

# Voorwoord

De serie *Statistiek voor de psychologie* is, zoals de titel al zegt, een inleiding in de statistiek toegespitst op opleidingen psychologie. De serie bestaat uit vier delen. Het algemene uitgangspunt is dat studenten aan de opleidingen psychologie, die niet allereerst komen voor de statistiek, toch gemotiveerd worden om het vak statistiek met goed gevolg af te leggen. De serie doet dit door in 'gewone taal' statistische begrippen uit te leggen, maar ook door je met behulp van opdrachten te stimuleren zelfstandig met de stof om te gaan.

Meer concreet heeft de serie de volgende kenmerken. Elke analyse wordt behandeld volgens dezelfde structuur, het 'elementaire rapport'. Dat maakt het leren veel gemakkelijker. Je wordt systematisch getraind op: datadesign en vraagstelling; betekenis en formulering van de conclusies; causale interpretaties en hun beperkingen. Je leert daarmee niet alleen het 'hoe' maar ook het 'wanneer', 'welke' en 'waarom' van de analyses op een praktische manier. Bij veel analyses wordt uitgebreid aandacht besteed aan visualiseren: het 'lezen' en inzichtsmatig redeneren met figuren. De stof begint extreem gemakkelijk (hoe bereken je een gemiddelde) maar eindigt op hoog niveau (dubbel multivariate repeated-measures ANOVA). In de eerste twee delen wordt uitvoerig stilgestaan bij de beperkingen van het intuïtieve statistische denken. Er is ook veel aandacht voor praktische regels. De stof wordt uitgelegd met veel gevarieerde voorbeelden, van het dagelijks leven tot gepubliceerd onderzoek. Met name onderzoeken die prototypisch zijn voor een bepaald gebied worden gebruikt. Tot slot is de stof vijf jaar lang getest en verbeterd naar aanleiding van vragen, problemen en suggesties van studenten van de opleiding Psychologie van de KUN.

Op de productpagina van dit boek op [www.boomlemma.nl](http://www.boomlemma.nl) vind je onder andere de data van computeropgaven.

Bij dezen wil ik Maurits Geerts bedanken voor zijn bijdrage in het schrijven van de eerste versie van hoofdstuk 8. Daarnaast wil ik de student-assistenten bedanken die vanaf 1997 met hun commentaar, suggesties, tips en verbeteringen hebben bijgedragen aan dit boek. Ook veel studenten hebben hieraan bijgedragen, al was het maar door aan te geven welke passages ze niet snapten. Hun grootste bijdrage lag in het positieve commentaar. Het was altijd leuk om te horen dat studenten zeiden dat ze het boek duidelijk vonden, dat ze het vaak leuk vonden om te lezen, en dat ze mijn humor waardeerden. Dat stimuleerde me om ermee door te gaan.

# Inhoud

HOOFDSTUK 1	Design en analyse	11
1.1	Specificatie van designs	13
1.2	Definities	14
1.3	Het verschil tussen een afhankelijke variabele en een niveau van een within-subjectfactor	19
1.4	Kiezen van de analyse	22
1.5	Andere voorbeelden	26
1.6	Bronnen van variatie	30
1.7	Beknopt rapporteren en toelichten van conclusies	31
1.7.1	Beknopt rapporteren	31
1.7.2	Toelichten van conclusies	32
1.8	Controle van assumpties	33
1.9	Opgaven	34
HOOFDSTUK 2	Multiple-regressieanalyse (MRA)	37
2.1	Inleiding	37
2.1.1	Herhaling <i>Statistiek voor de psychologie, deel 1B</i>	38
2.2	Samenvatting	39
2.3	Kort voorbeeld van een MRA	40
2.4	Elementair rapport van een MRA	43
2.4.1	Design	44
2.4.2	Mate van controle	44
2.4.3	Geaggregeerde data	45
2.4.4	Schatters: de regressiegewichten	46
2.4.5	De hypothesen	47
2.4.6	Toetsing: de ANOVA summary table	49
2.4.7	Toetsing: de $p$ -waarden van de $b$ -gewichten	54
2.4.8	Beslissingen	54
2.4.9	Causale interpretatie	58
2.4.10	Samenvatting: elementair rapport MRA	59
2.5	Beknopt rapport van een MRA	61
2.6	Aansturen van SPSS <i>Regression</i>	62
2.7	Output van SPSS <i>Regression</i> (selectie)	66
2.8	Betekenis van de output van SPSS <i>Regression</i>	67

2.9	Het adolescentieonderzoek	69
2.10	Opgaven	72
HOOFDSTUK 3	Varianten van MRA	79
3.1	Gestandaardiseerde MRA	79
3.1.1	Samenvatting	79
3.1.2	Definitie van gestandaardiseerde regressiegewichten	80
3.1.3	Berekenen van gestandaardiseerde regressiegewichten in SPSS	83
3.1.4	Het effect van verandering van meeteenheid	83
3.1.5	Het voordeel van gestandaardiseerde regressie- gewichten	85
3.1.6	Het voordeel van ongestandaardiseerde regressie- gewichten	86
3.1.7	Causale interpretatie van regressiegewichten	86
3.1.8	De voorspellende waarde van een predictor	89
3.2	ANOVA als MRA met dummycodering	90
3.2.1	Dummycodering voor één factor	91
3.2.2	MRA met de dummycodering van één factor	94
3.2.3	Dummycodering van twee factoren	95
3.3	Meer over regressie	96
3.3.1	Hiërarchische regressie	96
3.3.2	Mediatieanalyse	97
3.3.3	Moderatie	97
3.4	Opgaven	99
HOOFDSTUK 4	GLM met één afhankelijke variabele	101
4.1	Inleiding	101
4.2	Samenvatting	102
4.3	Korte voorbeelden van GLM-Univariate	103
4.3.1	Alleen between-subjectfactoren: ANOVA	103
4.3.2	Alleen covariaten: MRA	104
4.3.3	Factoren en covariaten: ANCOVA	105
4.4	Elementair rapport van een ANCOVA	107
4.4.1	Design	107
4.4.2	Mate van controle	108
4.4.3	Geaggregeerde data	108
4.4.4	Hypothesen	109
4.4.5	Toetsing	113
4.4.6	Beslissingen	116
4.4.7	Causale interpretatie	117
4.4.8	Samenvatting: elementair rapport van een ANCOVA	118
4.5	Beknopt rapport van een ANCOVA of ANOVA	120

	4.6	Aansturen van SPSS <i>GLM-Univariate</i>	121
	4.7	Output van SPSS <i>GLM-Univariate</i> (selectie)	123
	4.8	Betekenis van de output van SPSS <i>GLM-Univariate</i>	124
	4.9	Meer over GLM-Univariate	126
	4.10	Opgaven	127
HOOFDSTUK	5	GLM met meerdere afhankelijke variabelen	131
	5.1	Inleiding	131
	5.2	Samenvatting	131
	5.3	Kort voorbeeld van GLM-Multivariate	132
	5.4	Elementair rapport van een MANOVA	135
	5.4.1	Design	136
	5.4.2	Mate van controle	137
	5.4.3	Geaggregeerde data	137
	5.4.4	Hypothesen	138
	5.4.5	Toetsing	143
	5.4.6	Beslissingen	145
	5.4.7	Causale interpretatie	147
	5.4.8	Samenvatting: elementair rapport van een MANOVA	147
	5.5	Beknopt rapport van een MANOVA	150
	5.6	Aansturen van SPSS <i>GLM-Multivariate</i>	151
	5.7	Output van SPSS <i>GLM-Multivariate</i> (selectie)	153
	5.8	Betekenis van de output van SPSS <i>GLM-Multivariate</i>	154
	5.9	Opgaven	157
HOOFDSTUK	6	GLM met een within-subjectfactor	161
	6.1	Inleiding	161
	6.2	Samenvatting	162
	6.3	Kort voorbeeld van een repeated-measures-MANOVA	163
	6.4	Elementair rapport van een repeated-measures-MANOVA	166
	6.4.1	Design	167
	6.4.2	Mate van controle	168
	6.4.3	Geaggregeerde data	169
	6.4.4	Hypothesen	170
	6.4.5	Toetsing	175
	6.4.6	Beslissingen	178
	6.4.7	Causale interpretatie	180
	6.4.8	Samenvatting: elementair rapport van een repeated-measures-MANOVA	183
	6.5	Beknopt rapport van een repeated-measures-MANOVA	185
	6.6	Aansturen van SPSS <i>GLM-Repeated Measures</i>	186
	6.7	Output van SPSS <i>GLM-Repeated Measures</i> (selectie)	191
	6.8	Betekenis van de output van <i>GLM-Repeated Measures</i>	192

6.9	Meer over repeated measures	194
6.10	Opgaven	197
HOOFDSTUK 7	GLM met within-subjectfactoren en meerdere afhankelijke variabelen	205
7.1	Inleiding	205
7.2	Samenvatting	206
7.3	Kort voorbeeld van een dubbel multivariate repeated- measures-ANOVA	206
7.4	Elementair rapport van een dubbel multivariate repeated-measures-ANOVA	208
7.4.1	Design	208
7.4.2	Mate van controle	209
7.4.3	Geaggregeerde data	209
7.4.4	Hypothesen	210
7.4.5	Toetsing	211
7.4.6	Beslissingen	212
7.4.7	Causale interpretatie	213
7.4.8	Samenvatting: elementair rapport van een dubbel multivariate repeated-measures-ANOVA	214
7.5	Beknopt rapport van een dubbel multivariate repeated-measures-ANOVA	216
7.6	Aansturen van SPSS <i>GLM-Repeated Measures</i>	217
7.7	Output van SPSS <i>GLM-Repeated Measures</i> (selectie)	220
7.8	Betekenis van de output van <i>GLM-Repeated Measures</i>	221
7.9	Opgaven	223
HOOFDSTUK 8	Non-parametrische toetsen	225
8.1	Inleiding	225
8.1.1	Wat zijn non-parametrische toetsen?	226
8.1.2	Wanneer een non-parametrische toets?	226
8.1.3	Waarom een non-parametrische toets?	226
8.2	Elementair rapport van een Mann-Whitney U-toets	229
8.2.1	Design	229
8.2.2	Mate van controle	230
8.2.3	Hypothesen	230
8.2.4	Geaggregeerde Data	231
8.2.5	Toetsingsgrootte en $p$ -waarde	232
8.2.6	Beslissing	232
8.2.7	Causale interpretatie	232
8.2.8	Samenvatting	233
8.2.9	De aansturing van SPSS	233
8.2.10	De uitvoer van SPSS	234

8.3	Elementair rapport van een $\chi^2$ -toets	234
8.3.1	Design	235
8.3.2	Mate van controle	235
8.3.3	Hypothesen	236
8.3.4	Geaggregeerde data	236
8.3.5	Toetsingsgrootheid, <i>df</i> en <i>p</i> -waarde	238
8.3.6	Beslissing	239
8.3.7	Causale interpretatie	239
8.3.8	Samenvatting	239
8.3.9	Aansturing van SPSS	240
8.3.10	De uitvoer van SPSS	241
8.4	Opgaven	242
HOOFDSTUK 9	Leerdoelen en zelftoetsen	245
9.1	Leerdoelen	245
9.2	Zelftoets 1	246
9.3	Uitwerkingen van zelftoets 1	253
9.4	Zelftoets 2	259
9.5	Uitwerkingen van zelftoets 2	268
Appendix	Overzicht van elementaire en beknopte rapporten	271
A.1	Elementair rapport MRA	271
A.2	Beknopt rapport MRA	273
A.3	Elementair rapport van een ANCOVA	273
A.4	Beknopt rapport van een ANCOVA	275
A.5	Elementair rapport van een MANOVA	276
A.6	Beknopt rapport van een MANOVA	278
A.7	Elementair rapport van een repeated-measures-manova	279
A.8	Beknopt rapport van een repeated-measures-MANOVA	281
A.9	Elementair rapport van een dubbel multivariate repeated-measures-ANOVA	282
A.10	Beknopt rapport van een dubbel multivariate repeated-measures-ANOVA	284
A.11	Elementair rapport van een Mann-Whitneytoets	285
A.12	Elementair rapport van een chi-kwadraattoets	286
Referenties		287
Register		289

# 1 Design en analyse

## *Achtergrond 1*

Een bekend probleem van scriptiestudenten is dat zij ‘verdrinken’ in hun data. Een beetje onderzoek bevat al snel tientallen variabelen en onderzoeken met honderden variabelen zijn bepaald geen uitzondering. In dit deel zullen steeds meer en steeds complexere data worden behandeld. De complexiteit van de analyses houdt daar gelijke tred mee. Om overzicht te behouden is het van het hoogste belang dat je bij elk onderzoek snel het design kan doorzien. Het design is je kompas bij het kiezen van een analyse.

## *Doel 1*

Na bestudering van dit hoofdstuk en het maken van de opgaven ben je in staat om, gegeven de beschrijving van de data en vraagstelling van een onderzoek, het **design** te specificeren en de juiste **analyse** te kiezen. Aan het einde van de cursus moet je dit proces hebben geautomatiseerd (= snel en zonder opzoeken).

Zie paragraaf 1.7 voor achtergrond 2 en doel 2.

## *Doorlopend voorbeeld*

Stel dat zestien personen een nieuwe taal moeten leren. Zij worden toegewezen aan drie verschillende groepen, die verschillende onderwijsmethoden krijgen: controle (geen onderwijs), methode A en methode B. Ieder subject wordt in vijf verschillende sessies getest, met tussenperiodes van een maand. Zij worden elke sessie getest op drie taken: lezen, schrijven en luisteren. Op elke sessie worden andere tests gebruikt om hun prestaties te meten. De onderzoeker heeft er echter voor gezorgd dat de vijf testjes voor lezen (R1 tot en met R5) alle even moeilijk zijn, dat de vijf testjes voor schrijven (W1 tot en met W5) even moeilijk zijn en dat de vijf testjes voor luisteren (L1 tot en met L5) even moeilijk zijn. Voor elke persoon worden geslacht en leeftijd vastgesteld aan het begin van het experiment. De data staan in tabel 1.1. Deze data zullen worden beschouwd met verschillende vraagstellingen.





## 1.1 Specificatie van designs

Bij het specificeren van het design van een onderzoek beschrijf je:

- de **afhankelijke variabelen** en hun meetniveaus (= meestal kwantitatief);
- de **within-subjectfactoren** en hun niveaus;
- de **between-subjectfactoren** en hun niveaus;
- de **covariaten** en hun meetniveaus (= kwantitatief);
- als er within-subjectfactoren zijn de **indeling van de metingen**.

*Covariaten* zijn kwantitatieve, onafhankelijke variabelen waarvan wordt onderzocht of zij correleren met een afhankelijke variabele. De *indeling van de metingen* is een tabel waarin voor elke afhankelijke variabele en elke cel van de within-subjectfactoren staat wat de naam is van de betreffende meting (= kolom) in de data-matrix. De overige begrippen zijn al behandeld.

In het design hoeven alleen variabelen te worden opgenomen die **relevant** zijn voor de vraagstelling. Deze variabelen moeten **consistent en uitputtend** worden geclassificeerd door het design.

Het design geeft aan wat de rol is van elke variabele in de analyse. Dat wordt bepaald door de **aard van de data** en door de **vraagstelling**.

Bij de afhankelijke variabele hoef je (anders dan in deel 3) niet op te schrijven hoeveel metingen er per persoon zijn. Die informatie staat als het goed is al in de indeling van de metingen. Het kan overigens geen kwaad het er toch bij te schrijven!

### *Toelichting*

In het design beschrijf je van elke SPSS-variabele de functie in de analyse. Je moet dus om te beginnen weten welke SPSS-variabelen er zijn! Het is bijna gênant om te zien hoeveel mensen deze stap over blijven slaan. Maar we zijn met statistiek bezig, dus het *gaat* om de data!

KIJK IN DE DATAMATRIX WAT DE VARIABELEN ZIJN.

Data kunnen vaak op meerdere manieren in een design worden gezet. Het hangt dan van de onderzoeksvraag af wat je moet doen. Bijvoorbeeld: eenzelfde variabele kan soms evengoed tot afhankelijke als tot onafhankelijke variabele worden gebombardeerd.

LEES GOED WAT DE ONDERZOEKSVRAAG IS.

### *Voorbeeld*

In het doorlopend voorbeeld is een zinnige onderzoeksvraag: worden lezen, schrijven en luisteren beïnvloed door de onderwijsmethode, geslacht, leeftijd, sessie, of interacties daartussen? Het design is in dat geval:

afhankelijke variabele 1	= Read (kwantitatief)
afhankelijke variabele 2	= Write (kwantitatief)
afhankelijke variabele 3	= Listen (kwantitatief)
within-subjectfactor 1	= Sessie (1, 2, 3, 4, 5)
between-subjectfactor 1	= Methode (controle, A, B)
between-subjectfactor 2	= Geslacht (man, vrouw)
covariaat 1	= Leeftijd (kwantitatief)

indeling van de metingen = (zie tabel 1.2)

*Tabel 1.2*

<i>Afhankelijke variabele</i>	<i>Sessie</i>	<i>Meting</i>
Read	1	R1
	2	R2
	3	R3
	4	R4
	5	R5
Write	1	W1
	2	W2
	3	W3
	4	W4
	5	W5
Listen	1	L1
	2	L2
	3	L3
	4	L4
	5	L5

Dit is trouwens één van de meest ingewikkelde voorbeelden die we tegen zullen komen.

## 1.2 Definities

Hoewel de meeste begrippen al behandeld zijn, geven we hieronder voor de zekerheid een lijst definities.

### *Datamatrix*

Een rechthoekige tabel met gegevens. In SPSS wordt hij getoond in de Data Editor.

Voorbeeld: Alles wat binnen de tabel in het wit staat.

*Subject/case*

Een rij (horizontaal) in de datamatrix. Meestal een persoon of dier.

Voorbeeld: Persoon 1-16.

*Variabele*

Een kolom (verticaal) in de datamatrix. Gewoonlijk bestaat een variabele uit scores die zijn verkregen met een of ander meetinstrument in een bepaalde conditie.

Voorbeeld: Methode, Geslacht, Leeftijd, R1-R5, W1-W5, L1-L5.

*Groep*

Een verzameling subjecten met een gemeenschappelijk kenmerk.

Voorbeeld: De groep subjecten met Methode = A is {Persoon 6, Persoon 7, Persoon 8, Persoon 9}.

*Nominale/categoriale/kwalitatieve variabele*

Een variabele waarvan het bereik (de mogelijke uitkomsten) bestaat uit slechts een beperkt aantal verbale labels: de categorieën. De categorieën moeten wederzijds uitsluitend zijn. De variabele kan zijn gecodeerd met scores die geen numerieke betekenis hebben, maar slechts categorieën symboliseren. De scores zijn meestal gehele getallen tussen 0 en 10 en indiceren een groepslidmaatschap zoals sekse, nationaliteit, ras, experimentele groep. Kenmerkend voor zo'n gecodeerde kwalitatieve variabele is dat je net zo goed een andere codering kunt gebruiken, zonder verlies van informatie. Een vuistregel om kwalitatieve variabelen te herkennen: bij zulke variabelen is het onzinnig om het gemiddelde te berekenen.

Voorbeeld: Methode, Geslacht. Merk op dat we Methode net zo goed hadden kunnen coderen met de getallen 1, 2, 3 in plaats van Controle, A, B.

*Interval-/continuel/kwantitatieve variabele*

Een variabele met scores die numerieke betekenis hebben. Vaak hebben de scores een decimale punt of hebben ze een groot bereik (bijvoorbeeld 0-40). Een kwantitatieve variabele meet een hoeveelheid – dat wil zeggen: een eigenschap die het subject 'meer' of 'minder' kan hebben en waarin het subject 'groot' of 'klein' is (bijvoorbeeld lengte, leeftijd, angst, responstijd). Vuistregel: bij een kwantitatieve variabele is het zinnig om het gemiddelde te berekenen.

Voorbeeld: Leeftijd, R1-R5, W1-W5, L1-L5.

*Onafhankelijke/verklarende variabele*

Een variabele waarvan wordt onderzocht of hij invloed heeft op een andere varia-

bele. Meestal is dat een kenmerk van een situatie (bijvoorbeeld experimentele conditie) of een kenmerk van de persoon bij aanvang van het onderzoek.

Voorbeeld: Methode, Geslacht, Leeftijd.

#### *Factor/between-subjectfactor*

Een **kwalitatieve, onafhankelijke variabele**. Een factor deelt de subjecten in in groepen en die groepen moeten met elkaar worden vergeleken.

Voorbeeld: Methode, Geslacht. De niveaus van de between-subjectfactor Methode zijn Controle, A, B. Elk niveau definieert een groep subjecten.

#### *Covariaat/predictor*

Een **kwantitatieve, onafhankelijke variabele**. In de gangbare analyses wordt aangenomen dat de afhankelijke variabele lineair samenhangt met de covariaat. Als er sterke redenen zijn om dat te betwijfelen, kan de betreffende variabele beter worden behandeld als between-subjectfactor dan als covariaat. Daartoe kan het nodig zijn het bereik van de variabele terug te brengen tot een klein aantal niveaus, bijvoorbeeld 'hoog', 'midden', 'laag'.

Of men de term covariaat of predictor gebruikt, hangt af van de analyse die men doet. Dat is niet logisch. Bij een ANCOVA gebruikt men de term 'covariaat'. Bij een multiple-regressieanalyse gebruikt men de term 'predictor'.

Voorbeeld: Leeftijd.

#### *Within-subjectfactor*

Een **kwalitatieve, onafhankelijke 'variabele' die verschillende waarden aanneemt bij hetzelfde subject**. Dit treedt op als een variabele onder verschillende condities is gemeten bij dezelfde subjecten. De within-subjectfactor geeft aan in welke conditie er is gemeten. Je spreekt pas van een within-subjectfactor als het de bedoeling is metingen uit de verschillende condities met elkaar te **vergelijken** door te toetsen of zij hetzelfde gemiddelde hebben.

In principe kunnen data die met een within-subjectfactor zijn verkregen op twee manieren worden weergegeven: 'onder elkaar' en 'naast elkaar'. Neem als voorbeeld de Readscore van Personen 1 en 2 tijdens Sessie 1 en 2 uit tabel 1.1. Onder elkaar ziet het er uit als in tabel 1.3.

*Tabel 1.3*

<i>Persoon</i>	<i>Sessie</i>	<i>Score</i>
1	1	1
1	2	2
2	1	4
2	2	4

Naast elkaar ziet het er uit als in tabel 1.4.

Tabel 1.4

Persoon	R1	R2
1	1	2
2	4	4

In SPSS moet je de data van een within-subjectfactor in principe **naast elkaar** zetten. Dat leidt tot de complicatie dat de **within-subjectfactor en de afhankelijke variabele niet als variabele in de datamatrix** staan. In het zojuist gegeven voorbeeldje zijn alleen Persoon, R1 en R2 variabelen in de datamatrix. Sessie en Read ontbreken. Zij kunnen pas bij de analyse worden gedefinieerd.

Wie wel als variabelen in de datamatrix staan, zijn de **metingen**. Dat is een variabele die de scores bevat op een afhankelijke variabele bij een specifieke cel van de within-subjectfactor(en). In het voorbeeldje zijn R1 en R2 de metingen. In SPSS heten dit ‘within-subject variables’ – niet te verwarren met de within-subjectfactor. Zoals een between-subjectfactor een indeling van subjecten is, zo is een within-subjectfactor **een indeling van de metingen**. Je moet een within-subjectfactor dan ook *definiëren* door aan te geven hoe hij de metingen indeelt. Dat wil zeggen: door aan te geven welke variabele uit de datamatrix hoort bij welk niveau van de within-subjectfactor. Een within-subjectfactor kan dan ook worden weergegeven als een *rij* scores of labels boven de datamatrix. Dit is gedaan bij de data van het doorlopend voorbeeld (tabel 1.1). *Conceptueel* is een within-subjectfactor een ‘variabele’. *Officieel* is een within-subjectfactor echter zelf geen variabele; een variabele is immers een kolom, geen rij.

Voorbeeld: Sessie. De niveaus van deze within-subjectfactor zijn 1, 2, 3, 4, 5. Deze condities moeten met elkaar worden vergeleken. Wat betreft de within-subjectfactor Sessie worden de variabelen ingedeeld als in tabel 1.5.

Tabel 1.5

Sessie	Variabelen
1	R1, W1, L1
2	R2, W2, L2
3	R3, W3, L3
4	R4, W4, L4
5	R5, W5, L5

#### *Afhankelijke of verklaarde variabele/criterium/measure*

Een variabele waarvan wordt onderzocht of hij wordt beïnvloed door een andere variabele. Vaak is de afhankelijke variabele een meting van gedrag, beleving of

prestatie aan het eind van het onderzoek. Afhankelijk van de overige elementen in het design heeft een afhankelijke variabele de volgende kenmerken.

- *Bij een between-subjectfactor.* De afhankelijke variabele is een variabele waarvan meerdere gemiddelden (of medianen, of frequentieverdelingen) worden berekend, namelijk in een gemiddelde in elke groep van de between-subjectfactor. Die gemiddelden worden dan met elkaar vergeleken. De nulhypothese is dat die gemiddelden gelijk zijn.
- *Bij een covariaat.* De correlatie tussen de covariaat en de afhankelijke variabele wordt berekend. De nulhypothese is dat die correlatie 0 is.
- *Bij een within-subjectfactor.* Een afhankelijke variabele is een verzameling variabelen waarvan de gemiddelden met elkaar worden vergeleken. De nulhypothese is daarbij dat al die gemiddelden gelijk zijn. In afwijking van de afspraak dat een ‘variabele’ een kolom in de datamatrix is, wordt hier toegelaten dat een afhankelijke variabele bestaat uit meerdere kolommen en dus eigenlijk een *verzameling variabelen* is. De betreffende kolommen noemen we **metingen** van de afhankelijke variabele.

Om variabelen te zien als metingen van één afhankelijke variabele is het nodig dat zij ‘commensurabel’ zijn, dat wil zeggen dat zij met een soortgelijk meetinstrument zijn verkregen. Om deze reden wordt een afhankelijke variabele die uit meerdere metingen bestaat in SPSS aangeduid als een ‘measure’. Dit woord dien je echter te vertalen als ‘maat’, niet als ‘meting’. Dus een maat is een afhankelijke variabele die bestaat uit meerdere metingen. Elk van die metingen bestaat uit een kolom scores. In de literatuur worden de begrippen ‘dependent variable’, ‘measure’ en ‘measurement’ niet consistent gebruikt.

Voorbeeld: Read, Write, Listen. Elke van deze afhankelijke variabelen bestaat uit meerdere, commensurabele metingen. Read bestaat bijvoorbeeld uit R1, R2, R3, R4 en R5. Wat betreft deze afhankelijke variabelen worden de metingen ingedeeld als in tabel 1.6.

*Tabel 1.6*

<i>Afhankelijke variabele</i>	<i>Meting</i>
Read	R1, R2, R3, R4, R5
Write	W1, W2, W3, W4, W5
Listen	L1, L2, L3, L4, L5

#### *Indeling van de metingen*

Een tabel waarin voor elke afhankelijke variabele bij elke cel van de within-subjectfactoren staat wat de naam is van de betreffende meting (= kolom) in de datamatrix.

Bij designs met een within-subjectfactor is de indeling van de metingen in SPSS nodig om de afhankelijke variabelen en de within-subjectfactoren eenduidig te

definiëren. Als de variabelen in de datamatrix bijvoorbeeld de namen Read1 en Read2 hebben, dan snapt SPSS niet dat dit respectievelijk de eerste en de tweede meting van Read zijn. ‘Read, wie is dat, ik ken alleen Read1 en Read2’. Dat lijkt stom, maar toch is het goed. Het hangt namelijk maar helemaal af van je vraagstelling of je de gemiddelden van Read1 wilt vergelijken met die van Read2. Als je ze niet wilt vergelijken, vormen ze *twee* afhankelijke variabelen *zonder* within-subjectfactor. Als je ze wel wilt vergelijken vormen ze *één* afhankelijke variabele *met* een within-subjectfactor. Het is prettig dat SPSS niet ongevraagd jouw vraagstelling bedenkt.

Voorbeeld: zie tabel 1.7.

*Tabel 1.7*

<i>Afhankelijke variabele</i>	<i>Sessie</i>	<i>Meting</i>
Read	1	R1
	2	R2
	3	R3
	4	R4
	5	R5
Write	1	W1
	2	W2
	3	W3
	4	W4
	5	W5
Listen	1	L1
	2	L2
	3	L3
	4	L4
	5	L5

Hiermee worden de twee eerder gegeven indelingen gecombineerd.

### 1.3 Het verschil tussen een afhankelijke variabele en een niveau van een within-subjectfactor

Bij designs met een within-subjectfactor wordt een afhankelijke variabele vaak verward met een niveau van een within-subjectfactor. Wat is het verschil? Het verschil zit hem uitsluitend in de **vraagstelling**. Je moet nagaan welke metingen volgens de vraagstelling met elkaar **vergeleken** moeten worden. Met ‘vergelijken’ bedoelen we hier dat er wordt getoetst of ze hetzelfde gemiddelde hebben.

Metingen waarvan de gemiddelden met elkaar moeten worden vergeleken, behoren tot **dezelfde afhankelijke variabele** en bij **verschillende niveaus van een within-subjectfactor**.

### *Voorbeeld*

Beschouw de variabelen R1 tot en met R5. Dat zijn vijf variabelen. Je kan ze op twee manieren in het design gebruiken:

- als vijf verschillende afhankelijke variabelen;
- als één afhankelijke variabele met een within-subjectfactor van vijf niveaus.

Als je voor het eerste kiest, dan zul je in de bijbehorende analyse antwoord krijgen op de vraag of Methode invloed heeft op R1, of Methode invloed heeft op R2, enzovoort. Als je voor het tweede kiest, dan zul je in de bijbehorende analyse antwoord krijgen op de vraag of de gemiddelden van R1 tot en met R5 verschillend zijn en of Methode invloed heeft op de somscore  $R1 + R2 + R3 + R4 + R5$ . Het kan niet allebei tegelijk. Om het design te specificeren moet je dus besluiten op welke van deze vragen je het liefste antwoord krijgt.

Als we de data van tabel 1.1 analyseren volgens het in paragraaf 1.1 gespecificeerde design met drie afhankelijke variabelen (Read, Write, Listen) en één within-subjectfactor (Sessie) dan geldt:

Er wordt getoetst of '3.8 = 4.2 = 5.3 = 4.5 = 4.1 maar dan in de populatie'. Deze gemiddelden gaan over dezelfde afhankelijke variabele, Read. Zij komen uit verschillende niveaus van de within-subjectfactor, Sessie.

Er wordt *niet* getoetst of '4.4 = 5.8 = 6.4 maar dan in de populatie'. Deze gemiddelden gaan over verschillende afhankelijke variabelen. Er wordt ook niet getoetst of '3.8 = 5.1 = 6.2 maar dan in de populatie'. Ook deze gemiddelden hebben betrekking op verschillende afhankelijke variabelen.

Als we de gemiddelden op Read, Write en Listen toch met elkaar willen vergelijken *dan moeten we het design anders specificeren*, namelijk:

afhankelijke variabele	= Taalvaardigheid (kwantitatief)
within-subjectfactor 1	= Taak (Read, Write, Listen)
within-subjectfactor 2	= Sessie (1, 2, 3, 4, 5)
between-subjectfactor 1	= Methode (controle, A, B)
between-subjectfactor 2	= Geslacht (man, vrouw)
covariaat 1	= Leeftijd (kwantitatief)
indeling van de metingen	= (zie tabel 1.8)



Tabel 1.8

Taak	Sessie	Meting
Read	1	R1
	2	R2
	3	R3
	4	R4
	5	R5
Write	1	W1
	2	W2
	3	W3
	4	W4
	5	W5
Listen	1	L1
	2	L2
	3	L3
	4	L4
	5	L5

Read, Write en Listen zijn nu niveaus van een within-subjectfactor in plaats van afhankelijke variabelen. Dit leidt tot:

Er wordt getoetst of '4.4 = 5.8 = 6.4 maar dan in de populatie'. Dit zijn nu immers metingen van dezelfde afhankelijke variabele bij verschillende niveaus van een within-subjectfactor. Ook wordt getoetst of '3.8 = 5.1 = 6.2 maar dan in de populatie'.

Als je even doordent, moet je concluderen dat we bij dezelfde data minstens *vier* verschillende designs hebben. In het eerste design worden R1 tot en met L5 beschouwd als vijftien verschillende afhankelijke variabelen zonder within-subjectfactor. In het tweede design (dat ik in paragraaf 1.1 specificerde) worden Read, Write en Listen beschouwd als drie verschillende afhankelijke variabelen met een within-subjectfactor Sessie van vijf niveaus (en  $3 \times 5 = 15$ ). In het derde design worden Read, Write en Listen beschouwd als drie niveaus van de within-subjectfactor Taak, die gekruist is met de within-subjectfactor Sessie van vijf niveaus. En dan is er nog de vierde mogelijkheid, om de sessies te zien als vijf afhankelijke variabelen met de within-subjectfactor Taak. Voor welk je design je moet kiezen, ligt niet besloten in de data, maar in de vraagstelling. Als je wilt weten of Read, Write en Listen hetzelfde gemiddelde hebben, en ook of de vijf sessies hetzelfde gemiddelde hebben, moet je kiezen voor het derde design.

In dit voorbeeld ligt het tweede design wat meer voor de hand dan het derde of vierde. Read, Write en Listen zijn gemeten met verschillende meetinstrumenten (ze zijn niet commensurabel). Er is geen moeite gedaan de testjes voor Read even

moeilijk te maken als die voor Write en Listen, of om ze anderszins vergelijkbaar te maken. Het is daarom weinig zinnig om te toetsen of ze hetzelfde gemiddelde hebben, want waarschijnlijk hebben ze dat niet, maar wat zegt dat? Ook uit de vraagstelling blijkt niet dat ze vergeleken moeten worden. Eerder is het zo dat ze naast elkaar bestudeerd moeten worden als verschillende afhankelijke variabelen.

## 1.4 Kiezen van de analyse

Op grond van de specificatie van het design kun je kiezen welke analyse moet worden gedaan. Op den duur moet je dat automatisch kunnen. In het begin kun je gebruikmaken van figuur 1.1, figuur 1.2 en figuur 1.3. Die lijken misschien wat ingewikkeld, maar de plattegrond van het centrum van jouw woonplaats is moeilijker. Net als het centrum van steden is de naamgeving van de diverse analyses historisch gegroeid.

Je hoeft voorlopig alleen figuur 1.1 te kennen. Figuren 1.2 en 1.3 zijn gegeven voor de volledigheid, misschien heb je er later nog wat aan. Alleen de *bovenste twee parten* van figuur 1.1 zijn belangrijk (25% en 50% grijs). De onderste twee parten (10% en 0% grijs) staan er voor de volledigheid, om de relaties met eerder behandelde stof aan te geven.

In de figuur worden analyses benoemd met gebruikelijke afkortingen, die jij ook mag gebruiken: ANOVA = analysis of variance; ANCOVA = analysis of covariance; MANOVA = multivariate ANOVA; GLM = general linear model. Wat dat allemaal betekent, daar hebben we het nog wel eens over. Leer de volgende associaties.

*eerste helft design* ⇒ *eerste helft analysenaam*

meerdere afhankelijke variabelen	⇒ multivariaat
within-subjectfactoren	⇒ multivariaat + repeated measures
<b>beide</b>	⇒ dubbel multivariaat + repeated measures

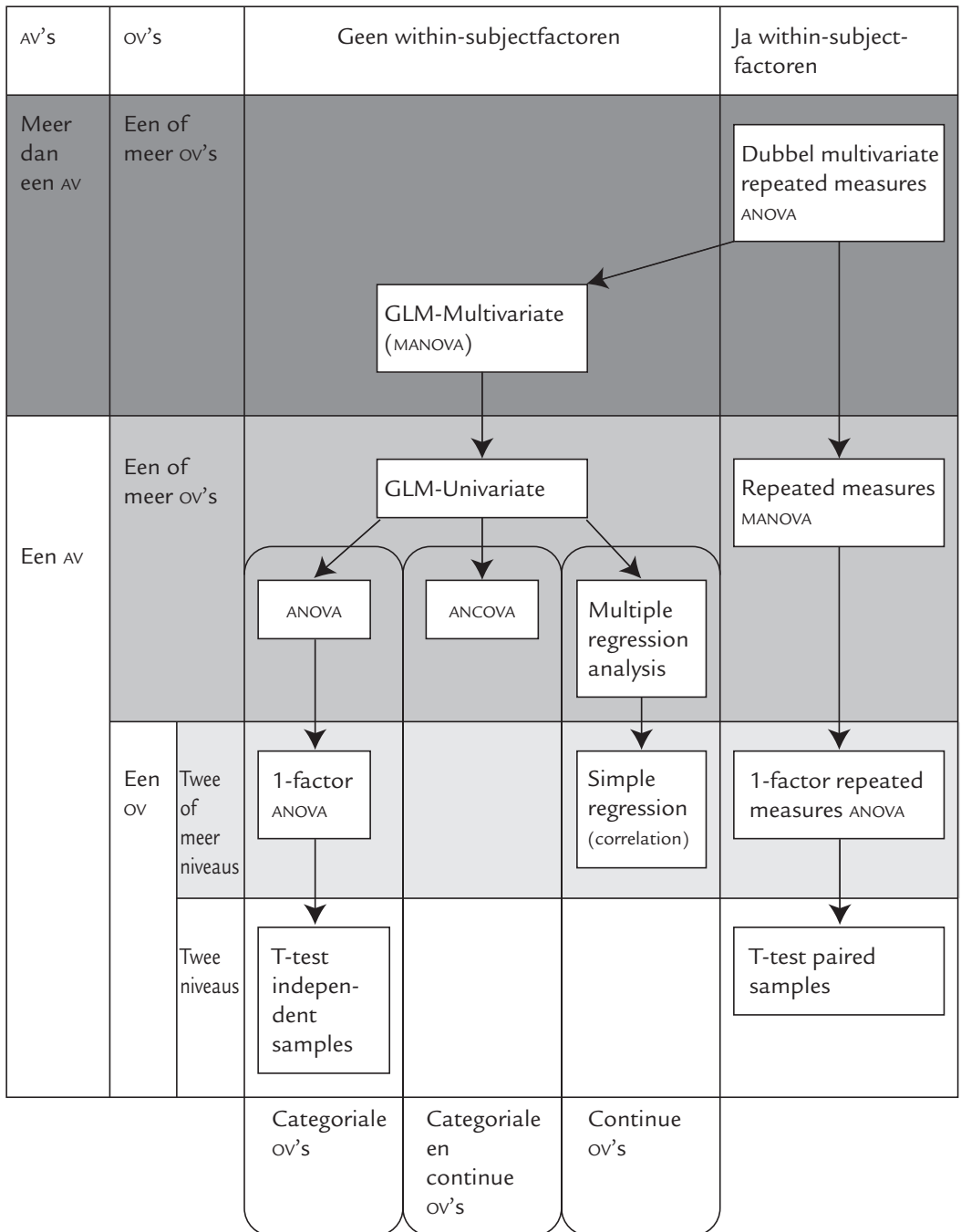
*tweede helft design* ⇒ *tweede helft analysenaam*

between-subjectfactoren	⇒ ANOVA
covariaten	⇒ multiple-regressieanalyse
<b>beide</b>	⇒ ANCOVA

### *Voorbeeld*

In het doorlopende voorbeeld zijn de afhankelijke variabelen kwantitatief, dus we moeten figuur 1.1 hebben. Er zijn meerdere afhankelijke variabelen, dus we moeten de bovenste helft hebben. Er is een within-subjectfactor, dus we moeten de rechterhelft hebben. In dit gebied staat nog maar één analyse, dus die moeten we hebben: ‘doubly multivariate repeated measures ANOVA’. Via de bovenstaande associatieregels kom je op de naam ‘dubbel multivariate repeated-measures-ANCOVA’. Dat is ook goed, nog beter zelfs.

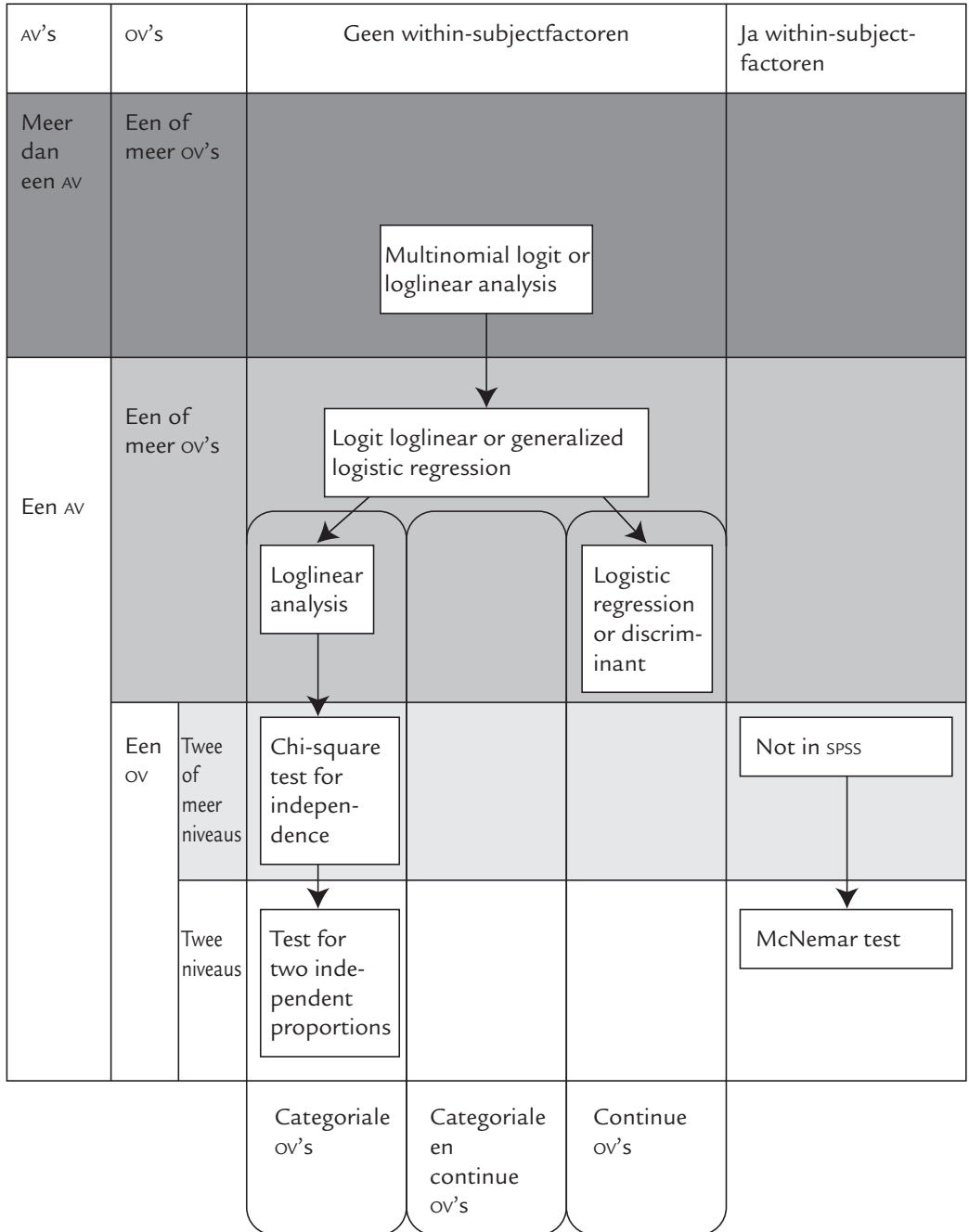
*Figuur 1.1 Analyses voor continue afhankelijke variabelen (DV = dependent variable; IV = independent variable)*



*Figuur 1.2 Analyses voor ordinale afhankelijke variabelen*

AV's	ov's	Geen within-subjectfactoren			Ja within-subjectfactoren
Meer dan een AV	Een of meer ov's				
Een AV	Een of meer ov's	<div style="border: 1px solid black; padding: 5px; text-align: center;">Not in SPSS</div>			
	Een ov	Twee of meer niveaus	<div style="border: 1px solid black; padding: 5px; text-align: center;">Kruskal-Wallis</div>	<div style="border: 1px solid black; padding: 5px; text-align: center;">Spearman rank correlation</div>	<div style="border: 1px solid black; padding: 5px; text-align: center;">Friedman repeated measures</div>
		Twee niveaus	<div style="border: 1px solid black; padding: 5px; text-align: center;">Mann-Whitney U</div>		<div style="border: 1px solid black; padding: 5px; text-align: center;">Wilcoxon matched pairs or Signs test</div>
		Categoriale ov's	Categoriale en continue ov's	Continue ov's	

*Figuur 1.3 Analyses voor categoriale afhankelijke variabelen*



## 1.5 Andere voorbeelden

### *Voorbeeld 1*

Zie het doorlopend voorbeeld van deel 3 van deze serie, hoofdstuk 1. Van een groep kinderen werd de sociale isolatie en de treiterstatus bepaald. Sociale isolatie werd gemeten met een schaal. Qua treiterstatus werden de kinderen ingedeeld in 'non-involved', 'bully' of 'victim'. Deze drie groepen werden met elkaar vergeleken op gemiddelde sociale isolatie. De data zien er uit als in tabel 1.9.

*Tabel 1.9*

<i>Kind</i>	<i>Treiterstatus</i>	<i>Sociale isolatie</i>
1	Non-involved	1.75
2	Bully	0.75
3	Non-involved	0.25
4	Victim	1.00
5	Victim	2.00
...	...	...

Het design is daarom:

afhankelijke variabele = Sociale isolatie (kwantitatief)  
 between-subjectfactor = Treiterstatus (bully, victim, non-involved)

De aangewezen analyse is ANOVA (onderdeel van SPSS GLM-Univariate).

### *Voorbeeld 2*

Zie voorbeeld 1. Stel dat daarnaast van elk kind het gymnastiekcijfer werd bepaald. Men vraagt zich af of dit eveneens door treiterstatus wordt beïnvloed. De data zien er uit als in tabel 1.10.

*Tabel 1.10*

<i>Kind</i>	<i>Treiterstatus</i>	<i>Sociale isolatie</i>	<i>Gymnastiek</i>
1	Non-involved	1.75	7
2	Bully	0.75	8
3	Non-involved	0.25	7
4	Victim	1.00	6
5	Victim	2.00	5
...	...	...	...

Het design is dan:

afhankelijke variabele 1	= Sociale isolatie (kwantitatief)
afhankelijke variabele 2	= Gymnastiekcijfer (kwantitatief)
between-subjectfactor	= Treiterstatus (bully, victim, non-involved)

Er is hier sprake van twee verschillende afhankelijke variabelen omdat sociale isolatie en gymnastiekcijfer naast elkaar worden bestudeerd. Zij worden niet met elkaar vergeleken. Dat wil zeggen: er wordt niet onderzocht of het gemiddelde van sociale isolatie gelijk is aan het gemiddelde gymnastiekcijfer. Weliswaar is het technisch mogelijk om dat toch te doen, maar conceptueel zou dat onzin zijn: sociale isolatie en gymnastiekcijfer zijn niet commensurabel. Daarom zijn het twee verschillende afhankelijke variabelen.

De aangewezen analyse is MANOVA (onderdeel van SPSS GLM-Multivariate).

### *Voorbeeld 3*

Zie voorbeeld 2. Bij deze data kan men evengoed een andere vraag stellen: hebben de gymnastische prestaties van een kind invloed op de sociale isolatie? Daarnaast blijven we geïnteresseerd in het effect van treiterstatus op sociale isolatie. Het design is dan:

afhankelijke variabele	= Sociale isolatie (kwantitatief)
between-subjectfactor	= Treiterstatus (bully, victim, non-involved)
covariaat	= Gymnastiekcijfer (kwantitatief)

De data zijn hetzelfde, maar de vraagstelling is anders. Daarom is het design anders. Gymnastiekcijfer wordt nu opgevat als onafhankelijke variabele in plaats van afhankelijke variabele. Het gevolg is dat de drie groepen van treiterstatus niet meer zullen worden vergeleken op gymnastiekcijfer. In plaats daarvan wordt bepaald wat de correlatie is tussen gymnastiekcijfer en sociale isolatie.

De aangewezen analyse is ANCOVA (onderdeel van SPSS GLM-Univariate).

### *Voorbeeld 4*

Stel dat men treiterstatus meet op een schaal van 1 tot 20, waarbij alle tussenliggende gehele getallen mogelijk zijn. De bullies scoren hoog, de victims laag, de non-involved midden. Stel dat men daarnaast van elk kind een maat voor ego-sterkte verkrijgt en dat op deze variabele de scores 0.00 (= laag), 0.05, 0.10, 0.15, ..., 0.95, 1.00 (= hoog) mogelijk zijn. Men vraagt zich af of treiterstatus en ego-sterkte invloed hebben op sociale isolatie. De data zien er uit als in tabel 1.11.

Tabel 1.11

Kind	Treiterstatus	Sociale isolatie	Egosterkte
1	10	1.75	0.35
2	18	0.75	0.90
3	12	0.25	0.65
4	5	1.00	0.55
5	6	2.00	0.75
...	...	...	...

Het design is dan:

afhankelijke variabele	= Sociale isolatie (kwantitatief)
covariaat 1	= Treiterstatus (kwantitatief)
covariaat 2	= Egosterkte (kwantitatief)

Treiterstatus is in dit geval kwantitatief in plaats van kwalitatief en daarom een covariaat in plaats van een between-subjectfactor.

De aangewezen analyse is multiple-regressieanalyse (in SPSS mogelijk met *GLM-Univariate*, maar je krijgt gedetailleerdere output met *Regression*).

#### Voorbeeld 5

Zie voorbeeld 1. Stel dat men ook een meting van sociale isolatie krijgt na twee jaar en na vier jaar. Noem de drie metingen van sociale isolatie respectievelijk SI1, SI2 en SI3. De vraag is nog steeds of treiterstatus invloed heeft op sociale isolatie, maar daarnaast is de vraag ook of het gemiddelde van sociale isolatie verandert in de loop van de tijd. De data zien er uit als in tabel 1.12.

Tabel 1.12

Kind	Treiterstatus	SI1	SI2	SI3
1	Non-involved	1.75	1.50	1.75
2	Bully	0.75	1.00	1.25
3	Non-involved	0.25	0.50	0.50
4	Victim	1.00	0.75	1.25
5	Victim	2.00	1.50	1.25
...	...	...	...	...



Het design is dan:

afhankelijke variabele	= Sociale isolatie (kwantitatief)
within-subjectfactor	= Tijd (direct, na 2 jaar, na 4 jaar)
between-subjectfactor	= Treiterstatus (bully, victim, non-involved).
indeling van metingen	= (zie tabel 1.13)

Tabel 1.13

Afhankelijke variabele	Tijd	Meting
Sociale isolatie	Direct	SI1
	Na 2 jaar	SI2
	Na 4 jaar	SI3

De drie variabelen SI1, SI2 en SI3 zijn commensurabel: er kan zinvol worden getoetst of hun gemiddelde gelijk is. Bovendien is het, zoals blijkt uit de vraagstelling, de bedoeling dat inderdaad te doen. *Daarom vormen SI1, SI2 en SI3 samen één afhankelijke variabele met een within-subjectfactor van drie niveaus.* Deze afhankelijke variabele noemen we Sociale isolatie en de within-subjectfactor noemen we Tijd. Die namen staan niet in de datamatrix. We hebben ze zelf verzonnen. Om duidelijk te maken wat we er precies mee bedoelen, geven we de indeling van de metingen.

De aangewezen analyse: repeated-measures-MANOVA (onderdeel van SPSS GLM-Repeated Measures).

#### Voorbeeld 6

Zie voorbeeld 5. Stel dat de drie metingen van sociale isolatie zijn gedaan met totaal verschillende schalen, die ook een andere meeteenheid en range hebben. Het is dan niet meer zinvol om te onderzoeken of de gemiddelden van SI1, SI2 en SI3 gelijk zijn. Het antwoord op die vraag zou namelijk meer zeggen over je meetprocedures dan over het concept sociale isolatie. De data zien er uit als in tabel 1.14.

Tabel 1.14

Kind	Treiterstatus	SI1	SI2	SI3
1	Non-involved	1.75	150	17.5
2	Bully	0.75	100	12.5
3	Non-involved	0.25	050	05.0
4	Victim	1.00	075	12.5
5	Victim	2.00	150	12.5
...	...	...	...	...

Het design is dan:

afhankelijke variabele 1	= SI1 (kwantitatief)
afhankelijke variabele 2	= SI2 (kwantitatief)
afhankelijke variabele 3	= SI3 (kwantitatief)
between-subjectfactor	= Treiterstatus (bully, victim, non-involved)

Merk op dat de datamatrix er vrijwel hetzelfde uitziet als in het vorige voorbeeld. In dit geval zijn de variabelen SI1, SI2 en SI3 echter drie verschillende afhankelijke variabelen zonder within-subjectfactor.

De aangewezen analyse is MANOVA (onderdeel van SPSS GLM-Multivariate).

## 1.6 Bronnen van variatie

Tot nu toe is er in deze serie geprobeerd om enigszins consequent te zijn in het gebruik van termen. Maar in deze wrede wereld lopen veel mensen rond die statistiek gebruiken en die geen boodschap hebben aan jullie of mij. En die andere termen gebruiken voor hetzelfde, of dezelfde termen voor iets anders. Dat gebeurt ook in SPSS, dus daar zul je mee moeten leven.

In deel 3 ben je bij ANOVA al verscheidene ‘bronnen van variatie’ tegengekomen: Between, Within, Total. Dit onderscheid bestaat ook bij geavanceerdere analyses. Er worden echter veel verschillende namen voor gebruikt, afhankelijk van de voorkeur van de auteur en de analyse in kwestie. We zetten ze even op een rijtje. Termen die onder elkaar staan hebben min of meer dezelfde betekenis. Hoe lager, hoe algemener de betekenis van de term.

*Tabel 1.15*

<i>Between</i>	<i>Within</i>	<i>Total</i>
Regression	Residu	Total
Hypothesis	Error	Total
Model	Error	Total
Explained	Unexplained	Observed

Verder heeft SPSS tegenwoordig de onaangename tendens om de term Total te gebruiken op een manier die afwijkt van vrijwel alle statistiekboeken. Als er in SPSS ‘Corrected Total’ staat, bedoelen ze daarmee de normale Total, en de SPSS ‘Total’ is onzin die je moet negeren voor onze toepassingen.

## 1.7 Beknopt rapporteren en toelichten van conclusies

### *Achtergrond 2*

Tot nu toe waren de boeken in deze serie opgebouwd rond elementaire rapporten. Die zijn deels een didactisch hulpmiddel om de stof te structureren. Zo zullen ze ook in dit deel nog worden gebruikt. In dit deel gaan we een stap verder: Je moet het rapport kunnen samenvatten voor de resultatensectie van een artikel. Dat zullen we een ‘beknopt rapport’ noemen. In de voorafgaande delen deden we soms al moeilijk over de precieze formulering, maar vanaf nu komen we pas echt los.

In sommige artikelen wordt bij de beslissingen ook genoemd hoe groot de gemiddelden zijn waar de toets over gaat. Dat zullen we ‘toelichten’ noemen. Het geeft ons een fraaie gelegenheid om te controleren of je weet waar je het over hebt.

### *Doel 2*

Na bestudering van deze paragraaf weet je wat er in de rest van de cursus wordt bedoeld met een ‘beknopt rapport’ en het ‘toelichten’ van conclusies.

### 1.7.1 Beknopt rapporteren

Bij deze spreken we af: een **beknopt rapport** is een **lopend verhaal** met de **essentiële conclusies** van de analyse, dat **zelfstandig leesbaar** is. Het omvat:

- de specificatie van het **design** dat is gebruikt;
- de naam van de het soort **analyse** dat is gedaan;
- de belangrijkste elementen van de **toetsing** (dat zijn  $F$ ,  $dfs$ ,  $p$ -waarde en  $R^2$  of  $eta^2$  – bijvoorbeeld ‘ $F(1,11) = 6.56, p = .03, R^2 = .374$ ’);
- de **beslissingen** in woorden; hierbij moeten de termen ‘significant’ en ‘non-significant’ worden gebruikt in plaats van ‘in de populatie’.

### *Toelichting*

In een beknopt rapport noem je in het algemeen *niet*: de mate van controle (gaat over de methode), de nulhypothesen (de lezer wordt geacht statistiek te hebben gehad en dat zelf te kunnen bedenken), formules (idem), de causale interpretatie (hoort in de discussie), dat je SPSS hebt gebruikt (irrelevant, het mag ook SAS zijn). Bij ongebruikelijke analyses mag hiervan worden afgeweken.

In artikelen worden de waarden van  $R^2$  of  $eta^2$  vaak niet genoemd. Bij het schrijven van je scriptie kan het daarom goed gebeuren dat je begeleider zegt dat dit ‘in de praktijk niet gebruikelijk’ is. Maar mensen die het kunnen weten, zeggen al jaren met goede redenen dat je het zeker moet doen (Cohen, 1988; Tabachnick & Fidell, 2001a, 2001b; Aron & Aron, 1994). Daarom moet je in deze cursus aanleren om het wel te doen. En je zou de wetenschap een dienst bewijzen als je te zijner tijd je scriptiebegeleider ervan kan overtuigen dat die ‘gebruiken’ dan verkeerd zijn.

In het hele boek zal ik hanteren dat ik de grenzen van ‘matig’ en ‘groot’ leg bij

een  $R^2$  van respectievelijk .10 en .20. Cohen (1988) hanteert iets andere grenzen, die verschillend zijn voor MRA en ANOVA, wat ik hier wil vermijden.

Tot nu toe heb je geleerd bij de beslissingen in een elementair rapport gebruik te maken van de frase 'in de populatie'. Daarmee gaf je een letterlijke verwoording van de hypothese die werd geaccepteerd. In artikelen is dat echter ongebruikelijk. Daar gebruikt men bij het verwoorden de termen 'significant' en 'non-significant' in plaats van 'in de populatie'. Dat laatste *moet* je in een beknopt rapport ook doen, want een beknopt rapport is bestemd voor een artikel. Bij een elementair rapport mag je beide formuleringen gebruiken, hoewel de voorkeur hier vanaf nu uitgaat naar de termen 'significant' en 'non-significant'. Maak niet de fout dat je schrijft 'de gemiddelden verschillen significant in de populatie'. Dat is onzin, want 'significant' betekent 'hoogstwaarschijnlijk generaliseerbaar naar de populatie'. Je zou dus voor jezelf de regel kunnen hanteren dat vanaf nu de term 'in de populatie' moet worden vermeden in het verwoorden van de beslissingen. Let wel: dit geldt alleen voor de beslissingen. Wanneer we praten over de betekenis van hypothesen is het nog steeds onvermijdelijk om de frase 'in de populatie' te gebruiken.

(Weetje: van de APA moet voorafgaande aan een decimale punt een cijfer worden geschreven, behalve als het getal tussen -1 en 1 *moet* liggen; dus  $F = 0.67$  maar  $p = .67$ .)

#### *Voorbeeld van een 'beknopt rapport'*

Een 1-weg-ANOVA op de data van jongens van 8 tot 11 jaar met treiterstatus (victim, bully, non-involved) als between-subjectfactor en sociale isolatie als afhankelijke variabele, liet zien dat er een significant hoofdeffect was van treiterstatus op sociale isolatie ( $F(2,39) = 7.60, p < .01$ ). Dit effect is sterk ( $R^2 = .28$ ).

### 1.7.2 Toelichten van conclusies

Het toelichten van conclusies houdt in dat je bij de conclusie van een toets bespreekt wat de **richting** van het effect is, en daarbij de **steekproefwaarden** vermeldt van de parameters waar die toets iets over zegt. Dat wordt in het algemeen alleen gedaan als de toets een significant resultaat leverde (Tabachnick & Fidell, 2001a, 2001b; Aron & Aron, 1994). Dat zal ik ook hier als standaard hanteren. Specifieker:

- Bij een significant hoofdeffect van een factor noem je op zijn minst de twee gemiddelden die het meest van elkaar verschillen. Bijvoorbeeld: 'De buikdansers verdienden gemiddeld het meest ( $m = 200$  euro/uur) en de statistiekleraren het minst ( $m = 5$  euro/uur).'
- Bij een significante interactie noem je op zijn minst de twee verschillen die het meest van elkaar verschillen. Bijvoorbeeld: 'De wiskundigen schreven thuis meer bladzijden ( $m = 5$ ) dan op hun werk ( $m = 1.5$ ) en bij de Beotiërs was dat andersom (respectievelijk  $m = 0$  en  $m = 40$ ).'
- Bij een significante covariaat noem je het betreffende *b*- of *bèta*-gewicht en de

betekenis van de richting. Bijvoorbeeld: ‘Een toename in de productie ging samen met een afname in de kwaliteit ( $\beta = -0.5772$ ).’

*Voorbeeld van een ‘beknopt rapport met toelichting van conclusies’*

(Het gedeelte toelichting is gecursiveerd.)

Een 1-weg-ANOVA op de data van jongens van 8 tot 11 jaar met treiterstatus (victim, bully, non-involved) als between-subjectfactor en sociale isolatie als afhankelijke variabele, liet zien dat er een significant hoofdeffect was van treiterstatus op sociale isolatie ( $F(2,39) = 7.60, p < .01$ ). Dit effect is sterk ( $R^2 = .28$ ). *De slachtoffers hadden gemiddeld de hoogste sociale isolatie ( $M = 1.50$ ) en de pestkoppen de laagste ( $M = 0.50$ ). De niet-betrokken jongens zaten daartussenin ( $M = 0.73$ ).*

## 1.8 Controle van assumpties

In alle procedures die onder GLM vallen, wordt normaalverdeeltheid van de residuen verondersteld. Bij analyses met covariaten wordt bovendien lineariteit verondersteld. Vanzelfsprekend moet worden onderzocht of aan deze assumpties is voldaan, voordat men met de resultaten naar buiten treedt. Hoe men dat controleert zal hier niet worden besproken. In deze fase van de studie is het naar mijn mening eerst van belang dat men begrijpt welke vragen met een analyse kunnen worden beantwoord. Pas dan zal men interesse kunnen opbrengen voor het controleren van de assumpties. Daarom verwijst ik hiervoor naar het boek van Tabachnick & Fidell (2001b).