

# Nederlandse Samenvatting

Dit proefschrift, getiteld *Van Structural Equation Modeling naar Next-Generation Sequencing: Het Evoluerende Landschap van de Moderne Gedragsgenetica*, gaat over de ontologie van latente psychometrische variabelen en over de genetica van intelligentie. Hieronder geef ik een samenvatting van de voorgaande acht hoofdstukken.

In hoofdstuk 2 wordt de techniek van het *structural equation modeling* geïntroduceerd, zoals deze toegepast wordt in het klassieke tweelingonderzoek. Eerst bespreek ik de basismethode voor het analyseren van overeenkomsten tussen familieleden om conclusies te kunnen trekken over de effecten van niet-geobserveerde genetische factoren en omgevingsfactoren. Daarna komt de implementatie van verschillende modellen in *structural equation modeling* aan bod voor genetisch-informatieve studies, waaronder simplex en latente groeicurve modellen. Vervolgens worden *structural equation*-modellen besproken die speciaal in de genetica zijn ontwikkeld, waaronder factoranalysemodellen met gemeenschappelijke en onafhankelijke factoren, modellen voor de interactie tussen genotype en omgeving, sekse-limitatiemodellen, en causale modellen waarbinnen wordt getracht de richting van causale effecten te bepalen. Het hoofdstuk eindigt met een discussie over het opnemen van gemeten genetische varianten in associatieanalyses gebaseerd op *structural equation modeling*.

Hoofdstuk 3 gaat over het gebruik van genetisch-informatieve analyses op het niveau van afzonderlijke items om meer te weten te komen over de ontologie van latente gedragsfenotypes (bijvoorbeeld depressie en algemene cognitieve vaardigheden). Hiervoor wordt gekeken naar de mediërende rol die ze spelen met betrekking tot genetische en omgevingsinvloeden. De beschikbaarheid van genetisch-informatieve data op itemniveau maakt het mogelijk om 1) te onderzoeken wat de empirische implicaties zijn van de realistische interpretatie van latente psychometrische eigenschappen, te weten hoe de geobserveerde covariatie gemedieerd wordt door genetische en omgevingsinvloeden, en 2) de dimensionaliteit van de latente covariantiestructuren (zowel genetisch als van de omgeving) te bestuderen die leidt tot de geobserveerde itemcovariatie. Ik merk op dat de typische problemen in het vaststellen van dimensionaliteit in psychometrisch onderzoek een gevolg kunnen zijn van de verschillen in de dimensionaliteit van genetische en omgevingscovariantiestructuren, en stel voor om genetisch-informatieve data op itemniveau te gebruiken als een hulpmiddel om meer inzicht te krijgen in de fenotypische dimensionaliteit.

In hoofdstuk 4 wordt de methodologie uit hoofdstuk 3 gebruikt om de ontologie en de genetische en omgevingsetiologie van de *internalizing syndrome dimensions* van de *Child Behavior Checklist* te onderzoeken (CBCL; Achenbach, 1991; Verhulst, Van der Ende, & Koot, 1996). De resultaten suggereren dat de verschillende syndromen van de CBCL beter begrepen kunnen worden als een samengesteld geheel van vrij te schatten genetische en omgevingsinvloeden dan als causale entiteiten die zouden leiden tot de waargenomen covariatie tussen symptomen. Bovendien geven de resultaten aan dat er een gemeenschappelijke genetische basis is voor angst, depressie en teruggetrokken gedrag, waarbij het onderscheid tussen deze syndromen voornamelijk ontstaat door individuspecifieke omgevingsfactoren. De bevindingen worden besproken in de context van de bekende moeilijkheid een onderscheid te maken tussen verschillende diagnostische categorieën zoals angst en depressie.

In hoofdstuk 5 wordt dezelfde methodologie gebruikt om 1) de houdbaarheid van de realistische interpretatie van de *Big Five* persoonlijkheidsdimensies (McCrae & Costa, 2008)

te onderzoeken, en 2) om de structuur van de genetische en omgevingscovariantiematrices te analyseren die ten grondslag liggen aan de geobserveerde covariatie van de persoonlijkheidsitems van de *NEO Five Factor Inventory* (NEO-FFI; Costa & McCrae, 1992). Het is opmerkelijk dat deze covariantiematrices vergelijkbare (vijf-factor) structuren vertonen, in tegenstelling tot de resultaten voor de CBCL. Echter, de latente persoonlijkheidsdimensies lijken de genetische en omgevingseffecten op de items niet volledig te verklaren, zoals te verwachten zou zijn onder de realistische interpretatie van de *Big Five*. Ik bespreek de implicaties hiervan voor de substantieve interpretatie van de *Big Five*.

Hoofdstuk 6 geeft een overzicht van de methodologie die gebaseerd is op het modelleren van genetische covariantiestructuren voor onderzoek naar angst en depressie bij kinderen, en het geeft een overzicht van de relevante bevindingen. De vragen die aan bod komen gaan verder dan alleen het inschatten van de bijdragen van genetische en omgevingsfactoren aan angst en depressie, wat relatief simpel is. Het blijkt dat er relatief consistente evidentie is voor a) kleine tot verwaarloosbare sekseverschillen in de genetische etiologie van deze aandoeningen, b) een belangrijke rol voor genetische factoren in het verklaren van stabiliteit door de tijd heen, c) een bijdrage van genetische factoren aan de comorbiditeit tussen angst en depressie, d) een mogelijke rol van genotype-omgevingsinteractie, e) een rol voor een correlatie tussen genotype en omgeving, en f) een ondergeschikte, misschien zelfs helemaal geen etiologische rol voor interactie tussen broertjes en zusjes.

De hoofdstukken 7-9 gaan over de genetica van intelligentie. In hoofdstuk 7 worden alle longitudinale data met betrekking tot verbale, non-verbale en algemene intelligentie samen geanalyseerd, zoals verzameld in het Jonge Nederlands Tweelingen Register (Bartels et al., 2007) in 2009. Ik heb een simplex model gebruikt om de genetische en niet-genetische oorzaken van de temporele stabiliteit van de data te analyseren. Gegeven deze uitkomsten bespreek ik vervolgens hoe de al bestaande longitudinale data het beste gebruikt kunnen worden in de context van *gene finding studies*. De hoge mate van stabiliteit van de gesommeerde genetische factoren geeft aan dat dezelfde set genen ten grondslag ligt aan de variatie in intelligentie tijdens de kinderjaren; dit rechtvaardigt het gebruik van een lineaire combinatie van scores van de verschillende leeftijdsgroepen voor genetisch associatieonderzoek.

De resultaten uit hoofdstuk 7 zijn gebruikt voor het modelleren van het fenotype in de associatiestudies van hoofdstuk 8 en 9. In hoofdstuk 8 wordt een onderzoek besproken waarin de associatie wordt getest tussen normale intelligentie en veelvoorkomende *single-nucleotide polymorphisms* (SNP's) in 43 genen die betrokken zijn bij intellectuele stoornissen. Het uitgangspunt was redelijk eenvoudig, namelijk dat de genetische variatie die van invloed is op continue, polygenetische eigenschappen (bijvoorbeeld normale intelligentie) op dezelfde plekken van het genoom kan liggen die ook verantwoordelijk zijn voor vergelijkbare monogenetische fenotypes (bijvoorbeeld intellectuele stoornissen). Hoewel er voor de afzonderlijke SNP's geen significanties werden gevonden, wezen de analyses met SNP's wel op een verrijking van de set van kandidaatgenen voor polymorfismes met betrekking tot intelligentie. Dit is het eerste onderzoek dat laat zien dat genen die betrokken zijn bij monogenetische aandoeningen met betrekking tot intellectueel functioneren ook normale intelligentie beïnvloeden.

In hoofdstuk 9 wordt het werk uit hoofdstuk 8 uitgebreid tot 168 genen die betrokken zijn bij intellectuele stoornissen, nu ook met gebruik van *next-generation exon sequencing*. Het gaat daarbij om de mogelijke effecten van zeldzame genetische variatie. In overeenstemming met de literatuur werd geen verrijking gevonden van de set van

kandidaatgenen voor mutaties die betrokken zijn bij normale intelligentie. Dit resultaat wordt besproken in de context van de literatuur.