

# VU Research Portal

## Measurement and Prediction in Heterogeneous Populations

Maij-de Meij, A.M.

2008

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Maij-de Meij, A. M. (2008). *Measurement and Prediction in Heterogeneous Populations*.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Samenvatting

In de sociale en gedragswetenschappen zijn veel attributen niet direct observeerbaar. Op basis van geobserveerde responses op bijvoorbeeld tests en vragenlijsten worden uitspraken gedaan over cognitieve capaciteiten, persoonlijkheid en attitudes. Met behulp van modellen met latente variabelen worden de associaties tussen de manifeste, geobserveerde, variabelen verklaard. Er is een algemeen raamwerk van modellen ontwikkeld waarbij de kans op een item response werd gerelateerd aan een onderliggende latente variabele. Deze latente variabele kan continu zijn of discreet. In de item response theorie (IRT) beschrijft een continue latente-trek variabele de kwantitatieve verschillen tussen personen. Kwalitatieve verschillen worden in kaart gebracht door een discrete latente-klasse variabele. Het centrale thema in dit proefschrift is de toepassing van mixture IRT modellen. Deze modellen voegen de latente-trek en latente-klasse variabele in één model samen. Er wordt een aantal homogene subgroepen, latente klassen, onderscheiden. Binnen deze latente klassen zijn verschillende latente-trek modellen van toepassing, in die zin dat de item parameterschattingen over klassen mogen verschillen. De mixture IRT modellen beschrijven de heterogeniteit binnen en tussen groepen. Op deze manier is het mogelijk om gelijktijdig te kijken naar kwantitatieve en kwalitatieve verschillen in antwoordgedrag van personen.

In Hoofdstuk 1 werd een aantal interpretaties van kwalitatieve verschillen gegeven. Er kunnen kwalitatieve verschillen voorkomen in het gemeten construct, zoals de situationele bepaaldheid van een eigenschap. Kwalitatieve verschillen kunnen ook ongerelateerd aan het gemeten construct zijn. Echter, invloeden op het antwoordgedrag die niet direct gerelateerd zijn aan de latente trek kunnen mogelijk wel inhoudelijk interessant zijn. Denk hierbij bijvoorbeeld aan antwoordtendensen en sociale wenselijkheid.

In Hoofdstuk 2 werd het construct "self-disclosure" nader onderzocht. Self-disclosure heeft betrekking om het delen van informatie over jezelf met anderen. Een mixture IRT

model met drie latente klassen werd geïdentificeerd als best passende model. De latente trek gaf de algemene neiging weer om meer of minder self-disclosure te tonen. De personen in de latente klassen verschilden in hun patroon van self-disclosure, afhankelijk van de ontvanger van de disclosure. Dit betekent dat self-disclosure situatie afhankelijk is, waarbij de situatie betrekking heeft op de persoon naar wie de disclosure plaats vindt. Hiermee werd het concept selectiviteit in self-disclosure geïntroduceerd. Het onderzoek heeft laten zien hoe mixture IRT modellen inzicht kunnen geven in de situationele bepaaldheid van persoonlijkheidseigenschappen.

Vervolgens werd in Hoofdstuk 3 een latente DIF-detectie methode vergeleken met een manifeste DIF-detectie methode. Er werd gekeken hoe de twee methoden presteerden wat betreft het identificeren van DIF items. De latente DIF-detectie methode gaat uit van het fitten van een mixture IRT model, waarna wordt gekeken welke item parameters significant verschillen over klassen. De gangbare manifeste DIF-detectie methoden onderzoeken het verschillend functioneren van items over manifeste groepen. Het voordeel van de latente DIF-detectie methode is dat vooraf geen aannames gedaan hoeven te worden over de bron van de bias. Echter, een manifeste variabele kan wel worden opgenomen in het model als indicator voor de latente klasse variabele. Dit biedt tevens de mogelijkheid om de relatie met de latente-klasse variabele te onderzoeken.

Met een simulatie studie werd de identificatie van DIF items onderzocht met beide DIF-detectie methoden. Gevarieerd werden steekproefgrootte en de sterkte van de relatie tussen de bron van de bias en de manifeste variabele. De resultaten lieten zien dat de manifeste DIF-detectie methode minder goed werkte naarmate de relatie zwakker werd tussen de manifeste variabele en de bron van de bias. De latente DIF-detectie methode deed het goed, ook in de condities waarin de relatie tussen de manifeste variabele en de bron van de bias zwakker was. De latente DIF-detectie methode, zonder manifeste variabele als indicator voor de latente-klasse variabele, deed het alleen beter in de conditie waar de correlatie tussen de bron van de bias en de manifeste variabele nul was. In de andere condities deed de latente DIF-detectie methode met de manifeste indicator het beter. De studie laat zien hoe mixture IRT modellen perspectieven bieden voor het verbeteren van DIF-detectie, met als bijkomend voordeel dat vooraf geen aannames gedaan hoeven te worden over de bron van de bias.

Kwalitatieve verschillen tussen de latente klassen kunnen ook optreden door een

verschillende betekenis van de response categorieën van de items. In Hoofdstuk 4 werden de schalen extraversie en neuroticisme onderzocht afkomstig van de Amsterdamse Biografische Vragenlijst (Wilde, ?). Aansluitend bij onderzoek van Smit, Kelderman en Van der Flier (?) werd er in deze studie een grotere dataset onderzocht, en werden geen restricties gelegd op het aantal te identificeren latente klassen. Voor beide persoonlijkheidschalen werden drie latente klassen gevonden. In overeenstemming met eerder onderzoek hadden de verschillen tussen de klassen allereerst te maken met het verschillend gebruik van de "?" categorie. Personen uit een van de latente klassen leken de response categorie te vermijden. Deze latente klasse kon tevens gekarakteriseerd worden door een lage neiging tot sociaal wenselijk antwoorden. Met name voor de neuroticisme schaal was een duidelijke voorkeur voor de "?" categorie waar te nemen voor personen uit een van de latente klassen. Deze latente klasse werd verder gekenmerkt door personen uit ethnische minderheidsgroepen, en een hogere neiging tot sociaal wenselijk antwoorden. Tenslotte bestond de derde latente klasse met name uit autochtone Nederlanders met een hogere neiging sociaal wenselijk te antwoorden.

Een belangrijke vraag was of mixture IRT modellen ook gebruikt kunnen worden om een betere voorspelling mogelijk te maken van iemands criterium score. Voor een deel van de steekproef in Hoofdstuk 4 waren criterium data aanwezig, die overeen kwamen met respectievelijk extraversie en emotionele stabiliteit. De voorspellende waarde van het mixture IRT model werd vergeleken met de voorspelling van het criterium onder de aanname van een een-klasse model. Uit de resultaten konden geen eenduidige conclusies getrokken worden wat betreft de waarde van mixture IRT modellen voor predictie van een extern criterium. Voor de extraversie schaal liet de trekschatting onder het een-klasse model hogere correlaties met het criterium zien dan de trekschatting onder het mixture IRT model. Voor de neuroticisme schaal lagen de resultaten andersom. Trekschattingen onder het mixture IRT model voorspelden het criterium beter dan schattingen onder aanname van een een-klasse model.

In Hoofdstuk 5 werd dieper ingegaan op de vraag of mixture IRT modellen een betere voorspelling van een criterium score mogelijk maken. Over het algemeen worden personen toegewezen aan latente klassen op basis van de hoogste kans gegeven het antwoordpatroon van de persoon. Vervolgens wordt de latente trekwaarde toegewezen, die voor die persoon voor de betreffende latente klasse geschat is. De zekerheid van toewijzing

aan een latente klasse kan echter sterk verschillen over personen. Daarom werd als alternatief een gewogen trekschatting voorgesteld. De klasse-specifieke trekschattingen werden in dat geval gewogen met de corresponderende latente-klasse kansen. In geval van hoge zekerheid van toewijzing (hoge klasse kans) zou het verschil tussen wegen en toewijzen klein moeten zijn. Dit verschil zou moeten toenemen naarmate de klasse kansen meer gelijk over de klassen verdeeld zijn, en/of wanneer de klasse-specifieke trekschattingen meer van elkaar verschillen over klassen.

Uit een simulatie studie bleek dat een gewogen latente trekschatting een hogere correlatie met de gesimuleerde latente trek en criterium score had dan de toegewezen latente trekschatting. In vergelijking met de gewogen latente trekschattingen, leverden trekschattingen op basis van een een-klasse model over het algemeen lagere correlaties op met gesimuleerde latente trek en criterium scores. Op basis van de resultaten onder verschillende condities werd geconcludeerd dat een gewogen trekschatting vooral beter voorspelt bij kortere tests en wanneer er grote verschillen in item parameters tussen de klassen zijn. Voor de twee persoonlijkheidsschalen uit Hoofdstuk 4 werden eveneens gewogen en toegewezen trekwaarden geschat. Hiervoor werden de klasse-specifieke trekschattingen eerst op eenzelfde schaal over klassen gebracht. Voor de extraversie schaal waren de correlaties van het criterium met de gewogen latente trekschattingen hoger dan de correlaties van het criterium met de toegewezen trekschattingen. Beide waren echter lager dan de trekschattingen op basis van een een-klasse model. Voor de neuroticisme schaal lagen de correlaties van de gewogen en toegewezen latente trekschattingen dicht bij elkaar. Alle correlaties op basis van het mixture IRT model waren hoger dan de trekschattingen op basis van het een-klasse model.

Kwalitatieve verschillen zijn vanuit verschillende perspectieven besproken, met name gericht op situationele bepaaldheid, DIF en antwoordtendensen. Verschillen in manifestaties van persoonlijkheidseigenschappen zijn in kaart te brengen. Dit is getoond voor self-disclosure. Het toepassen van mixture IRT modellen geeft tevens mogelijkheden om invloeden in kaart te brengen die niet direct betrekking hebben op de te meten latente trek. De resultaten van de latente DIF-detectie methode bieden goede perspectieven voor toekomstig onderzoek. Conclusies over de waarde van mixture IRT modellen als methode om schattingen van de latente trek en voorspelling van criterium scores te verbeteren kunnen nog niet eenduidig getrokken worden.