagnostici. Ze kunnen helpen om intelligente acties genuanceerd te beschrijven. De complexe combinatie van individuele-verschillen dimensies met op elk vrij veel ontwikkelingsstappen van bijvoorbeeld Demetriou et al. is niet gebruiksvriendelijk. De beschrijving en toetsing van ontwikkelingschalen behoren tot het domein van onderzoekers en niet tot dat van practici. Als er al een instrument beschikbaar is, wordt het nauwelijks gebruikt. Alternatieve instrumenten voor praktische en emotionele intelligentie en competenties worden gebruikt, maar de betrouwbaarheid en construct en *incremental* validiteit moeten verder uitgezocht worden. Hun extra bijdrage voor het voorspellen en verklaren van criteria is vooralsnog beperkt en mijn vermoeden is dat het zo zal blijven.

In de praktijk zijn de algemene intelligentietests met hoge G-ladingen en de Differentiële-Geschiktheids-Tests het bruikbaarst. Er zijn in Nederland meer dan 500 tests en instrumenten beschikbaar en geëvalueerd. Ze meten vooral individuele verschillen in persoonlijkheid, intelligentie, geschiktheid en prestatie.

Operationalisatie van intelligentie en geschiktheid wordt aan itemschrijvers overgelaten en psychometrici gaan van hun items uit om die aan eigen modeltoetsen te onderwerpen. Het meten is door Spearman uitgewerkt in een factoranalytisch model dat een eerste dominante factor zocht en vond. Thurstone zocht bij de 56 tests naar een eenvoudige structuur waarbij elke afzonderlijke test laadde op één factor en niet op de zes andere. Guilford probeerde zoveel mogelijk onafhankelijke factoren te vinden. De *positive manifold* van de vele tests werd door Vernon en Carroll in een hiërarchisch schema geordend. Meten van intelligentie is zo ook een oefening in factoranalyse geworden. De informatieanalyse van items en taken speelde een beperkte en niet-succesvolle rol (Sternberg) bij testontwikkeling. Hij kan helpen te ontdekken waarom personen falen of slagen op tests en taken.

Prestaties worden gedefinieerd door deskundigen die curricula maken en onderwijsdoelstellingen schrijven. Meting daarvan staat onder het regime van de klassieke en vooral moderne testtheorie. Deze stellen in staat om moeilijkheidsgraad en discriminatiewaarde en samenhang van de items onderling en hun schaalkenmerken te bepalen. Bij de IRT wordt dat de vraag of de test een unidimensionele schaal vormt. IRT staat los van de specifieke vaardigheden die nodig zijn in een bepaald onderwijsdomein, en van de opbouw en structuur van onderwijsdoelstellingen.

Operationalisaties met behulp van informatieverwerkingsprocessen en elementaire cognitieve taken en om IQ en cognitie te analyseren zijn niet gemakkelijk te verzinnen en bijgevolg empirisch te realiseren. Ze bieden bovendien geen winst voor het voorspellen van

criteria. De klassieke tests en vragenlijsten zullen waarschijnlijk nog geruime tijd de dienst uitmaken.

Omdat er geen dwingende interpretaties zijn voor intelligentie- en geschiktheidstests kan de diagnosticus zelf uitmaken welke operationalisatie, meting en instrumenten passen bij de vraag van zijn cliënt: is een algemene index voldoende: een G-verzadigde test; is differentiatie nodig bijvoorbeeld met het oog op een vakken of beroepskeuze: een Differentiële Aanleg Test? Als het om nog specifiekere vragen gaat is er Guilford met tests voor de vele intelligenties. Schoolprestaties kunnen gemeten worden met Cito-toetsen en werkprestaties worden meestal gemeten aan de hand van output en beoordeeld door chefs en managers. Oordelen voldoen doorgaans minder aan psychometrische criteria dan output metingen. Cito-toetsen tonen een nadruk op adequaat meten en niet op inhoud. Die moet geborgd worden door itemschrijvers. Deze zouden de items niet goed verdelen over onderwijsdoelstellingen: bijvoorbeeld meer aandacht voor kennis dan voor toepassen en evalueren op basis van zelfbedachte criteria. Leerlingen die voldoen aan de toetsnormen zullen niet zo vaak bij de diagnosticus verschijnen; wel zij die niet voldoen. Om het probleem in kaart te brengen heeft hij naast de toetsen van kennis en inzicht ook kennis van onderwijsdomeinen en informatie van ouders, leraren en de sociale context nodig.

De Nederlandse diagnosticus heeft dus beschikking over een ruim aantal intelligentie-, geschiktheids- en prestatietests. Ze zijn alle beschreven en van de meeste is een oordeel over de kwaliteit beschikbaar in Evers et al. (2000) en de 'Aanvullingen' die door de Commissie Testaangelegenheden van het NIP gemaakt worden (Cotan).

2. Ontwikkeling

De stand van zaken bij intelligentie en cognitie is vergelijkbaar met die bij typerend gedrag. Dit houdt in dat dezelfde ontwerpen - meestal correlatieve en (quasi-) experimentele - en statistische modellen en data-analysetechnieken - vooral factoranalyse en geavanceerde technieken: SEM, tijdserie analyse en *growth modeling* gebruikt worden. De laatste schieten soms de conceptuele structuur van de theorie voorbij. Evenals bij de studie van de ontwikkeling van de persoon domineert onderzoek naar stabiliteit van intelligentie en geschiktheden. Het concept van ontwikkelingsstadium lokte studies naar overgangen uit: zijn ze geleidelijk of plotseling? Ontwikkeling van prestaties is ingebouwd in de hiërarchische organisatie van onderwijsdoelstellingen, bijvoorbeeld Blooms taxonomie van doelen voor het cognitieve domein van 1956, geactualiseerd door Krathwohl (2002). Er is geen ontwikkelingstheorie voor prestaties in domeinen als taal, rekenen, wiskunde, enzovoort. Verandering in prestatie kan bijvoorbeeld gekoppeld worden aan het bereiken van conceptueel geordende onderwijsdoelstellingen, maar ook vakinhoudelijk kunnen niveaus van inzicht beschreven worden, bijvoorbeeld bij rekenen, wiskunde, natuur- en scheikunde, talen en geografie.

Ontwikkeling vastgelegd met behulp van tests De 'ontwikkeling' van individuele verschillen in intelligentie en geschiktheid wordt geanalyseerd aan de hand van testcores. Sinds de

invoering van het deviatie IQ - een leeftijdgebonden normering - vallen *absolute* IQ verschillen op jongere leeftijd buiten beschouwing. Op oudere leeftijd wordt in longitudinaal en cross-sectioneel onderzoek vastgesteld of er *achteruitgang* is. Zoals eerder vermeld nemen de Gf en informatieverwerkende capaciteiten af en blijft de Gc min of meer gelijk. Deze studies worden uitgevoerd met G-verzadigde tests en verschillen tussen gemiddelden van diverse leeftijden worden getoetst. Bij een grote reeks longitudinale gegevens worden tijdseriemodellen gepast (zie bijvoorbeeld Glass et al., 1975). Deze modellen kunnen in principe het verloop van gedragingen beschrijven maar ze worden meestal gebruikt om effecten van interventies te bepalen. Rangorde stabiliteit van intelligentie en geschiktheidstests is een deel van de betrouwbaarheidsanalyse. De correlaties zijn meestal hoog, want als ze laag zijn, zijn ze onbetrouwbaar. In dat geval wordt de test óf niet gebruikt óf aangepast. Dit leidt er toe dat intelligentie en geschiktheid tot de stabielste gedragskenmerken worden gerekend.

De concepten differentiatie en integratie kunnen als ontwikkelingshypothesen worden beschouwd. Ze worden echter zelden gebruikt om de verandering van de structuur van intelligentie door de tijd heen te beschrijven. De concepten hebben betrekking op *structurele verandering*. De termen zijn door de Duitse psycholoog Werner gebruikt om de verandering in motoriek bij jonge kinderen te beschrijven. Later zijn ze op cognitieve ontwikkeling toegepast. Als er meer factoren nodig zijn om de verandering in de tijd te beschrijven kun je van differentiatie spreken. Als meer gedragingen in de loop der tijd binnen één factor gaan vallen is er sprake van integratie. De concepten worden gebruikt als beschrijving en toegelicht met voorbeelden maar er is nauwelijks onderzoek om deze concepten empirisch te realiseren. Structurele stabiliteit wordt vaak voorondersteld. Er wordt van uitgegaan dat correlaties en factorstructuur door de tijd heen niet veranderen. Dit verschilt met het onderzoek van typerend gedrag waarin structurele stabiliteit regelmatig onderzocht is. Verandering door de tijd heen van schoolprestaties, opleidingen en beroep wordt niet als een stabiliteitsvraagstuk of een geordende ontwikkeling opgevat. De verandering wordt bewaakt of gevolgd tegen de achtergrond van het beheersen van specifieke inhouden en het behalen van doelstellingen. Voor werknemers wordt er bijvoorbeeld een Persoonlijk-Ontwikkelings-Plan (POP) gemaakt maar daar komt geen ontwikkelings- of individueel verschillen concept aan te pas, laat staan een toetsing, want het is immers een management en sturingsinstrument. Prestaties van leerlingen worden bewaakt met het leerlingvolgsysteem van het Cito. Het laatste is beperkt tot het basisonderwijs, maar zou ook in het voortgezet en hoger onderwijs en bij werknemers toegepast kunnen worden. Bij dat systeem wordt het Rasch model gehandhaafd waarbij een schaal voor bijvoorbeeld de basisschoolperiode wordt ontworpen. Er wordt niet aan een ontwikkelingsschaal gedacht.

Voor operationaliseren en meten van de ontwikkeling van prestaties van werknemers en leerlingen is geen theoretisch kader geformuleerd. Het zijn de inhouden van de werkzaamheden, de vakken en de doelstellingen die het operationaliseren met behulp van

taken en items bepalen. Als er toetsen gebruikt worden, zijn ze genormeerd per leeftijd en is geen vergelijking *over* leeftijdsgroepen nodig om de scores te interpreteren.

Ipsatieve en proces stabiliteit zijn nauwelijks thema's voor onderzoek bij intelligentie en geschiktheid. Men kan de voorspelling van het IQ van adolescenten op basis van de Bayley-scores en de verandering in aandacht en habituatie opvatten als pogingen om de heterotypische stabiliteit van intelligentie te bepalen.

Stadiumovergangen De Piagetiaanse cognitieve ontwikkeling veronderstelt een verloop in stadia. De analyse gaat op dezelfde manier in zijn werk als bij stadia van persoonsontwikkeling. In de sterke stadiumopvatting zijn stadia categorieën die niet overlappen, in een vaste volgorde verschijnen en waarbij iedere persoon eindigt op hetzelfde niveau. Het ging Piaget niet om individuele verschillen maar de ontwikkeling van het abstracte *epistemic subject*. Zijn veronderstellingen worden niet volledig door onderzoek gesteund. De antwoorden van naastliggende stadia overlappen nogal eens, zodat een probabilistisch model realistischer is dan een deterministisch, strikt categorisch model. De overgang naar een volgend stadium is in zo'n model niet abrupt, maar glijdend. Het vaste en gelijke eindpunt: formeel operationeel denken en redeneren, deed zich wat later voor, rond het 18^{de} tot zelfs het 21^{ste} levensjaar en op het gebied van eigen expertise; niet zonder meer daarbuiten. Het voorstel om een postformeel stadium te beschrijven is min of meer aanvaard maar het is niet als apart stadium aangetoond. Piaget zou dat geen nieuw stadium noemen maar een uitwerking of generalisatie die op de formele operaties volgt.

Instrumentatie van ontwikkelingstests en sociale context: Instrumentatie voor cognitieve ontwikkeling vraagt observatie en/of het uitvoeren van taken en opdrachten die het mogelijk maken om iemand aan een stadium toe te wijzen. Dat kan met deterministische modellen zoals de Guttman schaal. Het is gezien vele bekende en toevallige invloeden realistischer om waarschijnlijkheidsmodellen te gebruiken. Er zijn overigens nauwelijks gestandaardiseerde instrumenten om personen aan cognitieve stadia toe te wijzen. De diagnosticus zal ermee instemmen dat er stadia zijn want cliënten kunnen niet alles kennen en leren op elke leeftijd. Hij heeft echter niet veel middelen om ze aan cognitieve stadia toe te wijzen. Als er al procedures zijn worden ze zelden gebruikt omdat het niet nodig is voor het beantwoorden van de vraag of omdat het zichtbaar is in welk stadium iemand verkeert. De combinatie van Piagetiaanse taken en intelligentietests heeft een profiel, opgespannen door zes dimensies, opgeleverd met op elke dimensie een betrekkelijk groot aantal stadia. Zo beschouwd is dit model mogelijk nog complexer dan het SOI van Guilford en nauwelijks bruikbaar voor praktische diagnostiek. Demetriou en collega's hebben taken van Piaget overgenomen een aangepast om hun theorie over de individuele verschillen x ontwikkeling combinatie te beschrijven. Case heeft een procedure beschreven om de geheugenruimte vast te stellen: is één dimensie in stelling gebracht of twee onafhankelijke die gecombineerd kunnen worden volgens een regel?

Er is in Nederland even belangstelling geweest voor het ontwerpen en toetsen van ontwikkelingsschalen. Kerssies et al. (1989) hebben een ordinale schaal ontwikkeld voor het

Piagetiaans sensomotorisch stadium dat uit zes stappen bestaat. De schalen zijn te gebruiken bij diepzwakzinnige pupillen en kinderen tot ongeveer 2 jaar (Rensen, 1994). Het instrument maakt het mogelijk om het ontwikkelingsniveau te bepalen op verschillende gedragsdomeinen. Elk domein is geordend in de zes stadia van de sensomotorische ontwikkeling. Deze verschillen kwalitatief en geen van de stappen kan overgeslagen worden. De schalen hebben betrekking op de volgende Piagetiaanse sensomotorische domeinen:

- 1 Visueel volgen en objectpermanentie.
- 2 Middelen om gewenste gebeurtenissen te laten plaatsvinden.
- 3 Imitatie van geluiden en gebaren.
- 4 Operationele causaliteit.
- 5 Constructie van objectrelaties in de ruimte.
- 6 Schema's voor het met elkaar in verband brengen van objecten.

Er zijn zeven tot veertien taken en items per schaal. Met behulp van speelgoed worden reacties van baby's en pupillen uitgelokt die objectief gescoord en/of door deskundigen beoordeeld worden. De zeven tot veertien stappen zijn teruggebracht tot de zes niveaus van het sensomotorisch stadium. Rensen vond een substantiële samenhang tussen de zes ontwikkelingsniveaus en scores op de Bayley Ontwikkelings Schalen. Het bepalen van deze correlatie wijst erop dat deze complexe schaalprocedure als *voorspeller* van intelligentie wordt gebruikt bijvoorbeeld als validering terwijl het een theoriegeleide *beschrijving* van stappen in de vroege ontwikkeling is. Een ontwikkelingsschaal moet 'zich bewijzen' door samen te hangen met een individuele verschillen instrument.

Een ander voorbeeld is van Sijtsma en Verweij (1992). Zij hebben een Mokkenschaal, een Item Respons Functie (IRF) met dubbele monotonie gebruikt, waarin zowel de items als de proefpersonen geordend zijn langs een ordinale schaal. Daarmee konden zij de ontwikkeling van het seriëren, het maken van reeksen volgens een bepaalde regel door kinderen beschrijven en verklaren als een natuurlijk ontwikkelingsverloop van eenvoudig naar ingewikkeld.

De Catastrofe theorie en Niet Lineaire Dynamische modellen zijn voor ontwikkeling van koude cognitie beperkt gebruikt en de aandacht ervoor is verminderd. Plotselinge overgangen zijn spectaculair, denk aan religieuze bekingen, revoluties en plotselinge weersveranderingen. Ze zijn indrukwekkend omdat kleine veranderingen in de onafhankelijke controle variabelen grote effecten kunnen hebben: een beetje geld uitgeven om een loterij te winnen. In het intelligentie- en cognitiedomein zijn de onderwerpen minder spectaculair, bijvoorbeeld de overgang van pre- naar concreet operationeel en ook niet zo duidelijk te zien.

Voorgaande voorbeelden en die van Ego-ontwikkeling hebben betrekking op een ontwikkelingsschaal voor één gedragsdomein. Er zijn daarnaast complexe modellen ontwikkeld om verschillende typen veranderingen van meer, en op een bepaalde manier gekoppelde gedragingen door de tijd heen te beschrijven. Dan volstaat een rangorde op één

schaal niet meer. Het eerder genoemde voorbeeld van de twee concurrerende groeiers is nog betrekkelijk eenvoudig omdat ze op elkaar afgestemd zijn. Als het gaat om meerdere gedragingen die verschillende (niet alleen lineaire) relaties hebben en zich ontwikkelen met verschillende snelheden dan zijn complexe modellen nodig om dat te beschrijven. Statistici en modelbouwers zijn wel zover maar de inhoudelijke theorievorming niet. Voorbeelden:

(1) Molenaar et al. (2009) stellen voor om individuele ontwikkeling als een niet stationair tijdserie proces te beschrijven. Verder hebben Bergman et al. (2002) de individuele persoon voorgesteld als een aantal ontwikkelende systemen met verschillende snelheden. De ontwikkelingspaden kunnen individueel verschillen.

(2) McArdle et al. (2009) combineren groeicurve modelering en Rasch-schaling om vergelijkbare testcores te verkrijgen van verschillende instrumenten die taal en geheugen meten. Zij proberen een gemiddeld groeipad te beschrijven waar elk individu een variatie op is.

De discussie over deze modellen vindt plaats in gespecialiseerde tijdschriften en boeken (*Psychometrika*; Collins & Sayer, 2001; Bijleveld et al., 1998). Ze worden weinig toegepast op concrete gedragingen. De modellen zijn soms verder gevorderd dan theorie, operationalisatie en instrumentatie. Ze helpen de diagnosticus niet.

Operationalisatie en meting van ontwikkelingsprocessen zijn dus beperkt. Versterking van en waardering voor de ontwikkelingsoriëntatie op IQ en cognitie kunnen aanzetten om empirische realisaties en typerende voorbeelden te bedenken voor deze gedragingen en hun verloop.

Alternatieve theorievorming brengt aspecten van intelligentie onder de aandacht die de expliciete uit de weg gaan, bijvoorbeeld Gardners natuurlijke intelligentie, competenties, EI en EQ. Er zijn items en taken voor ontworpen en geanalyseerd en experts beoordelen werknemers en hun producten. Er wordt soms meer aandacht aan de factoranalyse dan aan de aard van de vragen en taken besteed.

Ontwikkeling van prestaties wordt niet als ontwikkelingsverschijnsel opgevat. De Cito-toetsen worden per jaar gelijkgesteld: 'geëquivaalend'. Prestatietoetsen worden zelden vergeleken omdat ze verwijzen naar verschillende curricula. Er is bijvoorbeeld geen ontwikkelingstheoretisch kader voor tekstbegrip of vierkantsvergelijkingen oplossen. De SATs en Cito-toetsen stellen in staat een leerling te volgen, maar hij wordt steeds met zijn eigen leeftijdsgroep vergeleken. De operationalisaties verschillen per leerjaar om ze passend te maken bij de leerdoelen. Operationalisatie en meting van de ontwikkeling van prestaties zijn geen uitgesproken onderzoeksthema, maar het heeft wel zin. Blooms taxonomie sluit een natuurlijke ordening van capaciteiten in die lijkt op de epistemologische ontwikkeling bij kinderen en studenten.

Voor diagnostici is de ontwikkeling van prestaties een open deur. In de praktijk beschikken ze echter alleen over steekproefgegevens van prestatietoetsen en vergelijken ze de individuele prestatie daarmee om te zien of er achterstand is. Voorsprong bepalen ligt niet

zo voor de hand, omdat ze daarvoor toetsen zouden moeten gebruiken, die bij andere doelstellingen passen, bijvoorbeeld die van een volgend leerjaar.

Samenvatting en conclusie

De ontwikkeling van intelligentie, cognitie en geschiktheden wordt bestudeerd door testresultaten in de tijd te vergelijken en stadia van ontwikkeling te analyseren. Prestaties door de tijd heen worden vastgesteld met testtheoretische modellen en vergeleken met inhoud en doelstellingen van curricula. Absolute verschillen tussen gemiddelden worden voornamelijk getoetst bij ouderen, om het verloop - meestal achteruitgang - te bepalen van IQ en informatieverwerking met het ouder worden. Leeftijdverschillen bij kinderen en jeugdigen hoeven niet geïnterpreteerd te worden omdat tests per leeftijd genormeerd zijn. Rangorde stabiliteit is een onderdeel van de betrouwbaarheidsanalyse. Differentiatie en integratie van gedragingen: structurele stabiliteit wordt vermeld als vorm van ontwikkeling, maar wordt zelden empirisch gerealiseerd en getoetst. Gegevens over operationalisaties voor ipsatieve en processtabiliteit zijn ons niet bekend. In de studies worden longitudinale en cross-sectionele ontwerpen en conventionele multivariate technieken gebruikt.

Cognitieve ontwikkeling kan voorgesteld worden als het doorlopen van een aantal stadia, geoperationaliseerd aan de hand van items en taken. Modellen die het mogelijk maken om te toetsen of er bij zulke taken van een ordinale ontwikkelingschaal sprake is winnen terrein. Daarnaast ontwikkelen statistici complexe modellen om de individuele ontwikkeling van meer gedragingen met verschillende ontwikkelingssnelheden tegelijkertijd te beschrijven. Deze modellen gaan verder dan de conceptuele structuur van de ontwikkelingstheorieën. Er zijn dan ook weinig toepassingen en ze bieden de diagnosticus geen handvatten. Het is mogelijk dat hij op basis van ervaring complexe ontwikkelingspaden vanzelfsprekend vindt en ze probeert vast te stellen in gedrag van cliënten.

3. Sociale context

Er zijn operationalisaties van de natuurlijke context in de vorm van voor de hand liggende categorieën, zoals gezin, broers en zussen, school, buurt, stad, land, (sub)cultuur. Contexten kunnen ook met behulp van een aantal kenmerken beschreven worden. Veel kenmerken van sociale contexten hangen samen intelligentie en cognitie. Er is geen handzame en algemeen aanvaarde taxonomie van contexten. Gemanipuleerde contexten, bijvoorbeeld therapieën zijn even complex als de natuurlijke: welke ingrediënten van therapie en training zijn werkzaam voor intelligentie, prestatie en cognitie? In gedragsmodificatiestudies bestaat de context uit precies omschreven manipuleerbare stimuli. Scholing, training en therapie zijn vormgegeven, gestileerde en gecontroleerde contexten. Deze zijn zo complex dat je niet weet wat je precies aanricht. Dat is de reden dat er naar effectieve ingrediënten van sociale contexten wordt gezocht. Het operationaliseren en meten van zulke contexten is niet afgerond. De sociale context wordt bestudeerd als oorzaak ($S \rightarrow R$) van verschillen en veranderingen in intelligentie, geschiktheid, cognitie en prestatie. Ontwerpen en statistische modellen die bedoeld zijn om oorzaken op te sporen, zijn (quasi-) experimentele designs en

ermee verbonden variantieanalyses (Kirk, 1968; Shadish et al., 2002). De samenhang van natuurlijke sociale contextkenmerken met intelligentie en cognitie wordt onderzocht met deze ontwerpen. Er zijn, evenals voor ontwikkeling, complexe statistische modellen beschikbaar, bijvoorbeeld complexe pad- en multi-level analyses die samenhangende omgevingsfactoren correleren met een (on)gewenst gedrag, bijvoorbeeld (jeugd)criminaliteit of succes als manager.

Samenvatting en conclusie

Allerlei kenmerken van de gemanipuleerde en natuurlijke sociale context zijn mogelijke operationalisaties van die context. Deze worden met behulp van schalen uitgedrukt en gecorreleerd met intelligentie en cognitie. De bijbehorende statistische modellen zijn variantie- en correlatieve analyses. Daarnaast zijn er complexe modellen om de vele beïnvloedende variabelen in een structuur onder te brengen. De theorievorming is daarin minder ver dan de mogelijkheden van statistische modellen, zoals pad- en multi-level analyses en structurele vergelijkingstechnieken.

4. Reflectie en evaluatie

De onderzoeksdesigns zijn bij intelligentie en cognitie (quasi-) experimentele en correlatieve ontwerpen. Statistische analyses en testtheorie bestaan uit gebruik van gesofisticeerde en conventionele multivariate technieken, variantie analyse en IRT modellen. Enkele zijn verder gevorderd dan de theorie en spelen voornamelijk een rol in de discussie tussen statistici onderling. Ze zijn niet nuttig voor de diagnosticus. Instrumenten voor intelligentie, geschiktheid en prestatie zijn beschikbaar en de meeste voldoen aan de kwaliteitseisen.

Operationalisatie, empirische specificatie en meting van individuele verschillen in intelligentie en geschiktheid tonen een gelijktijdige uitwerking van intelligentietheorie en multivariate analyse, vooral factoranalyse. De eerste theoretici waren ook methodologen. De correlatiematrix van tests en items staat meerdere interpretaties toe: één factor, een aantal factoren en hiërarchische ordeningen van factoren. Je kunt evenals bijvoorbeeld bij het voorspellen van misdaad kiezen: één dimensie: het aantal misdaden, een *simple structured* profiel: demografische kenmerken, IQ, BF, of een complexe hiërarchie. Operationalisatie van intelligentie en cognitie is een taak van itemschrijvers. Soms zijn het onderzoekers die theorie geleid een instrument ontwikkelen, meestal echter inhoudelijk deskundigen. Psychometrici laten het schrijven van items aan theoretici en inhoudelijke experts over en gaan aan de slag met hun modellen om de structuur van de items te analyseren en te toetsen. Prestatiemeting (school en werk) staat vrijwel buiten de theorie en is het domein van itemschrijvers en psychometrici. Dit is zowel een sterkte vanwege de zorg voor psychometrische kwaliteit van de instrumenten, als een zwakte omdat er geen of een beperkte analyse van de inhoudelijke structuur van prestaties plaatsvindt.

De diagnosticus is een gebruiker en heeft nauwelijks invloed op de inhoud en structuur van de tests en toetsen. Inzicht in wat de individuele verschillen veroorzaakt is het domein van

specialisten: psychometrici, theoretici, onderwijskundigen en personeelschefs. De diagnosticus kiest de producten eclecticisch met de vraag van de cliënt als leidraad. Als hij een algemene indruk wil verkrijgen dan is een G-verzadigde test geschikt. Gaat het om advies voor een opleiding, beroep of functie dan is een multipele intelligentie- of geschiktheidstest op zijn plaats. Een genuanceerde interpretatie van intelligentie en geschiktheid door de twee-bij-twee combinaties van factoren, zoals bij de Big Five, is nooit voorgesteld maar in beginsel mogelijk. De factoren van de PMA correleren ook al zijn ze naar een simpele structuur te roteren. Als er ontwikkeling toegevoegd wordt op een aantal primaire factoren dan nuanceert dat de interpretatie van intelligentie en cognitie. In de praktijk van de diagnostiek wordt dit overigens niet gedaan want het is tijdrovend en complex. De ontwikkelingsstappen op de factoren zijn bovendien waarschijnlijk moeilijk te onderscheiden.

De IQ-, geschikheids- en prestatietests domineren omdat ze efficiënt, objectief en gemakkelijk te scoren zijn. Er zijn andere procedures: observatie, experimentjes (Piaget) en ondervraging, onder meer het diagnostisch interview. Deze helpen bij een beschrijving van het intelligent handelen en cognitief gedrag van de cliënt: een fenomenologische, gedetailleerde en zo volledig mogelijke weergave van zijn gedrag in een situatie. De waarde van de laatste voor het voorspellen, controleren en beslissen van het gedrag wordt betwist, maar is onder bepaalde condities gelijkwaardig aan conventionele tests. De diagnosticus moet zelf uitmaken of het protocol - meestal een voorgeschreven testbatterij - volstaat of dat andere procedures bijdragen om de vraag van de cliënt te nuanceren en vervolgens te beantwoorden.

De analyse van informatieverwerking van intelligentie- en geschiktheidstest items en taken helpt eveneens bij het beschrijven van wat er gebeurt bij het uitvoeren van een taak. Deze capaciteiten hangen samen met intelligentie en helpen om intelligentie in stappen, processen uiteen te leggen. De uitwerking is niet gestandaardiseerd en de analyse tijdrovend. Het stap voor stap uiteenrafelen van het verwerken van informatie-eenheden bij een complexe taak imponeert als een atomistisch, rationeel systeem maar zal niet gemakkelijk in concreet gedrag terug te vinden zijn. Ze zijn even moeilijk aan te tonen als het bestaan en de opeenvolging van ontwikkelingsstadia. Daarom gebruikt een diagnosticus zo'n analyse zelden of nooit.

Operationalisatie en meten van prestatie is het domein van onderwijskundige en testtheoretische experts: curricula en IRT modellen maken de dienst uit. De diagnosticus weet dat de instrumenten testtheoretisch adequaat zijn. Of inhoud en doelstellingen goed gerepresenteerd zijn, wordt soms betwist, omdat het schrijven van items en het maken van taken voor de 'hogere' doelstellingen (Bloom, 1956; Krathwohl, 2002) ontbreekt. Kennis items domineren.

Operationaliseren en meten van ontwikkeling worden vertaald als rangorde stabiliteit van intelligentie- en geschiktheidstests. De waarden zijn doorgaans hoog. Kwalitatieve ontwikkeling is vooral met behulp van ordinale schalen beschreven. Het onderzoek sluit aan bij specifieke thema's zoals sensomotorische fase en Ego-ontwikkeling en is beschrijvend

van aard. Daarnaast zijn complexe modellen geconstrueerd om de ontwikkeling op meer attributen tegelijkertijd te beschrijven. Deze hebben geen invloed op de diagnostische praktijk.

De sociale context is geoperationaliseerd in voor de hand liggende categorieën. Ze worden wat preciezer onderscheiden aan de hand van kenmerken die opgevat worden als factoren in een multiple persoonlijkheidstheorie. Er is geen taxonomie, geen bruikbare classificatie van categorieën en hun kenmerken.

De diagnosticus beschikt over voldoende instrumenten om individuele verschillen in intelligentie en cognitie betrouwbaar vast te stellen. Onderzoek van ontwikkelingsschalen en -paden en de alternatieve voorstellen zijn beschrijvend interessant maar bieden geen nieuwe efficiënte en gebruiksvriendelijke instrumenten met toegevoegde validiteit voor voorspellen, controleren en beslissen. Prestatietests voor school zijn technisch van hoge kwaliteit maar beperkt door dominantie van psychometrische modellering. Voor prestaties in werk geldt voor een deel hetzelfde. Daarnaast is er het oordeel van chefs en de productiemeting zoals bijvoorbeeld de Hirsch index voor de output van wetenschappers.

Onderwerpen en namen Hoofdstuk I, II, III, IV, V en VI

Intelligentie, geschiktheid en (school)prestatie

Maximum performance

Cognitie

Epistemologie

Sofisten en Plato

Aristoteles en Thomas van Aquino: intelligentie als actieve kracht

Descartes: zekerheid; ingeboren, heldere en onderscheiden ideeën

Synthetische oordelen apriori en aposteriori (Kant)

Locke en Hume ('Zijn is waargenomen worden')

Associatiewetten

Pseudo-empirisch onderzoek (Smedslund)

Intuitionisme, rationalisme, idealisme, empirisme, pragmatisme

Eclecticisme

Prototypisch begrip van intelligentie (Neisser)

Scholastic Aptitude Tests (SATs), Grade Point Average (GPA), College Admission Tests (CATs)

Scholing

Academic, practical, contextual intelligence (Sternberg)

Spearman: G en S factoren

Positive manifold van intelligentietests en -items (Carroll)

PMA, *simple structure* (Thurstone)

SOI model van intelligentie (Guilford)

Hiërarchische modellen van intelligentie (Vernon, Carroll)

Oorzaken van variatie in IQ in de populatie

Elementaire Cognitieve Taken (ECTs)

Inspectietijd (IT) en reactietijd (RT en RTSD)

Gf en Gc van Horn en Cattell

Overerfbaarheidscoëfficiënt: h^2

Binnen en tussen gezinnen variantie: c^2 en e^2

Piaget: organismische theorie; stadia; assimilatie en accommodatie

Sociale context: natuurlijke en gemanipuleerde contexten

Compensatie- en verrijkingsprogramma's

Dishabituatie en later IQ

Gedragsgenetica

Flynn effect

Sociale constructie van intelligentie en cognitie

Natuurlijke intelligentie (Gardner)

Competenties (Greenspan en Driscoll)

Leerpotentieel en dynamisch testen

Emotionele Intelligentie (EI), Emotioneel Quotiënt (EQ)

Educational Testing Service (ETS: VS)

Cito
Comorbiditeit

Slotbeschouwing

Het diagnosticeren van kennen, intelligentie, cognitie, geschiktheden en prestaties van de cliënt draagt substantieel bij aan het beschrijven, voorspellen, controleren en beslissen over zijn vraag, bijvoorbeeld over zijn school- of beroepsloopbaan, geestelijke gezondheid en beroepskeuze. De diagnosticus moet eerst uitmaken of het zinvol is om intelligentie en cognitie te bepalen voor het beantwoorden van de vraag. Het is soms een reflex, een gedachteloze gewoonte om onmiddellijk het IQ te bepalen met een test.

De cliënt kan en mag U vragen of het nodig is dat je de batterij aan tests afneemt. Het antwoord dat U dat bij iedereen doet volstaat niet. 'Gelijke monniken gelijke kappen' gaat over iets anders. Uw cliënt leeft in een geïndividualiseerde wereld. Hij verwacht maatwerk en daar maakt menigeen reclame mee, bijvoorbeeld in de gezondheidszorg.

De geschiedenis van (niet) westerse concepten over intelligentie en cognitie en laat zien dat ze een prominente rol vervullen in veel culturen. Ze zijn verbonden met de sociale en economische status van een persoon. Het concept is in het Westen vooral gericht op succes in onderwijs en beroep en minder op inter- en intrapersoonlijk succes. Dat het bloed kruipt waar het niet gaan kan, blijkt uit het ontstaan van alternatieven: niet alleen de koude *academic* intelligentie maar vooral emotionele intelligentie en empathisch vermogen. De laatste is salonfähig gemaakt als extra verklaarde variantie in gedragscriteria, al lukt dat nauwelijks. Het gaat immers om het voorspellen van een criterium. Dat is het vertonen van gewenst gedrag en onderdrukken van ongewenst gedrag.

Theorievorming is een 'scepsele' van onszelf net als het scheppen van geld door banken. We kunnen het theoretiseren niet laten. Onze begrippen gaan verder dan wat we zien. Ze is te vergelijken met geldschepping van banken. Die maken geld uit het niets. Dat geldt ook voor theorievorming. Als het geld een keer gemaakt is moet het zaken voor ons regelen: bedrijven en instellingen van ziekenhuizen tot verpleegtehuizen op gang helpen, mensen aan een goede woning helpen, de staat ondersteunen bij haar sociale en economische taken. Geld maken kan niet zonder vertrouwen; het is geen casino of pyramidespel. Dat geldt ook voor theorie: ze moet ons helpen dingen te begrijpen, te verklaren, te 'handelen'. Ze maakt een deel van de (sociale) werkelijkheid intelligibel en bewerkbaar. Dat begrijpelijk, toegankelijk maken kan steeds weer nieuwe vormen aannemen en tot nieuwe inzichten leiden. Er is niet één theorie die alles dekt en die er toe leidt dat de kous af is. Intelligentietheorie bijt zich daarbij in haar eigen staart: ze dient intelligentie en cognitie zelf intelligibel, handelbaar en nuttig voor ons te maken.

Kennis over impliciete theorievorming van leken over individuele verschillen, ontwikkeling en de effecten van de sociale context op *maximum performance* is van belang voor de diagnosticus. Deze lekentheorie kan een deel van het probleem van de cliënt zijn. Daarvoor bestaat geen vragenlijst. De diagnosticus heeft kennis van studies maar moet het voor een deel zelf uitzoeken. Impliciet is niet expliciet. Dat houdt in dat uw cliënt er een eigen referentiekader op na houdt. Dat is soms maar half gearticuleerd en als zo'n kader vaag is en de cliënt er toch iets over moet zeggen neemt hij - als ieder van ons - een klaarliggende interpretatie over. Dat is zo'n *availability heuristic* van Kahneman. Alleen hier is het geen

bias waar je de cliënt op wil betrappen. Het is iets om te articuleren, te achterhalen, recht te doen. De rol van onderzoeker is een andere als die van de diagnosticus.

Bij de expliciete theorievorming domineert de individueel verschillen oriëntatie. Dit feit gecombineerd met de beschikking over tests en de modellen van de psychometrici zorgen ervoor dat het testen van intelligentie psychometrisch verantwoord gebeurt. De diagnosticus heeft te maken met IQ tests, vele met een hoog G-gehalte zoals de WAIS, Raven, Gf en Gc en met enkele geschiktheidstests (DAT, GATB). De opmerkingen en bespiegelingen over het IQ: zijn aard, aantal factoren, fundamentele eenheden (ECTs, EVP, RT, RTSD, enzovoort) beïnvloeden het diagnostisch onderzoek nauwelijks. Het kennis nemen van dit onderzoek kan nuttig zijn voor de diagnosticus om de uitslagen op tests te interpreteren maar hij is zich ervan bewust dat ze de predictie, controle en beslissing niet verbeteren en de klassieke tests niet gaan vervangen. Ontwikkeling van intelligentie en cognitie is onontkoombaar maar geen prominent onderwerp in de diagnostiek. IQ tests worden gemaakt om stabiel te zijn over de tijd maar evenals met de persoonlijkheid is er naast stabiliteit ook instabiliteit in intelligentie op individueel en groepsniveau. Van de sociale context zowel in de vorm van de gemanipuleerde in interventies en training als de natuurlijke, bijvoorbeeld ouders en school werd en wordt veel verwacht: verbetering van intelligentie, geschiktheid en school- en beroepsprestaties. Dit vooruitgangsgeloof leidde tot (quasi) experimenteel en correlatieel onderzoek om effecten en relaties te bepalen. De resultaten van compensatie- en verrijkingsprogramma's zijn in de publicaties altijd positief. Is er nooit een negatieve of nul resultaat en gaat er nooit iets mis bij onderzoek? De d- en r-waarden zijn overwegend bescheiden tot gemiddeld. Er zijn ook voorbeelden van een nul effect van grote compensatieprogramma's voor kinderen vanaf groep 1.

Deze staat van het kennisbestand van de psychologie over intelligentie en cognitie leidt niet tot zekere uitspraken over individuele verschillen, de ontwikkeling en de invloed van de sociale context. De diagnosticus kan niet anders dan transparant zijn over zijn vaststelling en diagnose over intelligentie en cognitief functioneren. Dat kan hij volgens de regels van het vak doen. Daarbij is hij onthand als hij op termijn een perspectief moet geven over verandering. Daar wil de cliënt iets over weten en vernemen. Hij wil ook weten wat de betekenis is van de te manipuleren context (welke training moet ik doen) de natuurlijke context. De cliënt wil immers niet een getal weten of hoever hij van een gemiddelde afwijkt. Hij wil iets bereiken dat waarde voor hem heeft. Dat is wat hij als criteriumgedrag beschouwt en dat hoeft niet parallel te lopen met wat een school of werkgever als het criterium ziet. De diagnosticus zit zo tussen twee vuren. Daarover transparant zijn is wat lastiger dan uitslagen van instrumenten doorgeven.

Nieuwe perspectieven op cognitie, bijvoorbeeld identificatie en analyse van de werking van informatieverwerkingsprocessen zijn aantrekkelijk en verhelderend zoals rationalistische, atomistische schema's en architecturen dat kunnen zijn. Het blijkt echter moeilijk om de conceptuele onderscheidingen tussen stappen in de processen en hun volgorde concreet vast te stellen. Pogingen om deze processen in tests of procedures onder te brengen blijken moeilijk en het resultaat voorspelt criteria niet beter dan de klassieke tests.

Informatieverwerkingstheorie gaat uit van een rationalistisch, atomistisch en logisch geordende structuur en een daarbij passende aanpak van intellectuele taken bijna zoals in chemische procestechnologie. Aard en groei van kennis en inzicht zijn grilliger dan deze theorie voorstelt. Ze beschrijft waarschijnlijk niet wat er werkelijk gebeurt bij intellectueel werk. Het voordeel is dat taken uiteengelegd worden in bestanddelen en daar kan men bij interventie en training zijn voordeel mee doen. Als de cliënt zijn informatieverwerking als een structuur voorgeschoteld krijgt kan hij een zeker ongemak voelen. Het gaat bij hem immers niet volgens het boekje. Er zijn grillig- en toevalligheden. Nijhoff zei het zo: '...en het toeval nam een binnenweg naar het doel'. Een ordelijk leven kan ongepland tot stand komen. Wat heeft de diagnosticus op zo'n opmerking te zeggen?

Als het over intelligentie gaat, volgt de vraag naar de aard. Dit thema is goed aan te pakken met het IQ want er is een score die variantie vertoont in de populatie. Waar kan die variantie allemaal aan gebonden worden? Is intelligentie genetisch bepaald en voor een deel door de omgeving? Het onderwerp blijft aanleiding tot debat. Jensen (1969) gooide de knuppel in het hoenderhok met zijn vraag of we het IQ kunnen laten toenemen (*How much can we boost IQ and Scholastic Achievement: Harvard Educational Review, 29, 1-123*). Herrnstein en Murray veroorzaakten in 1994 opnieuw een heftig debat. Ze argumenteerden dat er wel toename was maar de verschillen tussen de groepen in de bevolking gelijk bleven.

Kortom de programma's die de variantie tussen leerlingen zouden moeten verminderen - meer gelijkheid - hadden gefaald. Intussen probeerden ervaren onderzoekers een balans te vinden, bijvoorbeeld Neisser et al. en Gottfredsen in 1996 en 1997. Het debat luwt een tijd maar kan en zal weer opblazen zoals opnieuw in het werk van de inmiddels overleden auteurs Jensen en Rushton. Ze wijzen er opnieuw op dat de verschillen tussen de zwarte en blanke Amerikanen nog steeds even groot (1 SD) zijn als 40 jaar geleden. Ondanks bedreigingen en van beschuldiging van racisme hield Jensen zijn standpunt vol (zie Lynn, 2012). Helpen al die programma's dan niet? Helpen de programma's juist de leerlingen die het naar verhouding het minst nodig hebben: het Mattheus effect? De feiten wijzen in die richting.

Gedragsgenetici onderzoeken geen genen en geen omgeving met instrumenten. Ze werken niet met een witte jas in een laboratorium en ze operationaliseren niet kenmerken van omgevingen en situaties. Ze verdelen varianties en de variatie in genetische verwantschap neemt altijd een flinke hap uit de IQ variantie en dat is meer dan de variantie binnen en tussen gezinnen. Er is bij intelligentie en cognitie onderzoek met enige regelmaat van een vlucht vooruit naar disciplines met meer prestige: genetica, biochemie en neurologie. Daar hebben psychologen geen verstand van. Ze zullen het aan experts moeten overlaten. Dit debat biedt de diagnosticus geen nieuwe antwoorden op vragen over beschrijven, voorspellen en beslissen. Intelligentie en cognitie gaan over de relatie tussen kenner en de wereld inclusief medemensen en zichzelf. Dit is een ook epistemologisch vraagstuk. De diagnosticus kan zich afvragen wat de impliciete, onnadrukkelijke epistemische opvatting over die relatie is bij hem en zijn cliënt. De methodologisch empirisch-analytische stroming

domineert, maar onnadrukkelijk is in de theoretische oriëntaties de relatie tussen kenner en objecten, verschijnselen en andere personen geformuleerd. De cognitieve ontwikkeling volgens Piaget heeft een uitdrukkelijke relatie met de wijsgerige achtergrond, zowel in de architectuur - van reflexen naar formele operaties - als in de inhoud: bestudering van wijsgerige thema's als object permanentie, conservatie, perspectiefnemen, moraliteit en inzicht in experimenteren.

De relatie tussen kennen en (sociale) context wordt in concreet onderzoek experimenteel en correlatief onderzoek als een $S \rightarrow R$ relatie beschouwd. Er zijn ook daar uitwerkingen van de speciale relatie tussen kenner en context: de *affordances* van de Gibsons, de afstemming van de waarnemer/kenner in Brunswiks probabilistisch functionalisme. Je komt er niet onderuit een relatie tussen kenner en gekende te beschrijven. Dat is vanouds een taak van de filosofie. Maar wat heeft de diagnosticus te zeggen als de cliënt vraagt hoe hij intelligentie en het verloop opvat? Hoe ziet intelligentie en cognitie eruit en kan er iets veranderd worden? Kan ik uiteindelijk alles leren als ik de tijd krijg?

Wat de aard van intelligentie ook mag zijn, de diagnosticus weet dat de IQ, geschikheids- en prestatiescores criteria voorspellen in een groep en zelfs voor een cliënt, als hij niet ver buiten de regressielijn valt. De test scores valt niets te verwijten. Ze veroorzaken de verschillen tussen groepen niet, ze tonen ze. Als de cliënt een verklaring vraagt voor de onderscheiden in soorten intelligentie en voor de aard en beïnvloedbaarheid wat heeft de diagnosticus dan te zeggen? Een hoogleraar die op de TV als soundbite beweert dat je de helft van je geluk (IQ, persoonlijkheid) zelf in handen hebt en de helft in je genen zit (Dijksterhuis) mag je hem vragen geldt dat overal? Amerikanen zeggen immers alles zelf te maken. Is dat onjuist? Hoe komt u bij die 50%. Verdeelt u het eerlijk of kunt u dat uitrekenen. Wat is geluk, intelligentie, cognitie en persoonlijkheid eigenlijk? Wat zijn de consequenties van die verdeling voor mijn geluk, IQ, persoonlijkheid? Meestal is de cliënt niet zo mondig, maar dat kan verkeren, al zullen ze nooit gelijkwaardig zijn. De diagnosticus staat aan het roer.

Alternatieve opvattingen over intelligentie en cognitie drukken verzet uit tegen eenzijdigheid van IQ tests en SATs en wijzen erop dat er zelden meer dan 50% (onder gunstige, 'administratief foutloze' omstandigheden) van de criteriumvariantie gebonden wordt. Er wordt gewezen op het belang van competenties, emoties, praktische intelligentie en management kwaliteiten op intelligent en aangepast functioneren. Er zijn nieuwe instrumenten ontwikkeld die *incremental validity* zouden bezitten. Van sommige van deze tests moeten de psychometrische kwaliteiten verder onderzocht worden. De geschiktheids- en prestatietest lijken zich wat meer in de luwte te bevinden, maar dat kan van korte duur zijn. Denk bijvoorbeeld aan het verzet tegen de Cito-toetsen bij sommige schoolleiders in het basisonderwijs. Ze zijn niettemin nuttig voor selecteren en toewijzen van leerlingen en personeel en het alternatief om het over te laten aan het oordeel van leraren en de wensen van ouders heeft zo zijn bezwaren. Dat was voor De Groot c.s. (*Vijven en Zessen*) precies reden om naar analogie van het ETS het Cito op te richten. Wat kan een diagnosticus zeggen op de opmerkingen dat er alternatieven zijn en dat hij die volgens een cliënt onvoldoende

weegt. Hij heeft er daar weet van door media en de (semi)wetenschappelijke bijdragen in de media. Kan hij ze afwijzen of voor de cliënt valide duiden?

Het operationaliseren en meten van de drie capaciteiten lopen parallel aan die bij de persoonlijkheid. Operationaliseren is goede voorbeelden van intelligent gedrag zo te specificeren dat ze scoorbaar en daarna te meten zijn door ze te passen in een model. De eerste stap is een construct concreet maken in taken en opdrachten. Dat beslist wat er uit de analyses komt. Ook al sprak Spearman van de *indifference of indicator*, het type taak doet er toe. Zo is het gemakkelijker om de sterke G te vinden bij de Raven - allemaal hetzelfde type items - dan bij de WAIS of de DAT met 12 en 9 subtests met verschillende materialen. De tweede stap is door psychometrici voldoende uitgewerkt. Hun kleinste eenheid is het item met zijn psychometrische kenmerken (IRFs en ICCs) en dat past in een unidimensionele schaal. Ze bieden voor alle onderzoeksvragen modellen. De laatste zijn soms voor op inhoudelijke constructen en hun relaties. Voor dagelijks gebruik beschikt de diagnosticus over voldoende instrumenten die aan eisen van een zinvolle operationalisatie en van een meetmodel voldoen. De prestatietoetsen worden in gespecialiseerde instituten door inhoudelijk deskundigen gemaakt maar psychometrici hebben het voor het zeggen of items in de test opgenomen worden of niet. Operationaliseren bevat een concreet antwoord op wat intelligentie is. Dat is niet voor eens en voor al te omschrijven. Een cliënt altijd beweren dat er niet naar zijn sterke of zwakke kanten is gekeken. Wat zegt u daarop? In een klein taalgebied als het Nederlandse zijn voldoende instrumenten beschikbaar. Er is een expertoordeel over de psychometrische en inhoudelijke kwaliteiten.

De diagnosticus kan uit de intelligentie- en cognitie literatuur en zijn praktijkervaring afleiden dat

de eerste vraag is of het nodig is en het de cliënt helpt om zijn IQ, geschiktheid en prestatie te bepalen.

er veel theorieën en verklaringen zijn voor individuele verschillen en wat minder voor ontwikkeling en weer vele voor de invloed van de sociale context op intelligentie en cognitie. Deze kunnen vreedzaam naast elkaar bestaan. Er is geen noodzaak te kiezen voor de 'ware' of de minst 'belaste' theorie want dat is niet uit te maken. Er is geen cruciaal experiment. De diagnosticus 'moet' daarom niet een G-verzadigde test of PMA test gebruiken. De keuze hangt af van de vraag; die helpt te kiezen tussen een algemene of een gedifferentieerde intelligentie of geschiktheidstest. Hij is bereid die keuzevrijheden aan de cliënt uit te leggen.

de keuze voor prestatietoetsen duidelijk is omdat ze verbonden zijn aan curricula en aan functies bij bedrijven. Hij kan er niet omheen want ze zijn vaak voorgeschreven, denk aan de Cito-toetsen, testbatterijen van grote bedrijven en protocollen van instellingen. Hij kan dus zeggen tegen de cliënt dat hij door de molen moet omdat het in het protocol staat. Hij zal vervolgens het protocol moeten uitleggen als een zinvolle werkwijze.

deze toetsen psychometrisch van goede kwaliteit zijn maar niet steeds theorie over het inhoudelijk domein bevatten, bijvoorbeeld over leren rekenen, taalverwerving, wereldoriëntatie, verwerving van de 1^{ste}, 2^{de}, 3^{de} taal, inzicht in wiskunde, biologie, scheikunde en geologie.

de informatieverwerkingsstappen en -capaciteiten nuttig zijn om het gedrag bij het uitvoeren van een taak en detail te beschrijven maar ze bieden vooralsnog geen *incremental validity* ten opzichte

van de klassieke tests en toetsen. Informatietheorie imponeert als een solide architectuur: alles (de verschillende eenheden) op een rij en ook nog in de goede volgorde. Het werkelijke verwerken van informatie is grilliger. Een cliënt kan geïmponeerd raken door de dwingende logica terwijl hij weet dat hij zelden zo te werk gaat bij taken en problemen op intellectueel/cognitief en sociaal gebied. De diagnosticus kan het verschil tussen een werkelijk proces en de lineaire reconstructie achteraf uitleggen aan zijn cliënt; ook dat het laatste kan verhelderen wat de cliënt doet.

de hulp van niet psychologische disciplines interessant is maar ze bieden onvoldoende om de huidige diagnostische praktijk daarop te baseren en andere instrumenten te gebruiken. De cliënt kan andere disciplines aandragen als verklaring voor zijn vraag/probleem. Het staat niet in het protocol hoe de diagnosticus moet antwoorden, maar hij zal zijn werkwijze transparant moeten maken.

er eenzijdig nadruk ligt op individuele verschillen instrumenten. Ze zijn psychometrisch naar verhouding het best uitgewerkt. Er is zodoende een nadruk op stabiliteit en niet op ontwikkeling en verandering van cognitie. De afhankelijke variabelen bij interventies zijn vaak individuele verschillen constructen en hun relatie met de context wordt daardoor in een S (een interventie, een covariaat) $\rightarrow R$ (een afhankelijke variabele, bijvoorbeeld prestaties) schema geplaatst. De cliënt wordt daardoor vooral geïnformeerd over de stabiliteit van zijn intelligentie en cognitie terwijl hij juist wil weten hoe ze te verbeteren.

de verwachting dat diagnostiek snel verbeterd zal worden door nieuwe theorie en meetmethoden en gesofisticeerde modellen niet op korte termijn bewaarheid zal worden. Bescheiden verbeteringen in kleine stapjes liggen meer in de rede.

studies die informeren over IQ, geschiktheden en prestaties uitgevoerd zijn in steekproeven. De diagnosticus moet die informatie toepassen op één, zijn cliënt. De laatste mag vragen gaat het allemaal over hem of over een profiel geconstrueerd op basis van een steekproef gaat.

het denkbaar is dat het vinden, bepalen en toetsen van effecten en correlaties in steekproeven uiteindelijk een achterhoedegevecht zal blijken en een Bayesiaanse aanpak en de verfijning en disciplineren van het formuleren van casus(sen) een aanvaarde diagnostische praktijk wordt. Maatwerk dus, waar de cliënt om vraagt. Maar zover is het nog niet. Hoijtink et al. (2013) verdedigen dit standpunt voor diagnostiek en proberen te demonstreren dat hun aanpak beter is dan

(a) de dimensionele aanpak (volgens hen een diagnose is gebaseerd op de locatie van de persoon op een of meer latente variabelen)

(b) dan het diagnostisch perspectief (volgens hen het toewijzen van een persoon aan een reeks latente klassen)

(c) dan het constructivistisch perspectief (het passen van de responsen van een persoon bij een groep symptomen, bijvoorbeeld een persoonlijkheidsstoornis) en

(d) beter dan de causale aanpak met de latente trek als veroorzaker van verschillen op een theoretisch attribuut. Het antwoord op een item is een functie van de antwoorden op items die aan dat item voorafgaan in een causaal systeem.

(e) de Bayesiaanse evaluatie van diagnostische hypothesen conceptueel aantrekkelijk is en niet onder doet voor andere werkwijzen in diagnostisch onderzoek. De vraag is hoe doe je dat dan precies, voorbeelden van diagnoses met casussen graag!

Deze flexibele, dynamische aanpak vraagt veel van de diagnosticus en is vooralsnog niet praktisch haalbaar.

Referenties en geraadpleegde literatuur

- Anderson, J.R., Bothell, D., Burne, M.D., Douglas, S., Lebiere, C. & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 4, 1036-1060.
- Asbury, K., Wachs, T.D. & Plomin, R. (2005). Environmental moderators of genetic influence on verbal and nonverbal abilities in early childhood. *Intelligence*, 33, 643-661.
- Barnes, J.C., Beaver, K.M. & Boutwell, B.B. (2013). Average county-level IQ predicts county-level disadvantage and several county-level mortality risks rates. *Intelligence*, 41, 59-66.
- Bar-On, R. (1997). *The Emotional Intelligence Inventory (EQ-I): Technical Manual*. Toronto, Ontario, Canada: Multi-Health Systems.
- Bartholomew, D.J., Deary, I.J. & Lawn, M. (2009). A new lease of life for Thomson's Bonds Model of intelligence. *Psychological Review*, 116, 3, 567-579.
- Batterjee, A.A., Khaleefa, O., Khalil, A. & Lynn, R. (2013). An increase in intelligence in Saudi Arabia, 1977-2010. *Intelligence*, 41, 91-93.
- Beauducel, A. & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment*, 18, 97-112.
- Bebedek, M., Jauk, E., Sommer, M., Arendasy, M. & Neubauer, A.C. (2014). Intelligence, creativity, and cognitive control: the common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, 46, 73-83.
- Benson, N., Hulac, D.M. & Bernstein, J.D. (2013). An independent confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) Integrated: What do the process approach subtests measure? *Psychological Assessment*. DOI: 10.1037/a0032298.
- Bergman, L.R., Magnusson, D. & El-Khoury, B.M. (2003). *Studying individual development in an inter-individual context: a person-oriented approach*. New Jersey: L. Erlbaum
- Berry, J.W. & Bennett, J.A. (1992). Cree conceptions of cognitive competence. *International Journal of Psychology*, 27, 73-88.
- Bierman, K.L., Domitrovitch, R.L., Nix, R.L., Gest, S.D., Welsh, J.A., Greenberg, M.T., Blair, C., Nelson, K.E. & Gill, S. (2008). Promoting academic and socio-emotional school readiness: The Head Start REDI program. *Child Development*, 79,6,1802-1817.
- Bijleveld, C.C.J.H., Van der Kamp, L.J.Th., Mooijaart, A., Van der Kloot, W.A., Van der Leeden, R. & Van der Burg, E. (Eds.) (1998). *Longitudinal data Analysis: Designs, Models and Methods*. London: Sage Publications.
- Birren J.E. & Schaie, K.W. (Eds.), 2006. *Handbook of the Psychology of Aging*. Amsterdam Elsevier: Academic Press
- Bleichrodt, N., Drenth, P.J.D., Zaal, J.N. & Resing, W.C.M. (1984). *Revisie Amsterdamse Kinder Intelligentie tests: Instructie, Normen, Psychometrische gegevens*. Lisse: Swets & Zeitlinger.
- Block, J. & Kremen, A.M. (1996). IQ and Ego-Resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, 70, 349-361.
- Boom, J., Brugman, D. & van der Heijden, P.G.M. (2001). Hierarchical structure of moral stages by a sorting task. *Child Development*, 72, 2, 535-548.
- Boom, J., Hoijsink, H. & Kunnen, S. (2001). Rules in the Balance: Classes, strategies, or rules for the balance scale task. *Cognitive Development*, 16, 717-735.
- Borsboom, D. & Dolan, C.V. (2006). Why is G not adaptation: a comment on Kanazawa (2004). *Psychological Review*, 113, 2, 433-437.

- Bowman, D.B., Markman, P.M. & Roberts, R.D. (2002). Expanding the frontier of human cognitive abilities: So much more than (plain) g! *Learning and Individual Differences*, 13, 1, 127-158.
- Brainerd, C.J. (1978). The stage question in cognitive-developmental theory. *The Behavioral and Brain Sciences*, 2, 173-213.
- Brockman, J. (2015) (Samensteller). *Wetenschappelijk Onkruid: 179 hardnekkelig ideeën die vooruitgang blokkeren*. Amsterdam: Maven Publishing.
- Brooks, B.L., Holdnack, J.A. & Iverdon, G.L. (2011). Advanced clinical interpretation of the WAIS-IV and WMS-IV: Prevalence of low scores varies by level of intelligence and years of education. *Assessment*, 18, 2, 156-167.
- Burgaleta, M., Head, K., Alvares-Linera, J., Martinez, K., Escorial, S., Haier, R. & Colom, R. (2012). Sex differences in brain volume are related to specific skills not to general intelligence. *Intelligence*, 40, 60-68.
- Calvin, C.M., Fernandes, C., Smith, P. Visscher, P.M. & Deary, I. (2010). Sex, intelligence and educational achievement in a national cohort of over 175.000 11-years-old schoolchildren in England. *Intelligence*, 38, 424-432.
- Canivaz, G.L. & Watkins, M.W. (2010). Investigation of the factor Structure of the Wechsler Adult Intelligence Scale - Fourth edition (WAIS-IV): Exploratory and higher-order factor analyses. *Psychological Assessment*, 22, 4, 827-836.
- Carroll, J.B. (1993). *Human Cognitive Abilities. A survey of factor analytic studies*. Cambridge: Cambridge University Press.
- Carugati, F.F. (1990). Everyday ideas, Theoretical Models and Social Representations: The case of intelligence and its development. In: Gün R. Semin & K.G. Gergen (Eds.). *Everyday understanding: Social and Scientific Implications*. London: Sage Publications.
- Case, R., Demetriou, A. Platsidou, M. & Kazi, S. (2001). Integrating concepts and tests of intelligence from the differential and developmental traditions. *Intelligence*, 29,307-336.
- Ceci, S.J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, 27, 5, 703-722.
- Ceci, S.J. (1996) *A bio-ecological treatise on intellectual development*. Cambridge MS: Harvard University Press.
- Chapman, P.D., Minton, H.L. & Zenderland, L. (1991). Schools as Sorters: Lewis M. Terman, Applied Psychology and the intelligence test movement 1890-1930. *Isis*, 82 (2), 415-416.
- Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
- Coleman D. (1995). *The number one bestseller: Emotional Intelligence; Why can it matter more than IQ?* London: Bloomsbury Publishing.
- Collins, L.M. & Sayer, A.G. (Eds.) (2001). *New methods for the analysis of change*. Washington DC: American Psychological Association.
- Colom, R., Espinosa, M.J. & Garcia, L.F. (2001). The secular increase in test scores is a 'Jensen' effect. *Personality and Individual Differences*, 30, 553-559.
- Covey, S.R. (2005). *The seven habits of high effective people*. New York: Free Press.
- Cronbach, L.J. & Snow, R.E. (1977). *Aptitudes and Instructional Methods*. New York: Wiley.
- Danner, D., Hagemann, D., Schankin, A., Hager, M. & Funke, J. (2011). Beyond IQ: a latent trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence*, 39, 323-334.
- Dar-Nimrod, I. & Heine, S.J. (2011). Genetic Essentialism: on the deceptive determinism of DNA. *Psychological Bulletin*, 137, 5, 800-818.

- Das, J.P., Naglieri, J.A. & Kirby, J.R. (1994). *Assessment of cognitive processes*. Needham Heights MA: Allyn & Bacon.
- Das, J.P. & Naglieri, J.A. (1995). Individual differences in cognitive processing and planning: A personality variable? *Psychological Record*, 35, 6, 335-352.
- Deary, I.J. (2012) Intelligence. *Annual Review of Psychology*, 63, 453-482.
- Deary, I.J. (2014). The stability of intelligence from childhood to old age. *Current Directions in Psychological Science*, 23 (4) 239-245.
- Deary, I.J., Penke, L. & Johnson, W. (2010). The neuroscience of human intelligence. *Nature Reviews Neuroscience*, 11, 201-211.
- Deary, I.J., Lawn, M., Bartholomew, D.J. (2008). A conversation between Charles Spearman, Godfrey Thomson and Edward L. Thorndike: The international examinations inquiry Meeting 1931 -1938. *History of Psychology*, 112, 122-142.
- Demetriou, A. & Efklides, A. (1987). Experiential structuralism and neo-Piagetian theories: Towards an integrated model. *International Journal of Psychology*, 22, 679-728.
- Demetriou, A., Efklides, A. & Platsidou, M. (1993). The architecture and dynamics of the developing mind. *Monographs of the Society for Research in Child Development*, Vol. 58, 5-5- serial no. 234.
- Demetriou, A. & Raftopoulos, A. (1999). Modeling the developing mind: from structure to change. *Developmental Review*, 19, 319-368.
- Demetriou, A. & Raftopoulos, A. (Eds.) (2005). *Emergence and transformation in the mind: Modeling and measuring Cognitive Change*. Cambridge: Cambridge University Press.
- Demetriou, A., Spanoudis, G., Mouyi, A., Kazi, S. & Platsidou, M. (2013). Cycles in speed-working memory-G relations: Towards a developmental-differential theory of the mind. *Intelligence*, 41, 34-50.
- Diesfeldt, H.F.A. & Vink, M.T. (1989). Visuele patroonanalyse en redeneren. Ravens CPM bij volwassenen van hoge tot zeer hoge leeftijd. *Tijdschrift voor Gerontologie en Geriatrie*, 20, 241-247.
- Donders, F.C. (1868). On the speed of mental processes. (W.G. Koster, Translation). *Acta Psychologica*, 30, 412-431. (first published in 1868).
- Dutton, E. & Lynn, R. (2013). A negative Flynn effect in Finland, 1997-2009, *Intelligence* (2013). <http://dx.doi.org/10.1016/j.intell.2013.05.008>.
- Edwards, A.E. (1957). *Techniques of attitude scale construction*. Century Psychology Series. New York: Appleton-Century-Crofts.
- Embretson, S.E. (1998). A cognitive design system approach to generating valid tests: application to abstract reasoning. *Psychological Methods*, 3, 3, 380-396.
- Embretson, S.E. (1999). Cognitive psychology applied to testing. In: F.T. Durso (Ed.). *Handbook of Applied Cognition*. New York: John Wiley & Sons.
- Embretson, S.E. & Reise, S.P. (2002). *Item response theory for psychologists*. London: Lawrence Erlbaum Publishers.
- European Journal of Psychological Assessment* (whole number 2, 2009): Assessing cognitive Failures (Guest Eds.: A. Efklides and G.D. Sideridis).
- Ennis, M.D. & Overton, W.F. (2006). Cognitive Developmental and Behavioral-Analytic theory evolving into complementarity. *Human Development*, 49.(3, 143-173.
- Evers, A. & Lucassen, W. (1984). *Voorlopige Handleiding DAT '83*. Lisse: Swets & Zeitlinger.
- Evers, A. & Lucassen, W. (1991). *Handleiding DAT '83*. Lisse: Swets & Zeitlinger.
- Evers, A. & Lucassen, W. (2007). *Handleiding DAT 2007*. Amsterdam: Pearson.

- Evers, A., Van Vliet-Mulder, J. & Groot, C. (2000). *Documentatie van Tests en Testresearch in Nederland*. Deel I en Deel II. Assen/ Maastricht: Van Gorcum.
- Evers, A.V.A.M., Van Vliet-Mulder, J., & ter Laak, J. (1992). *Documentatie van Tests en Testresearch in Nederland*. Assen/ Maastricht: Van Gorcum.
- Fink, A. & Neubauer, A.C. (2001). Speed of information processing, psychometric intelligence: and time estimation as an index of cognitive load. *Personality and Individual Differences*, 30, 1009-1021.
- Flanagan, D.P., Genshaft, J. & Harrison, P.L. (1997). (Eds.). *Contemporary Intellectual Assessment: Tests and Issues*. London: the Guilford Press.
- Flavell, J.H. (1963). *The developmental psychology of Jean Piaget*. Van Nostrand, Princeton, New Jersey.
- Fleishman, E.A. (1975). Towards a taxonomy of human performance. *American Psychologist*, 30, 9, 1127-1149.
- Flynn, J.R. IQ gains over time: Toward finding causes. In: Neisser, U. (Ed.) (1998). *The rising curve: long term gains in IQ and related measures*. Washington: American Psychological Association (pp. 25-65).
- Flynn, J.R. (1987). Massive gains in IQ in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 2, 171-191.
- Flynn, J.R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.
- Flynn, J.R. & Rossi-Casé, L. (2011). Modern women match men on Raven's Progressive matrices. *Personality and Individual Differences*, 50, 799-803.
- Fodor, J.A. (1974). Special sciences (Or the disunity of science as a working hypothesis. *Synthese*, 28, 97-115.
- Føllesdahl, H. & Hagtvét, K.A. (2009). Emotional Intelligence: The MSCEIT from the perspective of generalizability theory. *Intelligence*, 37,1, 94-105.
- Funder, D.C. (2001, 2nd Edition). *The personality puzzle*. New York: WW. Norton & Company.
- Furnham, A.F. (1990). Common sense theories of personality. In: Gün R. Semin & K.J. Gergen (Eds.). *Everyday understanding: Social and Scientific Implications*. London: Sage publications.
- Furnham, A.F. (2001). Self-estimates of intelligence: culture and gender difference in self and other estimates of both general (G) and multiple intelligences. *Personality and Individual Differences*, 31, 1381-1405.
- Furnham, A., Fong, G. & Martin, N. (1999). Sex and cross-cultural differences in the estimate multifaceted intelligence quotient scores for self, parents and siblings. *Personality and Individual Differences*, 26, 1025-1034.
- Gardner, H. (1983). *Frames of Mind: The theory of multiple intelligences*. New York: Basic Books.
- Gardner, H. (1984). The seven frames of Mind. *Psychology Today*, 21-26. (June, 1984).
- Gardner, H. (1993). *Creating minds*. New York: Basic Books.
- Gardner, H. (1999a). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.
- Gardner, H. (1999b). *The disciplined mind*. New York: Basic Books.
- Garlick, D. (2002). Understanding the nature of the general factor of intelligence: the role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review*, 109, 1, 116-136.
- Gazzaniga, M.S. (1989). Organization of the human brain. *Science*, 245, 947-952.
- Gelman, L. & Williams, E.M. (1998). Enabling constraints for cognitive development and learning: Domain specificity and epigenesis. In: *Handbook of Child Psychology, fifth edition. Cognition, perception and language*. D. Kuhn and R.S. Siegler, Volume Editors (pp. 575-631).

- Gergen, K. (2010). The acculturated brain. *Theory and Psychology, 20*, 795-816.
- Gergen, K.J., Gloger-Tippelt, G. & Berkowitz, P. (1990). The cultural construction of the developing child. In: Gün R. Semin & K.J. Gergen (Eds.). *Everyday understanding: Social and Scientific Implications* London: Sage Publications.
- Gignac, G. E. (2010). Seven-factor model of emotional intelligence as measured by Genos ET: a Confirmatory factor analytic investigation based of self- and rater-report data. *European Journal of Psychological Assessment, 26*, 4, 306-316.
- Gignac, G.E. (2015). The magical number 7 and 4 are resistant to the Flynn effect: no evidence for increases in forward or backward recall across 85 years. *Intelligence, 48*, 85-95.
- Girard, T.A., Axelrod, B.N. & Wilkins, L.K. (2010). Comparison of WAIS-III short form for measuring Index and full-scale scores. *Assessment, 17*, 3, 400-405.
- Goddard, H.H. (1927). Who is a moron? *The Scientific Monthly, 1*, 1, 41-46.
- Golay, Ph. & Lecerf, Th. (2011). Orthogonal higher order structure and confirmatory factor analysis of th French WAIS-III. *Psychological Assessment, 23*, 1, 143-152.
- Goldstein, K. (1967). Kurt Goldstein: Autobiography. In: E.G. Boring & G. Lindzey (Eds.). *A History of Psychology in Autobiography* (Vol. V.). (pp. 147-167).
- Gotlieb, A. (2014). *Operatie 'werk Arthur de deur uit'*. Onder redactie van Joep Dohmen & Jeroen Wester. Amsterdam: Bertram+deLeeuw Uitgevers.
- Gottfredson, L.S. (1997). Mainstream science on intelligence. *Intelligence, 24*, 1, 13-23.
- Gottfredson, L.S. (2012). Resolute ignorance on race and Rushton. *Personality and Individual Differences (2012)*. <http://dx.doi.org/10.1016/j.paid.2012.10.021> (pp. 1-6).
- Greenspan, S. & Driscoll, J. (1997). The role of intelligence in a broad model of personal competence In: D. Flanagan, J. Genshaft & P. Harrison (Eds.). *Contemporary Intellectual Assessment: Tests and Issues* London: The Guilford Press (pp. 131-151).
- Grégoire, J. Coalson. D.L. & Zhu, J. (2011). Analysis of the WAIS-IV index score scatter using significant deviation from the mean index score. *Assessment, 18*, 2, 168-177.
- Grigorenko, E.L. (2002). In search of the genetic engram of personality. In: D. Cervone & W. Mischel. *Advances in the Science of Personality*. London: The Guilford Press (pp. 29-82).
- Grigorenko, E.L., Sternberg, R.J. & Ehrman, M.E. (1 997). A theory based-approach to the measurement of foreign language learning ability: the CAN AL-F theory and test. *The Modern Language Journal, 84*, 390-405.
- Guilford, J.P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J.P (1985). The structure of intellect model. In: B.B. Wolman (Ed.). *Handbook of intelligence: Theories, measurement and applications*. New York: Wiley. (Pp.225-266).
- Guttman, L. (1971). Measurement as a structural theory. *Psychometrika, 36*, 329-347
- Haase, C.M., Heckhausen, J. & Wrosch, C. (2012). Developmental regulation across the life span: Toward a new synthesis. *Developmental Psychology*. DOI: 10.1037/a0029231. 1-9.
- Harman, H.H. (1967). *Modern Factor Analysis* (2nd revised Edition). Chicago: The University of Chicago Press.
- Harris, J.R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review, 102*, 3, 458-489
- Harrison, P.L., Flanagan, D.P. & Genshaft, J.L. (1997). *An integration and synthesis of contemporary theories, tests, and issues in the field of intellectual assessment*. In: P. Harrison, D. Flanagan & J. Genshaft (Eds.). *Contemporary intellectual assessment: Theories, tests and issues*. (pp. 533-561).
- Hebb, D.O. (1949). *The organization of behavior*. New York: Wiley.

- Haworth, C.M.A., Wright, M.J., Luciano, M., Martin, M.G., de Geus, E.J.C., van Bijsterveldt C.E.M., Plomin, N. (2010). The heritability of general cognitive ability increases linearly from childhood to adulthood. *Molecular Psychiatry*, 15, 112-1120.
- Heatherton, T.F. & Weinberger, G.J.L. (Eds.) (1994). *Can Personality Change*. Washington DC American Psychological Association.
- Hegel, G.W.F. *Phaenomenologie des Geistes*. (Phenomenology of the Mind). Dutch translation of J. Hahn 'De zelfverantwoording van de Geest'. Utrecht: The Netherlands: Futile.
- Heidegger, M. (1927; 1972). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag (Time and Being).
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.
- Hemphill, J.F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, 58, 78-79.
- Herrnstein, R.J. & Murray, C. (1994). *The Bell Curve*. New York: Free Press.
- Hofstee W.K.B. (2002). The questionnaire construction of personality: pragmatics of personality assessment. In: H.I. Braun, D.N. Jackson & D.E. Wiley (Eds.). *The role of constructs in psychological and educational measurements* (pp. 9-37).
- Holdnack, J.A., Zhou, X., Larrabee, G.J. Milis, S.R. & Salthouse, Th. A. (2011). Confirmatory factor analysis of the WAIS-IV/WMS-IV. *Assessment*, 18, 2, 178-191.
- Hong, Y, Morris, M.W., Chiu, C. & Martmez, V.B. (2002). Multicultural Minds: a dynamic constructivist approach to culture and cognition. *American Psychologist*, 55, 709-720.
- Horn, J. (1988). Thinking about human abilities. In: J.R. Nesselrode & R. B. Cattell (Eds.). *Handbook of Multivariate Experimental Psychology* (2nd Edition). New York: Plenum Press.
- Horn, J. (1989). Models of intelligence. In: R.L. Linn (Ed.). *Intelligence: measurement, theory, and public policy*. Urbana: University of Illinois Press (pp. 29-73).
- Horn, J.L & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 2, 253-276.
- Hornung, C., Brunner, M., Reuter, R.A.P. & Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence*, 39, 210-221.
- Hox, J. (2002). *Multilevel analysis: Techniques and Applications*. Mahwah NJ: L. Erlbaum.
- Hunt, E. (2011). *Human Intelligence*. New York, NY: Cambridge University Press.
- Hunt, E., Lunneborg, C. & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology*, 7, 2, 194-227.
- Hunter, J.E. & Schmidt, F.L (1990). *Methods of meta-analysis: correcting for error and bias in research findings*. Newbury Park, CA: Sage.
- Hutton, U, Wilding, J. & Hudson, R. (1997). The role of attention in the relationship between inspection time and IQ in children. *Intelligence*, 24, 3, 445-460.
- Jauk, E., Benedek, M., Dunst, B. & Neubauer, A.C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of an empirical breakpoint detection. *Intelligence*, 41, 212-221.
- Jensen, A.R. (1987). Process differences and individual differences in some cognitive tasks. *Intelligence*, 11, 107-136.
- Jensen, A.R. (1998). The G-factor. *The Science of Mental Ability*. Westport CT, Praeger Publishers.
- Jensen, A.R. (2011). The theory of intelligence and its measurement. *Intelligence*, 39, 171-177.

- Joseph, J. & Ratner, C. (2010). The fruitless search for genes in psychiatry and psychology: Time to re-examine a paradigm? *Council for responsible genetics: Gene Myths Project*. Available on line: <http://councilforresponsiblegenetics.org/pageDocuments/INX6VC0254>.
- Kail, R.V. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, 109, 3, 490-501.
- Kail, R.V. (2000). Speed of information processing: Developmental change and links to Intelligence. *Journal of School Psychology*, 38, 1, 51-61.
- Kail, R.V. (2013). Developmental analyses of individual differences in intelligence: Comments on Demetriou et al. (2013)., *Intelligence* (2013), <http://dx.doi.org/10.1016/j.intell.2013.05.005>.
- Kaloi, E. (2011). Integrative thinking is the key: An evaluation of current research into development of adult thinking. *Theory and Psychology*, 10, 1, 1-17.
- Kan, K.J. (2011). *The nature of nurture: The role of gene-environment interplay in the development of intelligence*. PhD dissertation: Universiteit van Amsterdam.
- Kanazawa, S. (2004). General intelligence as domain-specific adaptation. *Psychological Review*, 111, 512-523.
- Kanazawa, S. (2008). Temperature and Evolutionary novelty as forces behind the evolution of general intelligence. *Intelligence*, 36, 99-108.
- Kanazawa, S. (2010). Evolutionary psychology and intelligence research. *American Psychologist*, 65 (4). 279-289.
- Kanazawa, S. (2011). Intelligence and physical attractiveness. *Intelligence*, 39, 7-14.
- Kaufman, A.S. (2000). Intelligence tests and school psychology: predicting the future by studying the past. *Psychology in the Schools*, 3, 1, 7-16.
- Kavšek, J. (2004). Predicting later IQ from infant visual habituation and dishabituation: a meta-analysis. *Applied Developmental Psychology*, 25, 369-393.
- Kendler, K.S. (2005). A gene for... The nature of gene actions in psychiatric disorders. *American Journal of Psychiatry*, 162, 1243-1252.
- Kenny, A. (2001). *A brief History of Western Philosophy*. Oxford: Blackwell Publishers.
- Kerssies, I.J., Rensen, F.S.X., Oppenheimer, L. & Molenaar, P.C.M. (1989). *De ordinale schalen voor het bepalen van de psychologische ontwikkeling in de sensomotorische periode*. Lisse: Swets & Zeitlinger.
- Khemlani, S. & Johnson-Laird, P.N. (2012). Theories of syllogism: a meta- analysis. *Psychological Bulletin*, DOI: 10.1037/a0026841.
- Kiep, M. & Spek, A.A. (2015). De AQ (Autism-Spectrum Qoutiënt) bij Nederlandse mannen en vrouwen met en zonder ASS (Autismespectrumstoornissen). *De Psycholoog*, Maart 2015, 41-49.
- Kievit, R.A. (2014). *Turtles all the way down? Psychometric approaches to the reduction problem*. PhD thesis Universiteit van Amsterdam. Enschede: Ipskamp drukkers.
- Kievit, R.A., Van Rooijen, H. Wicherts, J.M., Kan, K.J. Waldorp, L.J., Scholte, H.S. & Borsboom, D. (2012). Intelligence and the brain: A model-based approach. *Cognitive Neuroscience*, 3, 89-97.
- Kim, K.H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education*, 16, 1, 57-66.
- Kirk, R.E. (1968). *Experimental design: Procedures for the Social sciences*. Belmont CA: Brooks/Cole Publishing Company.

- Kitchener, K.S. & King, P.M. (1981). Reflective judgment. Concepts of justification and their relationship to Age and Education. *Journal of Applied Developmental Psychology*, 2, 1, 89-116.
- Klein, C., Arend, I.C., Beauducel, A. & Shapiro, K.L. (2011). Individuals differ in attentional blink: mental speed and intra-subject stability matter. *Intelligence*, 39, 27-35.
- Kooreman, A. & Luteijn, F. (1987). *Handleiding Groninger Intelligentie Test (GIT): Schriftelijke verkorte vorm*. Lisse: Swets & Zeitlinger.
- Kort, W., Compaan, L., Bleichrodt, N., Resing, W.C.M. Schittekatte, M., Boon, M., Vernui, G. & Verhaeghe, P. (2002). *WISC-III NL. Handleiding*. David Wechsler. Amsterdam: Dienstencentrum NIP.
- Kramer, D. (1983). Post-formal operations? A need for further conceptualization. *Human Development*, 26, 91-105.
- Kratwohl, D.R. (2002). A revision of Bloom's taxonomy: an overview. *Theory in Practice*, 41, 4, 212-218.
- Kraus, M.W., Piff, P.K., Mendoza-Denton, R., Rheinschmidt, M.L. & Keltner, D. (2012). Social class, Solipsism, and Contextualism: How the rich are different from the poor. *Psychological Review*, 119, 3, 546-572.
- Kuhn, T.S. (1962). The structure of scientific revolutions. *International Encyclopedia of Unified Science (Vol. 2.2)* Chicago: The University of Chicago Press.
- Laak, J. ter (1995). *Psychologische Diagnostiek: Inhoudelijke en Methodologische Grondslagen*. Lisse: Swets en Zeitlinger.
- Lanziano, T. & Curci, A. (2014). Incremental validity of Emotional Intelligence Ability in predicting Academic Achievement. *American Journal of Psychology*, 127, 4, 447-461.
- Lezak, M.D., Howieson, D.B. & Loring, D.W. (2004). *Neuropsychological Assessment* (4th Edition). Oxford: Oxford University Press.
- Liu, J., Yang, H., Li, L., Chen, T. & Lynn, R. (2012). An increase of intelligence measured by the WPPSI in China, 1984-2006. *Intelligence*, doi: 10.1016/j.intell.2012.01.008.
- Luria, A.R. (1973). *The working brain*. New York: Basic Books.
- Luria, A.R. (1980). *Higher cortical functions in man*. (2nd revised Edition). New York: Basic Books.
- Luteijn, F. & Van der Ploeg, A.E. (1983). *Handleiding GIT*. Lisse: Swets & Zeitlinger.
- Luteijn, F. & Barelds, D.H.P. (2004). *Handleiding GIT- 2, 2004*. Amsterdam: Harcourt Testpublishers.
- Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the Development Quotients of Infants. *Intelligence*, 37, 16-24.
- Lynn, R. (2011). *The chosen people: A study of Jewish intelligence and achievements*. Washington: Summit Publishers Augusta GA978-1-59368-036-7408.
- Lynn, R. (2013). Who discovered the Flynn effect? A review of early studies on secular increase of IQ. *Intelligence*, <http://dx.doi.org/10.1016/j.intell.2013.03.004>.
- Lynn, R. & Vanhanen, T. (Eds.). (2002). *IQ and the wealth of nations. Human evolution, behavior, and intelligence*. Westport: Praeger Publishers.
- Lynn, R. & Vanhanen, T. (2012). *Intelligence: a unifying explanatory construct for the social sciences*. Bristol, UK: Ulster Institute for Social Research. (530 pp., paperback). ISBN: 978-0-958811-8-2
- Lynn, R. & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence*, 38, 353-360.
- Maraun, M.D. (1997). Appearance and reality: is the Big Five the structure of trait descriptions? *Personality and Individual Differences*, 22, 629-647.

- Marketdate Enterprises Inc. (2010). *The market for self-improvement products & services*. Rockville, MD: Market data Enterprises Inc.
- Marshalek, D., Lohman, D.F. & Snow, R.E. (1983). The complexity continuum in the radix and hierarchical models of intelligence. *Intelligence*, 7,107-127.
- Mas, J.M.J. (1976). *WAIS, Nederlandstalige bewerking: aanvullende normering voor de leeftijdsgroep van 55-64 jaar*. Amsterdam: Swets & Zeitlinger.
- Maul, A. (2011). The factor structure and cross-test convergence of the Mayer-Salovey-Caruso model of Emotional Intelligence. *Personality and Individual Differences*, 50, 457-463.
- Mayer, J.D. Salovey, P. & Caruso, D.R. (2008). Emotional Intelligence, New ability or eclectic traits? *American Psychologist*, 63, 6, 503-517.
- Mayer, J.D. Salovey, P. & Caruso, D.R. (2001). Models of emotional intelligence. In R.J.Sternberg (Ed.). *Handbook of Intelligence*. Cambridge: Cambridge University Press. (pp. 396-420).
- McArdle, J. J., Grimm, K.J., Hamagami, F. Bowles, R.P. & Meredith, W. (2009). Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. *Psychological Methods*, 14, 2, 126-149.
- McDaniel, M.A. (2005). Big-brained people are smarter: a meta-analysis of relationship between in vivo brain volume and intelligence. *Intelligence*, 33, 337-346.
- Merleau-Ponty, M. (1945). *Phénoménologie de la Perception*. Paris Gallimard: translated by Colin Smith 'Phenomenology of Perception' (1958). London: Routledge Classics.
- Meyer, G.J., Finn, S.E., Eyde, L.D., Kay, G.G., Moreland, K.L, Dies, R.R., Eisman, E.J., Kubiszin, T.W. Reed, G.M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *American Psychologist*, 56, 2,128-165.
- Mihura, J.L., Meyer, G.J., Dumitrascu, N. & Bombel, G. (2012). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the Comprehensive System. *Psychological Bulletin* DOI: 10.103/a0029406 (pp. 1-58).
- Mikolajczak, M., Balon, N. Ruosi, M. & Kotsou, I. (2012). Sensitive, but not sentimental: Emotionally Intelligent people can put their emotions aside when necessary. *Personality and Individual Differences*, 52, 537-540.
- Mingroni, M.A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other Trends. *Psychological Review*, 114, 3, 806-829.
- Molenaar, P.C.M., Sinclair, K.O., Rovine, M.J., Ram, N. & Corneal, S.E. (2009). Analyzing developmental Processes on an individual level using non-stationary time series modeling. *Developmental Psychology*, 45,1, 260-271.
- Moore, D.S. (2003). *The dependent gene: The fallacy of 'nature versus nurture '*. New York: Owl Books. Henry Holt and Cy, LLC.
- Mugny, G. & Carugati, F. (1990). *Social representations of intelligence*. Cambridge: University of Cambridge Press.
- Murray, Ch. (2012). *Coming apart: The state of white America 1960-2010*. New York: Crown Forum.
- Murphy, N.A. & Hall, J.A. (2011). Intelligence and interpersonal sensitivity: a meta-analysis) *Intelligence*, 39, 54-63.
- Musek, J. (2007). A general factor of Personality: Evidence for the Big One in the Five-Factor-Model. *Journal of Research in Personality*, 41, 1213-1233.
- Nagel, E. (1961). *The structure of Science: Problems in the logic of scientific explanation*. London: Routledge & Keagan Paul.

- Naglieri, J.A. (1999). *Essentials of CAS Assessment*. New York: Wiley.
- Naglieri, J.A. & Das, J.P. (1997). *Das-Naglieri Cognitive Assessment System*. Ithaca, IL: Riverside.
- Nederlands Bijbelgenootschap Haarlem (2004). *Nieuwe Bijbelvertaling*; Haarlem: Em. Querido's Uitgeverij B.V. en Uitgeverij Jongbloed. (pp.1897-1949: Matteüs Evangelie).
- Neisser, U. (1976). *General, academic and artificial intelligence*. Hillsdale NJ: Erlbaum.
- Neisser, U. (Ed.). (1998). *The rising curve: long term gains in IQ and related measures*. Washington: American Psychological Association.
- Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloof, R., Sternberg, R.J. & Urbina, S. (1996). Intelligence: knows and unknowns. *American Psychologist*, *51*, February, 77-101.
- Nelson, J.M., Canivez, G.L. & Watkins, M.W. (2013). Structural and incremental validity of the Wechsler Adult Intelligence Scale –Fourth Edition with a clinical sample. *Psychological Assessment* DOI: 10.1037/a0032086, 1-13.
- Nevo, B. (1993). In search of a correctness typology for intelligence. *New Ideas in Psychology*, *11* 391-397.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Nguyen, A.T.D. & Benet-Martínez, V. (2012). Biculturalism and Adjustment: a meta-analysis. *Journal of Cross-Cultural Psychology*, DOI: 10.1177/0022022111435097.
- NICHD Early Child Care Research Network (2004). Does class size in the first grade relate to children's academic and social performance or observed class room processes? *Developmental Psychology*, *35*, 651-664.
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F. & Turkheimer, e. (2012). Intelligence: New findings and Theoretical developments. *American Psychologist* (January 2, 2012) preprint.
- Ollendieck, T.H., Grills, A.E. & King, N.J. (2001). Applying developmental theory to the assessment and treatment of childhood disorders: Does it make a difference. *Child Psychology and Psychotherapy*, *8*, 2, 304-314.
- Oppenheim, P & Putnam, H. (1958). Unity of science as a working hypothesis. *Minnesota studies in the Philosophy of Science*, *2*. 3-36.
- Ou, S.-R. (2005). Pathways of long-term effects of an early intervention program on educational attainment: Findings from the Chicago Longitudinal Study. *Applied Developmental Psychology*, *26*, 4, 578-611.
- Parker, J.D.A., Keefer, K.V. & Wood, L.M. (2011). Toward a brief multidimensional assessment of emotional intelligence: psychometric properties of the Emotional Quotient Inventory- Short Form. *Psychological Assessment*, *23*, 3, 762-777.
- Pearson, H. (2006). Genetics: what is a gene?. *Nature*, 441.
- Penke, L., Borsboom, D., Johnson, W., Kievit, R.A., Ploeger, A. & Wicherts, J.M. (2011). Evolutionary psychology and intelligence research cannot be integrated in the way Kanazawa (2010) suggested. *American Psychologist*, *66*, 9, 916-918.
- Perry, J. (1970). *Forms of intellectual and ethical development in the College Years. A Scheme*. New York: Academic Press.
- Perry, J.C. & Cooper, S.H. (1989). An empirical study of defense mechanisms: I. Clinical interview and life vignette ratings. *Archives of General Psychiatry*, *46*, 444-460.
- Piaget, J. (1969). *Structuralism*. Meppel: Boom. (1970: English Edition: Structuralism. New York: Harper & Row).

- Piaget, J. (1970). Piaget's Theory. In: *Carmichael's Manual of Child Psychology*, 3rd Vol. (pp. 703-733). New York: John Wiley & Sons Inc.
- Piaget, J. (1972). Intellectual evolution from Adolescence to Adulthood. *Human Development*, 15, 1, 1-12.
- Plato, Vertaling van zijn *Verzameld werk* (1980): Baarn.
- Primi, R. (2014). Developing a Fluid Intelligence Scale through a combination of Rasch modeling and Cognitive Psychology. *Psychological Assessment*, 26, 3, 774-788.
- Putnam, H. (1967). The nature of Mental states. Reprinted in N. Block (Ed., 1980). *Readings in the philosophy of psychology* (pp. 223-236). London: Methuen.
- Ramsden, S., Richardson, F.M., Josse, G. Thomas, M.S.C., Ellis, C. Shakeshaft, C., Seghier, M.L. & Price, C.J. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature*, 00 Month, 2011 Vol. 000, 1-4.
- Rattan, A., Savani, K., Naidu, N.V.R. & Dweck, C.S. (2012). Can everyone become highly intelligent? Cultural differences in and societal consequences of beliefs about the universal potential for intelligence. *Journal of Personality and Social Psychology*, DOI: 10.1037/a00229263. 1-17.
- Raven, J.C. (1938). *Progressive Matrices*. London: Lewis.
- Reed, T.E. & Jensen, A.R. (1992). Conduction velocity in a brain nerve pathway of normal adult correlates with intelligence level. *Intelligence*, 16, 259-272.
- Rensen, F.S.X. (1994). *De organisatie van sensomotorische ontwikkelingsprocessen. Een structurele vergelijking tussen normale en geretardeerde ontwikkeling*. Academisch proefschrift Universiteit van Amsterdam. Amsterdam: F.S.X. Rensen.
- Resing, W.C.M. (2006). *Zicht op potentieel: over dynamisch testen, variabiliteit in oplossingsgedrag en leerpotentieel van kinderen*. Oratie 3 maart, 2006, Universiteit, Leiden.
- Ribaupierre, A. de (2015). Why should cognitive developmental psychology remember that individuals are different? *Research in Human Development*, 12, 237-245.
- Rindermann, H. & Neubauer, A.C. (2001). The influence of personality on three aspects of cognitive performance: processing speed, intelligence and school performance. *Personality and Individual Differences*, 30, 829-842.
- Roberts, M.J. (2002). The relationship between extraversion and ability. *Personality and Individual Differences* 32, 517-522.
- Robinson, M.D. & Gordon, K.H. (2011). Personality dynamics: Insights from personality social cognitive literature. *Journal of Personality Assessment*, 93, 2, 161-176.
- Ross, M. (1989). Relation of Implicit Theories to the Construction of Personal Histories *Psychological Review*, 96, 341-357.
- Runquist, E.A. (1936). Intelligence test scores and school marks in 1928 and 1933. *School & Society*, 43, 301-304.
- Rushton, J.P. (1999). Secular gains are not related to the g factor and inbreeding depression – unlike Black-White differences: A reply to Flynn. *Personality and Individual Differences*, 26, 381-389.
- Rushton, J.P. & Jensen, A.R. (2010). The rise and fall of the Flynn effect as a reason to expect a narrowing of the Black-White IQ gap. *Intelligence*, 38, 213-219.
- Rutter, M. (2007). Gene-environment interdependence. *Developmental Science*, 10, 1, 12-18.
- Sackett, P.R., Kuncel, N.R., Arneson, J.J., Cooper, S.R. & Water, S.D. (2009). Does Socio Economic Status explain the relationship between Admissions tests and post-secondary Academic

Performance? *Psychological Bulletin*, 135, 1, 1-22.

Salahodjaev, R. & Azam, S. (2015). IQ and the weight of nations. *Personality and Individual Differences*, 87, 105-109. (Tashkent, Oesbekistan).

Salthouse, T.A. (2004a). Localizing age-related individual differences in a hierarchical structure. *Intelligence*, 32, 541-561.

Salthouse, T.A. (2004) Aging and Cognition. *Current Directions in Psychology*, 13, 1, 140-150.

Sattler, J.M. (1982). *Assessment on Children's Intelligence and Special Abilities*. London: Allyn & Bacon.

Schaie, K.W. & Zajani (2006). Intellectual development across adulthood. In: C. Hoare (Ed.) *Handbook Of Adult Development and Learning*. New York NY: Oxford University Press.

Schaffer, H.R. (2006). Key Concepts in Developmental Psychology. London: Sage Publications

Schlesinger, I.M. & Guttman, L (1969). Smallest space analysis of intelligence and achievement tests. *Psychological Bulletin*, 71, 93-100.

Schmidt, F.L. (2014). A general theoretical integrative model of individual differences in interests, abilities, personality traits, and academic and occupational achievement: a commentary on four recent articles. *Perspectives on Psychological Science*, 9, 2, 211-218.

Schmidt, F.L & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

Schmidt, F.L & Hunter, J.E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 1, 162-173.

Scholten, Y. (2015), geïnterviewd door G. Kindermans. Werken met mensen met een verstandelijke beperking en gedragsproblemen. *De Psycholoog*, Maart 2015, 33- 37.

Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J. & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25, 167-177.

Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Company.

Shayer, M., Ginsberg, D. & Coe, R. (2007). Thirty years on – a large anti-Flynn effect? The Piagetian test Volume & Heaviness norms 1975- 2003. *British Journal of Educational Psychology*, 77, 1, 25-41.

Shikishima, C., Yamagata, S. Hiraishi, K. & Sugimoto, Y. (2011). A simple syllogism-solving test: Empirical findings and implications for g research. *Intelligence*, 39, 89-99.

Siegler, R. S. (1997). *Emerging Minds*. New York: Oxford University Press.

Siegler, R.S. & Crowley, K. (1991). The micro-genetic method: a direct means for studying cognitive development. *American Psychologist*, 46, 4, 606-620.

Siegler, R.S. & Svetina, M. (2002). A micro-genetic/cross-sectional study of matrix completion: Comparing short-term and long-term long change. *Child Development*, 73, 6, 793-809.

Monographs of the Society for research in Child development. Vol. 58, 5-5-serial no 234.

Sijtsma, K. & Verweij, A.C. (1992). Mokken Scale Analysis: Theoretical considerations and an Application to transitivity tasks. *Applied Measurement in Education*, 5, 355-373.

Silvia, P.J. & Beaty, R.E. (2012). Making creative metaphors: The importance of fluid intelligence for creative thought. *Intelligence*, 40, 343-351.

Sjöberg, L. (2001). Emotional Intelligence: A psychometric analysis. *European Psychologist*, 6, 2, 79-85.

- Spearman, C. (1907). Demonstration of formulae for true measurement of correlation. *American Journal of Psychology*, 18, 161-169.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
- Stanley, S. (2012). Intimate distances: William James' introspection, Buddhist mindfulness and experiential inquiry. *New Ideas in Psychology*, 30, 201-211.
- Starmans, Chr. & Friedman, O. (2012). The folk concept of knowledge. *Cognition*, 124, 272-283.
- Sternberg, R.J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sternberg, R.J. (1985). Implicit theories of intelligence, creativity and wisdom. *Journal of Personality and Social Psychology*, 49, 4, 607-727.
- Sternberg, R.J. (1991). Death, Taxes and bad Intelligence Tests. *Intelligence*, 75, 257-269.
- Sternberg, R.J. (1997). The concept of intelligence and its role in lifelong learning and success. *American Psychologist*, 52, 8, 1030-1037.
- Sternberg, R.J. (1997) *Successful Intelligence*. New York: Plenum Press.
- Sternberg, R.J. (Ed.). (2000a). *Handbook of Intelligence*. New York: Cambridge University Press.
- Sternberg, R.J. (Ed.) (2000b). *Practical Intelligence in everyday life*. Cambridge: Cambridge University Press.
- Sternberg, R.J. (2014). I study what I stink at: Lessons learned from a career in psychology. *Annual Review of Psychology*, 65, 1-16.
- Sternberg, R.J. & Grigorenko, E.L. (2002). *Dynamic testing. The nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Sternberg, R.J. & Hubart, L (1992). The creative mind. *Nederlands Tijdschrift voor de Psychologie (Dutch Journal of Psychology)*, 47, 3, 288-300.
- Sternberg, R.J. & Zhang, L. (Eds.) (2001). *Perspectives on thinking, learning and cognitive styles*. Mahwah NJ: Lawrence Erlbaum.
- Sternberg, R.J., Conway, R. Ketron, J. & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, 41, 1, 17-55.
- Sternberg, R.J., Grigorenko, E.L. & Bundy, D.A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, 47 -54.
- Sternberg, R.J., Grigorenko, E.L. & Bundy, D.A. (1997). *Measuring human development*. Washington: DC World Bank.
- Sternberg, R.J., Wagner, R., Williams, WW & Horvath, J.A. (1995). Testing common sense. *American Psychologist*, 50, 912-927.
- Stinissen, J.P., Willems, P.J., Coetsier, P. & hulsman, W.L.L. (1970). *WAIS, Nederlandstalige bewerking. Handleiding*. Amsterdam: Swets & Zeitlinger.
- Stolarski, M., Zajenkowski, M. & Meisenberg, G. (2013). National Intelligence and personality: Their relationship and impact on national economic success. *Intelligence*, 41, 94-101.
- Störig, H. J. (2002, 26th Edition). *Geschiedenis van de filosofie (History of Philosophy)*. Utrecht, the Netherlands: Het Spectrum B.V.
- Strasser, S. (1965). *Bouwstenen voor een filosofische anthropologie*. (Building blocks of a philosophical anthropology). Hilversum/Antwerpen: Paul Brand.
- Subotnik, R.F., Olszewski –Kubilius, F. & Worrel, F.C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, 12, 1, 3-54.

- Taylor, F. (1911). *Principles of scientific management*. New York, London: Harper and Brothers
- Taylor, Ch. (1989). *Sources of Self: The making of modern identity*. Cambridge: Cambridge University Press.
- Templer, D.I. & Rushton, J.Ph. (2011). IQ, skin color, crime, HIV, AIDS, and income in 50 U.S. States. *Intelligence*, 39, 437-442.
- Tellegen, P. J.(2002). *De kwaliteit van de normen van de WAIS III*. *De Psycholoog*, september, 2002, 463-466.
- Tellegen, P.J. & Van Dijk, H. (2009). *NIO en Allochtonen*. Amsterdam: Boom: test uitgevers.
- Te Nijenhuis, J. (1997). *Comparability of GATB scores for immigrants and majority group members in the Netherlands*. Dissertatie Amsterdam: Vrije Universiteit.
- Te Nijenhuis, J & Van de Vijver, F.J.R. (2001). Onderzoek naar partijdigheid. In: N. Bleichrodt en F. van de Vijver . *Diagnostiek bij allochtonen: mogelijkheden en beperkingen van psychologische tests*. Lisse: Swets & Zeitlinger (pp. 61-86).
- Te Nijenhuis, J., Jongeneel-Grimen, B. & Armstrong, E.L. (2015). Are adoption gains on the g factor? A meta-analysis. *Personality and Individual differences*, 73, 56-60.
- Thorndike, L.L. (1927). *The measurement of Intelligence*. New York: Teachers College.
- Thorndike, E.L. The international examinations inquiry meeting 1931 -1938. *History of Psychology*, 11, 2, 122-142.
- Thurstone, L.L. (1925). A method of scaling psychological and educational tests. *Journal of Educational Psychology*, 16, 4, 433-451.
- Thurstone, L.L. (1935). *Vectors of Mind*. Chicago: University of Chicago Press.
- Thurstone. L.L. (1938). *Primary Mental Abilities*. Chicago: Chicago University Press.
- Thurstone, L.L. & Thurstone, T. C. (1930). A neurotic inventory. *Journal of Social Psychology*, 1, 1, 3-30.
- Thurstone, L.L. & Thurstone, T.C. (1941). Factorial studies of Intelligence. *Psychometric Monographs, Whole No. 1*. Chicago: University of Chicago Press.
- Trahan, L.H., Stuebing, K.H., Fletxher, J.M. & Hiscock, H. (2014). The Flynn Effect: A meta-analysis. *Psychological Bulletin*, doi.org/10/1037a0037173.
- Tucker-Drob, E.M., Cheung, A. & Briley, D.A. (2014). Gross domestic product, Science Interest and Science Achievement: A Person x Nation interaction. *Psychological science*, 25, 11, 2047-2057.
- Turkheimer, E., Haley, A., Waldron, M., Donofrio, B. & Gottesman, I.I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 22, 125-133.
- Uterwijk, J. (2001). De Nederlandstalige WAIS-III. *De Psycholoog*, juni, 2001, 288-294.
- Uterwijk, J. Klinkenberg, E. & Kooy, A. (2005). *Wechsler Adult Intelligence Scale: WAIS-III 2005*. Amsterdam: Harcourt Test Publishers.
- Van Bon, W.H.J. (1986). *Raven's coloured matrices. Nederlandse normen en enige uitkomsten van onderzoek*. Lisse: Swets & Zeitlinger.
- Van Boxtel, H.W., Snijders, J.Th. & Welten, V.J. (1982). *ISI-reeks: Interesse, Schoolvorderingen,, Intelligentie. Publikatie 7 Vorm III. Verantwoording van het prestatiegedeelte & Handleiding voor de gehele testreeks*. Groningen: Wolters-Noordhoff.
- Van Boxtel, H.W. & Hemker, B.T. (2009). *Wetenschappelijke verantwoording van de inteelgentietest Eindtoets Basisonderwijs*. Cito B.V. Arnhem.
- Van Boxtel, H.W., Engelen, R. & De Wijs, A. (2011). *Wetenschappelijke verantwoording Eindtoets Basisonderwijs 2010*. Cito, Arnhem, 2001.

- Van de Flier, H. & Boomsma-Suerink, J.L. (1990). *Handboek GATB update 3*. Utrecht: Stichting GATB Research.
- Van de Flier, H. & Boomsma-Suerink, J.L. (1994). *Handboek GATB 1002-B*, 2^e druk, Separate uitgave Normen GATB B 1002-B . Utrecht: Stichting GATB Research
- Van der Maas, H.L.J., Dolan, C.V., Grasman, R.P.P.P., Wicherts, J.M., Huizenga, H.M. & Raaijmakers, M.E.J. (2006). A dynamic model of general Intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, 113, 4, 842-861.
- Van Dijk, H. (2008a). *De NIO en de PRO*. Amsterdam: Boom test uitgevers.
- Van Dijk, H. (2008b). *NIO en Dyslexie*. Amsterdam: Boom test uitgevers.
- Van Haassen, P.P. De Bruyn, E.E.J., Pijl, Y.J., Poortinga, Y.H., Lutje Spelberg, H.C., Van der Steene, G., Coetsier, P., Spoelders-Claes, R. & Stinissen, J. (1986). *Wechsler Intelligence Scale for Children Revisited. Nederlandstalige uitgave, Deel I: Testinstructie. Deel II: Scoring en Normen. Deel III: Verantwoording*. Lisse: Swets & Zeitlinger
- Vernon, P.E. (1950; 1961, 2nd edition). *The structure of human abilities*. London: Methuen.
- Vernon, P.A. & Mori, M. (1990). Physiological approaches to the assessment of intelligence. In: CR. Reynolds & R. Kamphaus, M (Eds.). *Handbook of psychological and educational assessment of children: Intelligence and achievement*. (pp. 389- 402). New York: Guilford Press.
- Verouden, G., Stet, F. & Scheele, J. (1987). *Psychologische selectie van etnische minderheden* (intern rapport). Amsterdam: GG & GD, Sektor Bedrijfsgezondheid: team Arbeids- en Organisatiepsychologie.
- Verschueren, K. & Koomen, H. Redactie (2007). *Handboek diagnostiek in de leerlingbegeleiding*. Antwerpen/Apeldoorn: Garant.
- Visser, R.H.S., Van Vliet-Mulder, J. Evers, A. & ter Laak, J. (1982). *Documentatie van Tests en Testresearch in Nederland*. Amsterdam: NIP.
- Vock, M., Preckel, F. & Holling, H. (2011). Mental abilities and school achievement: a test of a mediation hypothesis. *Intelligence*, 39, 357-369.
- Von Stumm, S. & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy to adolescence. *Intelligence*, 48, 30-39.
- Wai, J. & Putallaz, M. (2011). The Flynn effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces. *Intelligence*, 39, 443-455.
- Wai, J. Putallaz, M. & Makel, M.C. (2012). Studying intellectual outliers: Are there sex differences and are the smart getting smarter? *Current Directions in Psychological Science*, 21, 6, 382-390.
- Walsh, W.B. & Betz, N.E. (1990). *Tests and Assessment*. NJ: Prentice Hall.
- Watkins, M.W. (2010). Structure of the Wechsler Intelligence Scale for Children- Fourth edition among a national sample of referred students. *Psychological Assessment*, 22, 4, 782-787.
- Watkins, M.W. & Smith, L.G. (2013). Long-term stability of the Weschler Intelligence Scale for Children- Fourth Edition. *Psychological Assessment*, DOI: 10.1037/a0031653. (p. 1-7).
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children. 4th Edition: Technical and interpretative manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale. 4th Edition: Technical and Interpretative Manual*. San Antonio, TX: Pearson.
- Wechsler, D. , Kaplan, E., Fein, D., Kramer, J., Delis, D. & Maerlander, A. (2004). *Wechsler Intelligence scale for Childeren- Fourth edition-Integrated*. San Antonio, TX: Harcourt.
- Weisz, J.R. & Yeates, O.W. (1981). Cognitive development in retarded and non retarded persons: Piagetian tests of the similar structure hypothesis. *Psychological Bulletin*, 90, 1, 153-178.

- Weisz, J.R. & Zigler, E. (1979). Cognitive development in retarded and non retarded persons: Piagetian test of the similar sequence hypothesis. *Psychological Bulletin*, 86, 832-851.
- Wesdorp, H. (1985). *Goed Onderwijs, wat is dat? Voorstudie Periodieke Peiling van het Onderwijsniveau*. Den Haag: Stichting voor Onderzoek van het Onderwijs (SVO).
- White, R., (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, 66, 297-333.
- Wicherts, J.M., Borsboom, D. & Dolan, C.V. (2010). Why national IQs do not support evolutionary theories of intelligence. *Personality and Individual Differences*, 48, 91-96.
- Widiger, T.A. & Samuel, D.B. (2005). Evidence-based assessment of personality disorders. *Psychological Assessment*, 53, 2, 278-287.
- Wiggins, G.P. (1993). *Assessing student performance, exploring the purpose and limits of testing*. San Francisco: Jossey Bass Publishers.
- Williams, B. (2002). *Truth and Truthfulness*. Princeton and Oxford: Princeton University Press.
- Woodley, M.A. (2011). Comment: Heterosis doesn't cause the Flynn effect: a critical examination of Mingroni (2007). *Psychological Review*, 118, 4, 689-693.
- Wraw, C., Deary, I.J., Gale, C.R. & Der, G. (2015). Intelligence in youth and health at age 50. *Intelligence*, 53, 23-32. (UK).
- Zell, E. & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science*, 9, 2, 111-125.
- Zwanepol, E.A., Berghuis, H., De Groot, E. & Luteijn, F. (2002). Vergelijkend onderzoek tussen de WAIS III en de verkorte GIT. *De Psycholoog*, 37, 2002, 581-587.

Auteur

Jan J.F ter Laak (1944) was werkzaam aan de Faculteit Sociale Wetenschappen van de universiteiten van Nijmegen, Tilburg, Leiden en Utrecht. Bij de afdeling Ontwikkelingspsychologie van de Universiteit Utrecht doceerde hij gedurende bijna 20 jaar psychologische diagnostiek. Hij was in de praktijk werkzaam te Enschede, Haarlem en Breda. Hij was lid van het hoofdbestuur van het Nederlands Instituut van Psychologen (NIP) voor psychologische diagnostiek, lid en voorzitter van de Commissie Testaangelegenheden het NIP (de Cotan) en voorzitter van de sectie Kinder & Jeugd van het NIP.