

PSYCHOSOCIALE ARBEIDSBELASTING
EN WERKSTRESS

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Veldhoven, Marcus Johannes Petrus Maria van

Psychosociale arbeidsbelasting en werkstress / Marcus
Johannes Petrus Maria van Veldhoven. - Lisse: Swets &
Zeitlinger
Proefschrift Rijksuniversiteit Groningen, - Met lit. opg.
- Met samenvatting in het Engels.
ISBN 90-265-1477-8
NUGI 685
Trefw.: werkbelasting / beroepsstress.

Omslagontwerp: Ivar Hamelink
Druk omslag: Casparie, IJsselstein
Druk: Grafisch Productiebedrijf Gorter B.V., Steenwijk

© Copyright 1996, Marc van Veldhoven en Swets & Zeitlinger B.V., Lisse

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd bestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of op enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16B Auteurswet 1912 j^o het Besluit van 20 juni 1974 St.B.351, zoals gewijzigd bij het Besluit van 23 augustus 1985, St.b. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 882, 1180 AW Amstelveen). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) dient men zich tot de uitgever te wenden.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

ISBN 90 265 1477 8
NUGI 685

Rijksuniversiteit Groningen

**PSYCHOSOCIALE ARBEIDSBELASTING
EN
WERKSTRESS**

Proefschrift

ter verkrijging van het doctoraat in de
Psychologische, Pedagogische en Sociologische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus dr. F. van der Woude
in het openbaar te verdedigen op
donderdag 4 april 1996
des namiddags te 4.00 uur

door

Marcus Johannes Petrus Maria van Veldhoven

geboren op 22 november 1963
te Sittard

Promotores:

Prof. dr. T.F. Meijman

Prof. dr. T.A.B. Snijders

Stellingen

1. Er bestaat geen ontwerp van arbeidstaken en -organisaties dat volledige preventie van werkstress garandeert. (*Dit proefschrift*)
2. Het is te hopen dat de critici van zelfbeoordelingsmethoden zelf hun beoordeling van de waarde van zelfbeoordelingsmethoden aanpassen naar aanleiding van objectieve resultaten over de waarde van dergelijke methoden.
3. Het construeren van zelfbeoordelingsschalen met item-response technieken komt neer op het ijken van de omgangstaal.
4. Arbeidsorganisaties zijn geneste systemen. Dit maakt observaties van individuen binnen groepen binnen organisaties per definitie afhankelijk, in de statistische betekenis van het woord. Het is vreemd te moeten constateren dat bij het grootste deel van het onderzoek dat tot nu toe op het gebied van arbeid en organisatie is verricht hiermee geen rekening is gehouden. In feite moet alles opnieuw.
5. De hoeveelheid verklaarde variantie in individueel ziekteverzuim op basis van sociaal-wetenschappelijke variabelen is teleurstellend klein. Binnen deze verklaarde variantie is de bijdrage van variabelen op het gebied van psychosociale arbeidsbelasting en werkstress bemoedigend groot. (*Dit proefschrift*)
6. Individuele problemen op het gebied van welbevinden zijn besmettelijker dan individuele problemen op het gebied van spanning. (*Dit proefschrift*)
7. Negatieve affectiviteit behoort geen ongemeten variabele te blijven in onderzoek naar de relatie tussen psychosociale arbeidsbelasting en werkstress.
8. Iedereen die de resultaten van werkbelevingsonderzoek van nu af aan interpreteert in termen van algemene onvrede, kan worden verweten de analyse van de problematiek te vertroebelen en hiermee de oplossing ervan tegen te houden. (*Dit proefschrift*)
9. Het bestaan van non-lineaire relaties tussen psychosociale arbeidsbelasting en werkstress stemt somber over de mogelijkheid tot de ontwikkeling van normen voor de arbeidsbelasting van individuele medewerkers. Voor groepen medewerkers lijken dergelijke normen wel mogelijk.
10. Wat de bloeddrukmeting is voor de arts, is de vragenlijstmeting voor de psycholoog.

(*Stellingen behorende bij het
proefschrift van Marc van Veldhoven,
"Psychosociale arbeidsbelasting en werkstress"*)

Voorwoord

Acht jaar geleden vertrok ik van de universiteit naar een bedrijfsgezondheidsdienst (BGD). Psychologen waren indertijd zeldzaam binnen de BGD-en. Ik leerde veel van de medische en technische disciplines in het veld, maar het was moeilijk om kennis en ervaringen uit te wisselen met vakgenoten. Bij gebrek aan passend opleidingsaanbod besloot ik om wetenschappelijk onderzoek als middel tot persoonlijke ontwikkeling te gebruiken, met het schrijven van een proefschrift als toen nog vage ambitie.

Na verloop van tijd gebeurde het, dat ik met Frank van Dijk en Theo Meijman werkte aan een artikel over het Periodiek Bedrijfsgezondheidkundig Onderzoek (PBGO). De strekking van het artikel was dat er in bedrijven in de loop van drie à vier jaar (het gebruikelijke interval tussen twee PBGO's bij hetzelfde bedrijf) zoveel gebeurt, dat een goede longitudinale analyse slechts beperkt mogelijk is. Tijdens de gesprekken over dit artikel ontstond de gedachte om een proefschrift te schrijven dat meer aansloot bij mijn oorspronkelijke vakgebied. Zo begon ik vier jaar geleden aan het slaan van een brug tussen de praktijk van een BGD en de psychologie, met het (modulair) PBGO als vehikel. Ik had natuurlijk gewaar-schuw moeten zijn.

Nu, vier jaar later, bestaan BGD-en niet meer. BGD-en werden arbodiensten, en de wijzigingen in de wetgeving over arbeidsomstandigheden, ziekteverzuim en arbeidsongeschiktheid hebben de kern van de dienstverlening fundamenteel gewijzigd. De interesse voor psychosociale arbeidsbelasting en werkstress is fors toegenomen omdat deze onderwerpen raken aan de kern van de nieuwe arbo- en verzuimdienstverlening. In de huidige situatie van elkaar beconcurrerende arbodiensten is het onmogelijk een dergelijk onderwerp te laten liggen.

Het PBGO bestaat ook niet meer, is overcompleet geraakt na de wettelijke verplichting tot Risico-Inventarisatie en -Evaluatie (RI&E) en Periodiek Arbeidsgezondheidkundig Onderzoek (PAGO).

Een vierde kerndiscipline deed zijn intrede in de arbodiensten: de arbeids- en organisatiedeskundige. Contact met vakgenoten is veel beter mogelijk dan enkele jaren geleden, bijvoorbeeld via de opleiding tot en de beroepsvereniging voor arbeids- en organisatiedeskundigen.

Het goed uitvoeren van een langdurig project is in de afgelopen vier jaar inderdaad beperkt mogelijk gebleken: doel, context en voorwaarden zijn ondertussen volledig veranderd. Deze veranderingen hebben de nodige tussentijdse koersaanpassingen gevergd. Zo komt het dat dit geen proefschrift is over de ontwikkeling van een vragenlijstmodule in het kader van het PBGO, maar een proefschrift over het meten van psychosociale arbeidsbelasting en werkstress met vragenlijsten in zijn algemeenheid.

Dit alles zou ik niet hebben kunnen bereiken zonder de hulp van velen. Enkelen wil ik hier expliciet noemen.

In de eerste plaats is de inbreng van mijn promotor, Theo Meijman, bij het totstandkomen van dit proefschrift onmisbaar geweest. Hij bracht niet alleen jaren van ervaring in vragenlijstontwikkeling en arbeidspsychologie met zich mee. Ook een groot deel van de in dit proefschrifttraject ontwikkelde vragenlijst (de VBBA) is qua inhoud te danken aan zijn voorarbeid. Ik dank hem echter vooral voor de belangstelling, steun, creativiteit en humor die hij heeft getoond gedurende zijn promotiebegeleiding. Theo, ik voel me bevoorrecht met jou te hebben kunnen samenwerken.

Gedurende het afgelopen jaar gaf Tom Snijders, mijn tweede promotor, een nadrukkelijke methodologische wending aan de inhoud van dit proefschrift. Tom, het uiteindelijke resultaat (zeker wat betreft het derde deel van de tekst) zou zonder jouw inbreng absoluut niet zijn wat het nu is.

Werken aan een proefschrift vergt de nodige voorwaarden. Ik prijs me gelukkig dat de directie van de Arbodienst West-Brabant in Breda mij gedurende een zo lange periode in de gelegenheid heeft gesteld om een deel van mijn tijd hieraan te besteden. Zonder Jos van Collenburg en Philip Koster van Groos zou dit proefschrift nooit van de grond zijn gekomen. Ook de mogelijkheid om het proefschrift op deze manier uit te geven dank ik aan de Arbodienst West-Brabant. Jos en Philip, hartelijk dank voor al jullie steun en belangstelling.

Een belangrijke bijdrage aan dit proefschrift is geleverd door mijn collega's van de Arbodienst West-Brabant. Ik dank allen die zorgden voor contacten met bedrijven en instellingen, allen die zorgden voor verzameling, invoer en verwerking van vragenlijstgegevens, allen die zorgden voor het aanleveren van literatuur, en allen die mij hielpen bij het maken van de vragenlijst en de tekst. Dit proefschrift is een beetje van jullie allemaal.

Twee personen binnen de Arbodienst West-Brabant verdienen speciale vermelding voor hun bijdrage in de laatste fase van dit proefschrift. Carla Schellings sleepte mij door de onvermijdelijke laatste loodjes die een proefschrift met zich meebrengt heen. Carla, hartelijk dank voor je betrokkenheid en steun. Ik dank Heleen Geerards voor de deskundige ondersteuning bij het vormgeven van de definitieve tekst. Heleen, hartelijk dank dat je hiervoor de benodigde tijd wilde vrijmaken. Verder wil ik Margreet Dijkstra en Marjan de Kwaasteniet bedanken, die als stagiaires een bijdrage leverden aan de totstandkoming van dit proefschrift.

Buiten de Arbodienst West-Brabant verdient vooral de inbreng van twee Amsterdamse instituten hier vermelding. Ik dank Rick Fortuin en iedereen bij de Stichting Kwaliteitsbevordering Bedrijfsgezondheidszorg (SKB) voor alle mogelijkheden en stimulansen om in de afgelopen jaren te werken aan instrumentontwikkeling. Beste Rick, zonder jou en jouw mensen zou ik niet zoveel hebben kunnen leren als ik in de afgelopen jaren heb kunnen doen.

Dit proefschrift kreeg een forse startimpuls doordat ik twee jaar in de gelegenheid was om via het Senior Research Plan van het Nederlands Instituut voor Arbeidsomstandigheden (NIA) te werken aan de ontwikkeling van een vragenlijst voor het meten van psychosociale arbeidsbelasting en werkstress. Beste medewerkers van de (ondertussen opgeheven) afdeling onderzoek, bedankt voor jullie belangstelling tijdens de Amsterdam-periode van dit proefschrift.

Bij het NIA leerde ik Jaap Roorda kennen, die mij en Theo Meijman hielp bij de publicatie van een boek over de VBBA. Jaap, bedankt voor je bereidheid om dit proefschrift uit te geven via je huidige werkgever Swets en Zeitlinger.

Dit proefschrift had verder nooit kunnen bestaan zonder de bereidheid van bedrijven en instellingen, om mij onderzoek te laten doen met de VBBA en de voorlopers daarvan. Ik dank alle managers, medewerkers en P&O-ers voor de geboden medewerking en de constructieve commentaren.

Verder gaat mijn dank uit naar Jeannette, Marjolein en Jeroen. Mijn werken aan dit proefschrift is voor jullie niet eenvoudig geweest. Ik dank jullie voor de mogelijkheid om deze ambitie toch te kunnen vervullen.

Het grootste deel van dit proefschrift schreef ik niet in Amsterdam of Breda, maar bij mijn ouders in Sittard. PC's, boeken, artikelen en alle andere benodigde materialen sleepte ik voortdurend heen en weer tussen de diverse locaties. Waren de andere werkplekken hectisch, mijn ouders boden mij een rustige plek waar ik ongestoord kon schrijven. De steun en verzorging die zij mij in de afgelopen periode boden, is onbetaalbaar. Uit dank hiervoor draag ik dit proefschrift aan hen op.

Sittard, februari 1996

Voor mijn ouders.

Inhoudsopgave

Voorwoord

1.	Inleiding	1
1.1	Doel, belang en aanpak	1
1.2	Vraagstellingen	1
1.3	Begrippen en definities	2
1.4	Overzicht van de inhoud	5

DEEL I: Theorie

2	Psychosociale arbeidsbelasting	7
2.1	Sociotechniek	7
2.2	Handelingstheorie	8
2.3	Taakkenmerken-benadering	9
2.4	Michigan-model	10
2.5	Demand-control model	11
2.6	Vitamine-model	13
2.7	Belasting-belastbaarheid benadering	15
2.8	Integraal model Beehr en Newman	18
2.9	Vertretpunt voor de vragenlijstconstructie	20
2.10	Fundamentele methodologische verschillen tussen de theorieën	21
3	Het ontstaan van psychologische stress	25
3.1	Zes theorieën over het ontstaan van psychologische stress	25
3.1.1	Stress als respons	25
3.1.2	Stress als stimulus	26
3.1.3	Individuele verschillen-benaderingen	27
3.1.4	Transactioneel model	28
3.1.5	Zelfregulatie-model	29
3.1.6	Conservation of resources-model	30
3.1.7	Betekenis van de algemene stresstheorie voor het veld PSA en werkstress	31
3.2	Sociale vergelijkingstheorie	32
3.3	Symptoomperceptie-theorie	34
3.4	Betekenis van de algemene stresstheorie voor de vragenlijstconstructie	38

9	De voorspellende waarde van PSA en werkstress voor ziekteverzuim	163
9.1	Het voorspellen van ziekteverzuim	163
9.2	Vorbereiding van de analyse	166
9.3	Resultaten model 1 en 2	169
9.4	Resultaten model 3	172
9.5	Samenvatting en conclusie over het voorspellen van ziekteverzuim	173
10	Discussie	177
10.1	De structuur van PSA en werkstress	177
10.2	De relatie tussen PSA en werkstress	180
10.3	De voorspelling van ziekteverzuim	186
10.4	Integratie: een multi-level model voor de beleving en beoordeling van PSA en werkstress	188
10.5	Beperkingen van het onderzoek	191
10.6	Betekenis voor de praktijk	193
10.7	Slot deel III	201
	Literatuur	203
	Sammenvatting	215
	Summary	223

Bijlagen

1. LISREL-modellen hoofdstuk 7
2. LISREL-modellen hoofdstuk 9

Hoofdstuk 1: Inleiding

1.1 Doel, belang en aanpak

Dit proefschrift wil een bijdrage leveren aan praktijk en onderzoek op het gebied van psychosociale arbeidsbelasting en de gevolgen daarvan voor werknemers. Het maatschappelijk belang van dit thema is groot (Knepper, 1991; Smulders en Bloemhoff, 1991; Bijl en Bauduin, 1991; Houtman et al. 1991; De Heus en Diekstra, 1992; Winnubst en Schabracq, 1992; Schabracq en Winnubst, 1993; Diekstra et al., 1994; Houtman et al., 1995). Psychosociale arbeidsbelasting en werkstress zijn zowel in de wetenschap als in de maatschappij aanvaard als belangrijke thema's met betrekking tot arbeidsongeschiktheid, ziekteverzuim, en productiviteit. (Kalimo et al., 1987; Karasek en Theorell, 1990; Kompier en Pot, 1993). De thematiek van psychosociale arbeidsbelasting en werkstress is daardoor van belang voor arbodiensten, voorheen bedrijfsgezondheidsdiensten, die zich immers specifiek bezighouden met advisering over arbeidsomstandigheden in relatie tot verzuim en arbeidsongeschiktheid (Schabracq, 1992; Van der Klink, 1993; Van den Berg et al., 1993).

Het onderwerp psychosociale arbeidsbelasting en werkstress is enorm breed. Zowel de internationale als de nationale literatuur op dit gebied is omvangrijk. In het toepassingsgebied van waaruit dit proefschrift werd geïnitieerd, de arbodiensten c.q. bedrijfsgezondheidsdiensten, is het tot nu toe echter relatief onderbelicht. De ontwikkeling binnen het werkveld psychosociale arbeidsbelasting en werkstress is achtergebleven, vergeleken bij gebieden als veiligheid en fysisch-chemische arbeidsomstandigheden. Pas de afgelopen vijf jaar beginnen systematische activiteiten van de grond te komen, via kanalen als: welzijn in de Arbowet, psychologie van arbeid en gezondheid, en arbeids- en organisatie deskundigheid in de zin van de Arbowet. Het is een complicatie voor dit proefschrift om enerzijds te willen aansluiten bij de omvang en complexiteit van de vakliteratuur, anderzijds bij de beperkte staat van ontwikkeling van dit toepassingsgebied van arbo- c.q. bedrijfsgezondheidsdiensten.

1.2 Vraagstellingen

In dit proefschrift ligt het accent op een aantal methodologische en theoretische vragen die verband houden met het onderzoek naar psychosociale arbeidsbelasting en werkstress met vragenlijsten. In essentie wil dit proefschrift de waarde van vragenlijstonderzoek op het gebied van psychosociale arbeidsbelasting en werkstress aantonen. Meer in detail worden hiertoe de volgende vragen aan de orde gesteld:

Welke zijn de te onderscheiden dimensies bij onderzoek naar psychosociale arbeidsbelasting en werkstress met vragenlijsten?

In hoeverre wordt het onderzoek naar relaties tussen psychosociale arbeidsbelasting en werkstress met vragenlijsten bemoeilijkt door verstorende variabelen?

In hoeverre spelen sociale vergelijkingsprocessen een rol bij onderzoek naar psychosociale arbeidsbelasting en werkstress?

Als we proberen om individuele symptomen van werkstress te verklaren op basis van psychosociale arbeidsbelasting, hoeveel verklaring ligt dan op het niveau van de individuele werknemer, hoeveel op het niveau van de afdeling of de functie, en hoeveel op het niveau van de organisatie? Hoe zit de relatie tussen psychosociale arbeidsbelasting en werkstress in elkaar? Hoe is de relatie van psychosociale arbeidsbelasting en werkstress met ziekteverzuim?

Bij het beantwoorden van deze methodologische en theoretische vragen probeert dit proefschrift een brug te slaan tussen theorie en praktijk. Technische beschouwingen over schaalqualiteiten, verstorende variabelen, validiteit zijn voor de pragmaticus van belang om vertrouwen te kunnen hebben in de keuze voor een vragenlijstinstrument. Praktische beschouwingen over lengte, begrijpelijkheid, en nut van een vragenlijst sturen de theoreticus bij het uitbouwen van zijn instrumentarium en kennis.

Alhoewel dit proefschrift is ontstaan vanuit de optiek van een psycholoog werkzaam bij een arbodienst, kunnen wellicht ook aanpalende werkvelden hun voordeel ermee doen. Dit proefschrift is van belang voor al diegenen die bij arbodiensten, bedrijven, branche-organisaties, vakbonden, onderzoeksinstituten, universiteiten, overheid en adviesbureaus betrokken zijn bij onderzoek en beleid op het gebied van psychosociale arbeidsbelasting en werkstress.

1.3 Begrippen en definities

Er zijn twee problemen aan termen als werkstress, welzijn, werkdruk, en welbevinden. In de eerste plaats worden ze in de vakliteratuur soms breed, soms smal gebruikt. Onder werkdruk wordt in het ene geval alleen het tempo en de hoeveelheid van het werk verstaan, in het andere geval wordt de term gebruikt als koepelbegrip voor de totale arbeidsbelasting die iemand heeft. Onder werkstress wordt soms het hele terrein van oorzaken, gevolgen en intermediërende factoren verstaan waarlangs de werkomgeving de werknemer belast. Soms ook wordt werkstress gebruikt als term die de toestand beschrijft waarin iemand verkeert die gespannen is. De termen welzijn en welbevinden zijn nog minder helder, en worden afhankelijk van de context (economie, politiek, maatschappelijk werk, psychologie) anders gebruikt, en soms met elkaar uitgewisseld.

Het tweede probleem is dat in de academische setting een woord een andere betekenis kan hebben dan in de volksmond. Dit is het meest duidelijk bij de term

stress. Waar de meest gangbare academische opvatting van stress is, dat dit slaat op een verstoord evenwicht tussen belastende factoren en het vermogen om deze op te vangen, heeft de term stress in de populaire zin vooral de betekenis van een "gespannen toestand".

Aangezien dit proefschrift zich begeeft op het raakvlak van theorie en praktijk, en aangezien bij het maken van een vragenlijst wordt geprobeerd om een psychometrisch helder begrippenkader te ontwikkelen, is de kwestie van heldere begrippen voor dit proefschrift extra belangrijk.

Een aantal termen zal in dit proefschrift zo min mogelijk worden gebruikt, behalve daar waar literatuur wordt besproken waarin deze termen gangbaar zijn. Het gaat dan met name om de termen: arbeidssatisfactie, welzijn, en werkdruk. De term arbeidssatisfactie is enigszins uit de mode aan het raken. De strategie om diverse aspecten van arbeid, arbeidssituatie, en toestand van de werknemer uit te drukken in termen van tevredenheid werkt een vertekening van de afzonderlijke antwoorden in de hand door een gezamenlijke psychologische dispositie of antwoordtendens (Zapf, 1989). Karasek en Theorell (1990) noemen het uitsluitend werken met een dergelijke tevredenheidsbenadering zelfs contra-productief.

De term welzijn is door de specifieke invulling die hieraan is gegeven in de arbowet zeer beladen. Jammer is vooral, dat deze wettelijke benadering (ook wel de conditionele benadering genoemd) niet makkelijk aansluit bij de invulling die in het dagelijks taalgebruik aan de term welzijn wordt gegeven. We gebruiken de term welzijn dan ook zo min mogelijk, om zowel de wettelijke beladenheid als de verwarring met de omgangstaal te vermijden.

Werkdruk zullen we als term niet veel gebruiken in verband met de eerder genoemde onduidelijkheid over de breedte van deze term. Het is jammer dat werkdruk steeds meer als koepelbegrip wordt gebruikt. Hierdoor wordt het steeds moeilijker om werkdruk te onderscheiden van termen als arbeidsbelasting en werkstress. Een smalle betekenis in de zin van "hoeveel werk in hoeveel tijd" is het meest helder en zinvol.

Dan volgt nu een nadere definitie van de begrippen die we wel veelvuldig zullen hanteren in dit proefschrift. Dit zijn: psychosociale arbeidsbelasting, werkstress, spanning, en welbevinden.

In de eerste plaats psychosociale arbeidsbelasting. Hiermee wordt bedoeld: de verzameling van aspecten van de arbeidsinhoud en de arbeidssituatie (arbeidsverhoudingen, arbeidsomstandigheden en arbeidsvoorwaarden), die op het cognitieve, maar vooral ook op het motivationele en emotionele vlak invloed uitoefent op werknemers.

De term psychosociale arbeidsbelasting is handig omdat deze als koepelbegrip op één lijn kan worden geplaatst met andere koepelbegrippen voor typen arbeidsbelasting, zoals fysieke arbeidsbelasting en fysisch-chemische arbeidsbelasting.

Hierdoor krijgt psychosociale arbeidsbelasting qua naamgeving een logische plek naast andere mogelijke oorzaken van verzuim en arbeidsongeschiktheid waar arbodiensten hun advisering op richten.

De afgrenzing van psychosociale arbeidsbelasting ten opzichte van andere typen arbeidsbelasting is in zoverre problematisch, dat er best andere typen arbeidsbelasting voorstelbaar zijn, die een effect hebben op het cognitieve, motivationele en emotionele vlak. Denk bijvoorbeeld aan het effect van sommige toxische stoffen in de arbeidssituatie op het centrale zenuwstelsel. Een toevoeging aan de definitie van psychosociale arbeidsbelasting is dan ook, dat de invloed op het cognitieve, motivationele en emotionele vlak mede (of misschien zelfs voornamelijk) tot stand komt via de beleving van en beoordeling door de werknemer.

De term psychosociale arbeidsbelasting dekt af wat in veel andere teksten wordt aangeduid met werkstressoren. In het vervolg wordt de term psychosociale arbeidsbelasting afgekort tot "PSA".

Werkstress kan als term niet worden gemist vanwege de mate waarin deze term is ingeburgerd in wetenschap en maatschappij. In tegenstelling tot veel andere wetenschappelijke verhandelingen zullen we de term werkstress in dit proefschrift meer in de populaire zin gebruiken van een toestand waarin de werknemer verkeert.

We zullen zien dat deze toestand van werkstress, voor zover die kan worden gemeten met een vragenlijstmethode, operationeel kan worden beschreven in termen van arbeidsgerelateerde spanningsverschijnselen en arbeidsgerelateerd welbevinden. Dit zijn de andere twee termen die we veel zullen gebruiken.

Bij arbeidsgerelateerde spanningsverschijnselen (deze term zal meestal worden ingekort tot spanning) gaat de interesse in dit proefschrift niet zozeer uit naar de meer fysiologische verschijnselen van een werknemer, die zijn gekoppeld aan de in het werk geleverde inspanning, als wel naar de mentale interpretatie die een werknemer maakt van deze verschijnselen.

Arbeidsgerelateerd welbevinden (deze term zal meestal worden ingekort tot welbevinden) wordt hier bedoeld in de zin van psychologisch welbevinden bij de arbeid. Dit is te omschrijven als de mate waarin een werknemer een positieve dan wel negatieve opstelling heeft tegenover het werk en de organisatie.

Samengevat is dit het elementaire begrippenkader voor het proefschrift: psychosociale arbeidsbelasting is een determinant van werkstress, naast andere vormen van arbeidsbelasting. Werkstress kan worden beschreven in termen van spanning en welbevinden.

Deze paragraaf heeft niet de pretentie om een algemeen geldig begrippenkader aan te reiken. De doelstelling van deze paragraaf is slechts om in dit proefschrift helder te communiceren met de lezer. De genoemde definities zijn te beschouwen als werkdefinities.

1.4 Overzicht van de inhoud

Het proefschrift is opgebouwd uit drie delen. In het eerste deel gaan we in op de theoretische achtergronden van psychosociale arbeidsbelasting (hoofdstuk 2), werkstress (hoofdstuk 3), en het meten hiervan met vragenlijsten (hoofdstuk 4).

Daarna zullen we ons in het tweede deel gaan bezighouden met het afleiden van kenmerken van PSA en werkstress die op een betrouwbare en unidimensionele manier kunnen worden beoordeeld door werknemers zelf, aan de hand van een vragenlijst. We gaan daarna de beoordelingen van een groep werknemers over hun werk als handvat gebruiken om te komen tot een verzameling van vragenlijstschalen die aan een aantal technische criteria voldoet, en die het terrein PSA en werkstress inhoudelijk goed afdekt. Verder kijken we op welke manier deze verzameling van schalen het beste kan worden geordend.

In meer technische termen: deel II van dit proefschrift beschrijft de ontwikkeling van een vragenlijst voor het meten van PSA en werkstress. Eerst beschrijven we het vragenlijstontwerp (hoofdstuk 5), vervolgens het schaalconstructie-onderzoek (hoofdstuk 6).

In het derde deel gaan we in op de centrale methodologische en theoretische vragen van dit proefschrift (zie 1.2). Het beantwoorden van deze vragen geeft tevens indicaties over de validiteit van de in het tweede gedeelte geconstrueerde vragenlijst. In hoofdstuk 7 komen de convergente en divergente validiteit van de vragenlijstschalen aan bod. Dit gebeurt aan de hand van een dimensionaliseringsstudie. De geconstrueerde verzameling van schalen wordt in dit hoofdstuk samengevat tot een beperkt aantal dimensies. In hoofdstuk 8 komt de concurrente validiteit aan bod. In dit deel gaan we in op de vragen over verstorende variabelen, de rol van sociale vergelijking, het niveau van verklaring (individu, groep, organisatie), en de aard van de relatie tussen PSA en werkstress. In hoofdstuk 9 gaan we in op de predictieve validiteit van PSA en werkstress met betrekking tot ziekteverzuim. Het proefschrift sluit af met een discussie van de resultaten uit het derde deel. Gepoogd wordt om deze resultaten te integreren en in verband te brengen met de literatuur uit het eerste deel. Tevens wordt bekeken wat de beperkingen zijn van het onderzoek en wat de betekenis is van het onderzoek voor de praktijk van dienstverlening op het gebied van PSA en werkstress (hoofdstuk 10).

Voor een goed begrip van dit proefschrift is het van belang aan te geven in welke historische volgorde de tekst is ontstaan. De volgorde was dat eerst deel II werd geschreven (hoofdstuk 5 en 6), vervolgens deel III (hoofdstuk 7, 8, en 9) gelijktijdig met deel I (hoofdstuk 2, 3 en 4), en daarna tenslotte de inleidende en afsluitende teksten (hoofdstuk 1 en 10). De ontstaansvolgorde is herkenbaar in de tekst doordat de in hoofdstuk 1 t/m 4 gehanteerde terminologie eigenlijk pas in hoofdstuk 5 en 6 wordt opgebouwd.

Aan het begin van deel II kan de lezer dan ook qua gebruikte terminologie de oriëntatie kwijtraken. De lezer is hierbij gewaarschuwd, maar willen we tegelijkertijd geruststellen: aan het eind van deel II komt het allemaal weer goed.

De tekst in het tweede deel is gedeeltelijk eerder gepubliceerd (Van Veldhoven en Meijman, 1994). Aanzetten voor de inhoud van het eerste deel zijn in deze publicatie eveneens terug te vinden.

Uit het feit dat dit proefschrift in de hierboven beschreven volgorde is ontstaan, mag niet worden afgeleid dat bij het werken aan het tweede en derde deel niet is uitgegaan van wetenschappelijke theorievorming op het gebied van PSA en werkstress. Het is eerder zo dat de keuze is gemaakt om niet te werken vanuit één specifieke theoretische benadering, en in dit proefschrift de aandacht te concentreren op een aantal inhoudelijke en methodologische vraagstellingen die van belang zijn bij vragenlijstonderzoek naar PSA en werkstress, ongeacht het theoretisch vertrekpunt. De gevonden resultaten zijn vertaalbaar naar de meeste van de gangbare theorieën, zoals beschreven in het eerste deel.

Deel I: Theorie

Hoofdstuk 2: Psychosociale arbeidsbelasting

Een proefschrift dat zich wil bezighouden met het meten van PSA, zal zich moeten verdiepen in de theoretische noties over dit onderwerp. In dit hoofdstuk proberen we een aantal van de meest bekende manieren om PSA te benaderen systematisch te bespreken.

Door Kompier en Pot (1993) is een overzicht gegeven van een vijftal benaderingen op dit gebied, en de consequenties van deze theoretische benaderingen voor verschillen in methoden van onderzoek en diagnostiek. Deze vijf benaderingen komen hier terug, en worden aangevuld met een aantal andere. Achtereenvolgens komen aan bod: de sociotechniek (2.1), de handelingstheorie (2.2), de taakkenmerken-benadering (2.3), het Michigan-model (2.4), het demand-control model (2.5), het vitamine-model (2.6), de belasting-belastbaarheid benadering (2.7), en tenslotte het integraal model van Beehr en Newman (2.8).

We gaan in dit hoofdstuk niet diep op alle afzonderlijke theorieën in. We bespreken ze slechts om er een vertrekpunt uit te halen voor de vragenlijstconstructie in deel II (2.9). Hiertoe proberen we telkens per benadering de volgende vragen te beantwoorden:

Wat zou er volgens deze theorie gemeten moeten worden, en hoe? Op welke wijze kan werkstress voorkomen c.q. bestreden worden volgens deze theorie?

Aan het eind van dit hoofdstuk formuleren we kort een aantal fundamentele methodologische verschillen tussen de beschreven benaderingen (2.10).

2.1 Sociotechniek

Deze stroming bouwt verder op het werk van het Tavistock Institute of Human Relations in London (Trist en Bamforth, 1951). Hier was men vooral geïnteresseerd in de aansluiting tussen de technische organisatie en de sociale organisatie, met name op het niveau van kleinere werkeenheden en groepen. De eerste experimenten met het model deed men in de zware industrie (kolenmijnen).

In haar proefschrift beschrijft Haak (1994) hoe, geïnspireerd door het Tavistock Institute, sociotechnische tradities ontstonden in Groot-Brittannië, Noorwegen, Zweden en Nederland. De Nederlandse benadering kan worden getypeerd als een ontwerpbenadering. Haak onderscheidt in de Nederlandse sociotechniek vier ontwikkelingsfasen.

In de eerste fase vinden projecten en experimenten plaats, gebaseerd op het Britse model van gezamenlijk ontwerp van sociale en technische systemen. De tweede fase werd geïnitieerd vanuit het bedrijfsleven, en betrof de vergelijking van verschillende projecten op het gebied van werkstructurering. In de jaren tachtig lag de derde fase, waarin de Nederlandse sociotechniek de huidige vorm kreeg (De Sitter, 1981; Kuipers en Van Amelsvoort, 1990). Centraal in deze derde fase was de ontwikkeling van een sociotechnisch organisatiemodel en een ontwerpketen op basis van een analyse van het primaire productieproces. Uitgangspunt is het beter afstemmen van regelvermogen en regelbehoeften in het organisatie-ontwerp. In de huidige, vierde fase wordt de rol van de groepsdynamica en het ontwerp van de besturingsstructuur verder uitgewerkt.

De basisgedachte van de sociotechniek is om arbeidstaken en -situaties zo te ontwerpen dat ze geen werkstress oproepen. Werkstress is een gevolg van risico's in werk en werksituatie. De Sitter (1990) formuleert het als volgt (1990): *"Zo bezien is het kernprobleem van een werker een dynamisch probleem: aan welke kenmerken moet mijn werk voldoen, wil ik in staat zijn met mijn in de tijd veranderlijke werkproblemen te kunnen omgaan?"*. De werkstressrisico's die de arbeidssituatie bevat moeten worden vastgesteld en aangepakt. Dergelijke risico's moeten objectief en onafhankelijk van de werknemer worden gemeten (Christis, 1991).

2.2 Handelingstheorie

Deze theorie is afkomstig uit Duitsland. Voor Nederlandstalige inleidingen in de handelingstheoretische benadering en de daarbij behorende instrumenten, zie Roe en Zijlstra (1991) en Van Ouwerkerk, Meijman en Mulder (1994).

Arbeid wordt in de handelingstheorie opgevat als doelgericht handelen. Het doel komt hierbij tot stand op basis van eisen uit de arbeidssituatie en interpretaties door individuele werknemers van deze eisen.

In de werken van Hacker (1973; 1986) en Volpert (1983) staat het begrip "volledigheid van de arbeidstaak" centraal. Deze volledigheid heeft in de eerste plaats betrekking op de aard van de handelingsregulering die de arbeidstaak vereist. Als de werknemer, om zijn arbeidstaak doelmatig te realiseren, een beroep moet doen op zowel intellectuele, als perceptueel-conceptuele, als sensomotorische reguleringsmechanismen, wordt een arbeidstaak "hiërarchisch volledig" genoemd. De volledigheid heeft in de tweede plaats betrekking op de handelingsregulering in de tijd die nodig is om de taak doelmatig te realiseren. Als de taak voorbereidende, uitvoerende, organiserende, en controlerende taken bevat wordt de taak "sequentieel volledig" genoemd. Een dergelijke sequentiële volledigheid impliceert meestal ook een bepaalde mate van regelmogelijkheden in het werk.

De stellingname van de handelingstheorie is dat werknemers met volledige taken minder werkstress zullen hebben en meer zullen opleveren voor de organisatie en voor zichzelf. Het gaat erom functies zo te ontwerpen dat werknemers een zo volledig mogelijke functie hebben, passend bij de reguleringsmechanismen die ze kunnen en willen hanteren, en waarin zij zich kunnen handhaven en ontwikkelen.

Volgens de handelingstheorie is een uitgebreide meting nodig van diverse aspecten op het gebied van de arbeidssituatie, de arbeidstaak, de persoonlijke eigenschappen van de werknemer, en de gevolgen van het werk voor de werknemer, om aan te kunnen geven hoe een arbeidstaak zodanig kan worden ontworpen dat hij de gezondheid en ontplooiing van de werknemer bevordert. Methoden uit deze benadering zijn meestal zeer tijdrovend en omvangrijk (Roe en Zijlstra, 1991; Van Ouwkerk, Meijman en Mulder, 1994).

2.3 Taakkenmerken-benadering

De taakkenmerken-benadering komt uit een Amerikaanse traditie, met als belangrijke namen Turner en Lawrence (1965) en Hackman en Oldham (1975, 1980). Deze stroming is in ons land uitgewerkt door Algera (1981; 1983; 1989).

Het Job Characteristics Model (JCM) van Hackman en Oldham (1974; 1975) is een zeer invloedrijk model (geweest) binnen de arbeids- en organisatiepsychologie. Een functie wordt in dit model beschreven aan de hand van een vijftal objectieve taakkenmerken. Dit zijn: variatie, identiteit, zinvolheid, autonomie, en feedback. Deze vijf kenmerken beïnvloeden drie subjectieve ervaringen (Hackman en Oldham noemen dit "kritieke psychologische toestanden") bij de werknemer. In de eerste plaats de ervaren zinvolheid van het werk, in de tweede plaats de ervaren verantwoordelijkheid voor het resultaat uit het werk, en tenslotte de kennis van het behaalde resultaat uit het werk. Deze drie kritieke psychologische toestanden hebben weer gevolgen voor de motivatie, satisfactie, productiviteit, het ziekteverzuim, en het verloop.

Om het model te completeren, wordt de behoefte aan groei en ontwikkeling als variabele erin opgenomen. De behoefte aan groei en ontwikkeling zou de veronderstelde relaties tussen taakkenmerken, kritieke psychologische toestanden en gevolgen modereren.

De kritiek op het JCM concentreert zich op een aantal punten. In de eerste plaats is er de kwestie of taakkenmerken objectief dan wel middels zelfbeoordeling kunnen worden gemeten (Salancik en Pfeffer, 1978).

In de tweede plaats is er kritiek op de hoeveelheid en keuze van de fundamentele taakkenmerken (Fried en Ferris, 1986; 1987). Tenslotte is commentaar geleverd op de onderbouwing van de oorzakelijkheidsketen, zoals die in het model wordt aangegeven van taakkenmerken, via kritieke psychologische toestanden, naar gevolgen (Wall et al., 1978; Roberts en Glick, 1981; Fried en Ferris, 1987). Deze punten van kritiek op het JCM zijn nog steeds actueel. Ze komen in dit proefschrift uitgebreid aan bod. Zo gaan we in dit proefschrift in op de vraag van de te meten dimensies in PSA en werkstress. Ook de vraag naar de betrouwbaarheid en validiteit van zelfbeoordelingsmethoden komt aan bod.

Werkstress kan volgens het JCM voorkomen en bestreden worden door te zorgen dat de drie kritieke psychologische toestanden bij een werknemer aanwezig zijn. Dit kan alleen indien de vijf taakkenmerken in de functie aanwezig zijn. De behoefte aan groei en ontwikkeling van de individuele werknemer is hierbij een modererende variabele waarmee rekening gehouden dient te worden. Alhoewel het JCM strikt genomen moet worden geoperationaliseerd met zowel objectieve als subjectieve maten, hebben de meeste studies die uitgaan van dit model uitsluitend gewerkt met vragenlijsten. In de traditie van het JCM is veel vragenlijstonderzoek gedaan, met name op het gebied van arbeidssatisfactie en -motivatie.

2.4 Michigan-model

Het Michigan-model is afkomstig van een groep onderzoekers werkend bij het Institute of Social Research aan de universiteit van Michigan (VS). Het model geeft een centrale plek aan aspecten die zijn afgeleid uit de psychologische roltheorie (Katz en Kahn, 1978). Uitgangspunt voor het Michigan-model is de person-environment fit (P-E fit), oftewel de mate waarin werknemer en werkomgeving bij elkaar passen (French, Caplan en Van Harrison, 1982). Dit concept sluit aan bij het begrip "correspondence", zoals dat wordt gebruikt in de psychologische theorie over aanpassing aan het werk van Dawis en Lofquist (1984).

Het ontstaansproces van werkstress wordt beschreven als een keten die leidt van objectieve oorzaken (stressoren) naar objectieve gevolgen (strains) via subjectieve beleving (eerst van de stressoren daarna van de strains). Het verloop van het proces in deze keten wordt gemodereerd door twee factoren: de persoonlijkheid en de sociale steun van collega's en leiding.

Het Michigan-model is zeer invloedrijk geweest. Ook in Nederland is er een onderzoekstraditie in het kader van dit model, met als centrale plaatsen de vakgroepen arbeids- en organisatiepsychologie van de universiteiten in Nijmegen (Van Dijkhuizen, 1980; Reiche, 1982; Marcelissen, 1987; Buunk en De Wolff, 1988) en Utrecht (Winnubst, 1986; De Jong, 1987).

De kritiek op het Michigan-model heeft enerzijds betrekking op het model zelf (keuze van factoren, hoeveelheid factoren, gebrek aan dynamiek), anderzijds op het feit dat in het merendeel van het onderzoek nauwelijks aandacht wordt besteed aan objectieve, oorzakelijke factoren in de werkomgeving en aan objectieve stress-gevolgen (Kasl 1978, 1986; Christis, 1988; Buunk en De Wolff, 1988).

Om werkstress te voorkomen moet volgens het Michigan-model vooral de P-E fit worden bewaakt. Een analyse van de objectieve werksituatie alleen, zegt dan ook niet voldoende. Persoonskenmerken, de subjectieve beleving van het werk, de sociale relaties op het werk, en de ervaren psychologische en lichamelijke symptomen van spanning, zijn evenzeer van belang. In deze traditie wordt dan ook veelvuldig gebruik gemaakt van vragenlijsten (Caplan et al., 1975; Bergers, Marcelissen en De Wolff, 1986).

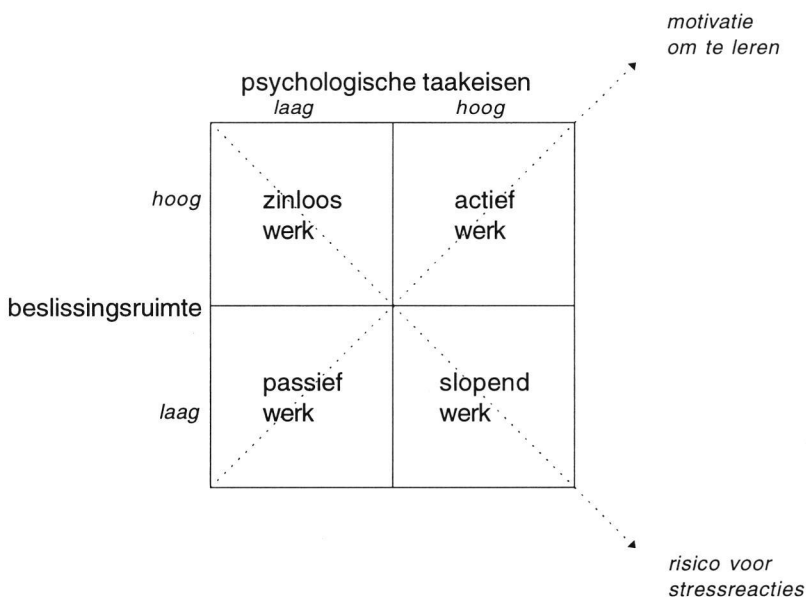
2.5 Demand-control model

Het demand-control model van Karasek (1979) vindt zijn oorsprong in Scandinavië en de VS. Het combineert elementen van de sociotechniek en de taakkenmerkenbenadering. In zijn oorspronkelijke vorm is het demand-control model eenvoudig. Karasek stelt dat er slechts twee fundamentele aspecten van invloed zijn op werkstress. In de eerste plaats is dat het aspect van de psychologische taakeisen ("psychological job demands"). Karasek (1979) omschrijft dit als "psychological stressors involved in accomplishing the workload" (p. 291). Het tweede belangrijke aspect is dat van de beslissingsruimte ("decision latitude", ook wel: "job control"). Karasek (1979) omschrijft dit als "the working individual's potential control over his tasks and his conduct during the working day" (p. 289-290).

Bij nadere beschouwing is het concept van beslissingsruimte niet eenduidig uitgewerkt door Karasek. Hij splitst dit concept op in twee theoretisch onderscheidbare deelconcepten, namelijk: de variatie in het beroep dat het werk doet op de vaardigheden van een werknemer ("skill discretion" oftewel "task variety"), en de mogelijkheden die een werknemer heeft om beslissingen te nemen over de uitvoering van de arbeidstaak ("task authority" oftewel "autonomy"). Als argumentatie voor de combinatie van deze twee deelconcepten tot het begrip "beslissingsruimte" voert Karasek het volgende aan: *"We have claimed that these (skill discretion and task authority) are mutually reinforcing aspects of work that often appear together in the workplace, in a combination which we call decision latitude. Because in our opinion a high level of skill gives the worker control over which specific skills to use to accomplish the task, we refer to such a measure as skill discretion, and often loosely label the decision latitude combination: 'control'. Skill utilization and decision authority are so closely related in empirical studies that they are often combined for analytic purposes in job design research"* (Karasek en Theorell, 1990; p. 58).

Op basis van de dimensies psychologische taakeisen en beslissingsruimte wordt een model gebouwd waarin vier cellen worden onderscheiden (zie figuur 2.1). Werknemers met hoge psychologische taakeisen en veel beslissingsruimte hebben banen met uitdaging, leermogelijkheden, en perspectief (in termen van Karasek: actieve banen). Tegenover dit soort banen liggen de passieve banen. Hierin hebben de werknemers zowel weinig psychologische taakeisen als beslissingsruimte. Dit soort banen is saai en biedt weinig mogelijkheden voor groei.

Vanuit het oogpunt van werkstress het meest belangrijk, zijn banen met veel psychologische taakeisen, maar weinig beslissingsruimte (slopend werk). Het model voorspelt juist bij werknemers in dit soort werk veel symptomen van werkstress. Tegenover deze banen staan banen met een laag risico op werkstress. Dit is het geval bij een combinatie van weinig psychologische taakeisen en veel beslissingsruimte (zinloos werk).



Figuur 2.1: Het demand-control model van Karasek (1979)

Het Karasek-model spreekt aan door zijn eenvoud, en is waarschijnlijk op dit moment het meest gebruikte onderzoeksmodel in het vakgebied. Alhoewel Karasek en Theorell (1990) in hun boek een enorme massa aan onderzoeksgegevens presenteren waarin het model wordt bevestigd, is in de literatuur de laatste jaren de kritiek fors toegenomen.

In veel onderzoek is de meest centrale stelling van het model, namelijk het interactie-effect van psychologische taakeisen met beslissingsruimte, niet bevestigd (Spector, 1987; Parkes, 1990; Warr, 1990a; Fletcher en Jones, 1993).

Karasek en Theorell (1990) hebben een uitbreiding voorgesteld van het oorspronkelijke model door de dimensie sociale ondersteuning eraan toe te voegen. Opnieuw wordt verondersteld dat er een interactie-effect zal optreden van weinig sociale steun met hoge psychologische taakeisen en weinig beslissingsruimte. Volgens dit "expanded demand-control model" is het mogelijk werkstress te voorkomen, door de combinatie van hoge psychologische taakeisen met weinig beslissingsruimte en weinig sociale ondersteuning te vermijden. Voor dit uitgebreide model is empirische ondersteuning gevonden in grootschalig onderzoek in Zweden (Johnson en Hall, 1988).

Ondanks de risico-benadering, wordt in de Karasek-traditie vooral gewerkt met vragenlijstonderzoek voor het vaststellen van PSA-kenmerken. De werkbeleving van individuele werknemers wordt vervolgens geaggregeerd naar het niveau van beroepen. Deze geaggregeerde gegevens worden gekoppeld aan andere gegevensbronnen op collectief niveau, zoals medische en ziektekosten-registraties. Het komt hierdoor dat de onderbouwing van het Karasek-model een epidemiologische inslag heeft.

2.6 Vitamine-model

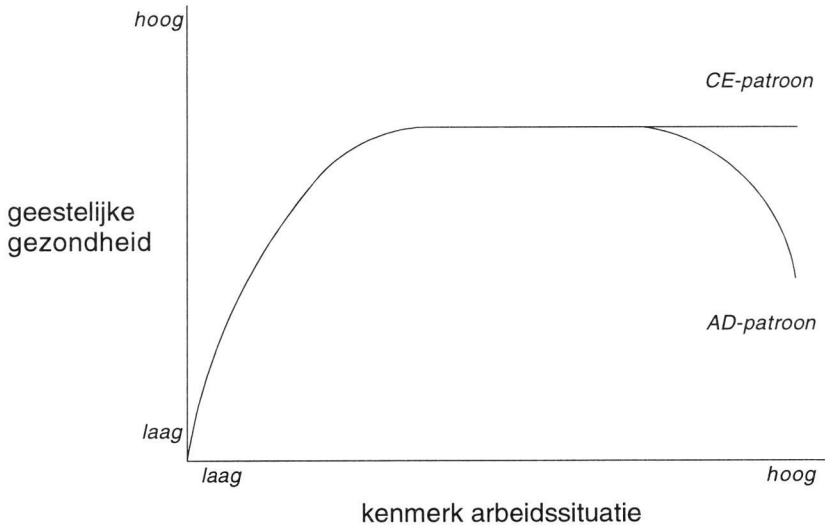
Door Warr (1987; 1994) is een model voorgesteld dat het effect van arbeid op de geestelijke gezondheid vergelijkt met het effect van vitamines op de lichamelijke gezondheid.

Warr besteedt ruime aandacht aan de operationalisatie van het begrip geestelijke gezondheid in relatie tot de arbeid. Een deel van de door hem ontwikkelde vragenlijstschalen wordt in deel II van dit proefschrift gebruikt bij de vragenlijst-constructie. In het kader van dit hoofdstuk gaat de interesse echter vooral uit naar de invulling die Warr geeft aan PSA.

Vergelijkbaar met vitamines hebben mensen in hun arbeid een dagelijkse dosis nodig van bepaalde arbeidskenmerken. Hiervan onderscheidt Warr er in totaal negen. Een te lage dosis van één of meer van deze negen factoren heeft negatieve gevolgen voor de werknemers.

De negen factoren zijn echter op te splitsen in factoren die bij een overdosis geen negatieve gevolgen hebben (vergelijkbaar met het effect van de vitamines C en E), en factoren die bij een overdosis wel negatieve gevolgen hebben (vergelijkbaar met het effect van de vitamines A en D). De AD- en CE-patronen van werkkenmerken in relatie tot de geestelijke gezondheid zijn weergegeven in figuur 2.2.

CE-factoren zijn volgens Warr: beloning, veiligheid, en sociale waardering. Hiervan kan men nooit een overdosis krijgen. Wel zwakt bij een hoge dosering het effect af, of neemt het effect niet meer verder toe. Dit is vergelijkbaar met effect van vitamine C, dat bij een overdosis door het lichaam weer wordt uitgescheiden.



Figuur 2.2: Het vitamine-model van Warr (1987)

AD-factoren zijn prestatie-eisen, afwisseling, duidelijkheid, regelmogelijkheden, gebruik van vaardigheden, en interpersoonlijk contact. De diverse AD-factoren hebben hun eigen 'toxische' reacties, vergelijkbaar weer met de uiteenlopende reacties op een overdosis vitamine A of D.

Te hoge prestatie-eisen, een te groot beroep op vaardigheden, en teveel regelmogelijkheden leiden volgens Warr tot overbelasting en stress. Afwisseling daarentegen heeft bij een overdosis andere effecten: dit leidt tot lage concentratie en prestatie. Teveel van deze factor leidt met andere woorden tot lage arbeidssatisfactie en arbeidsmotivatie. Teveel duidelijkheid in de functie leidt volgens Warr tot weinig gevoel van controle over het werk, en biedt weinig ruimte voor ontwikkeling. Tenslotte bespreekt Warr de factor interpersoonlijk contact. Bij een overdosis hiervan verliest men de persoonlijke controle en treden overbevolkingsreacties op.

Het vitamine-model is interessant, omdat het in tegenstelling tot andere modellen ervan uitgaat, dat diverse aspecten van werk en werksituatie tot andersoortige reacties zullen leiden, in plaats van dat ze allemaal in dezelfde richting werken. Volgens Warr is er sprake van differentiatie in de relaties tussen PSA en werkstress: er is niet één, algemeen geldig stress-mechanisme, maar er zijn er meer. Bovendien veronderstelt Warr dat de relatie tussen AD-factoren en gevolgen hiervan non-lineair is, in tegenstelling tot de meeste andere theorieën, die van lineaire verbanden uitgaan.

Warr (1990a) heeft zelf onderzoek gedaan dat de stellingnames van zijn model bevestigt. Behalve door Warr zelf is er nog niet zoveel onderzoek gedaan naar het vitamine-model. Een beperkte selectie van empirische evidentie wordt besproken in een artikel van Warr uit 1994. In Nederland heeft eveneens nog weinig onderzoek plaatsgevonden naar het vitamine-model. Uitzonderingen hierop vormen de onderzoeken van Furda (1995), De Jonge (1995) en De Jonge, Schaufeli en Furda (1995). In deze onderzoeken wordt op onderdelen bevestiging gevonden voor de predicties uit het vitamine-model.

In het derde deel van dit proefschrift wordt het vitamine-model gebruikt als leidraad voor het bestuderen van de aard van de verbanden tussen PSA en werkstress.

Het voorkomen en bestrijden van overbelasting en stress is voor Warr fundamenteel iets anders dan het creëren van leuk werk. Bij spanningsklachten is het vooral een kwestie van het optimaliseren van de arbeidsbelasting: niet teveel werk en niet te weinig, niet te moeilijk werk en niet te gemakkelijk, niet te veel regelmogelijkheden en niet te weinig. Bij het creëren van leuk werk is het eveneens van belang om de arbeidsbelasting te optimaliseren, maar op andere punten. Een arbeidstaak dient niet te weinig, maar ook niet teveel afwisseling te bieden. Hetzelfde geldt voor taakduidelijkheid en interpersoonlijk contact.

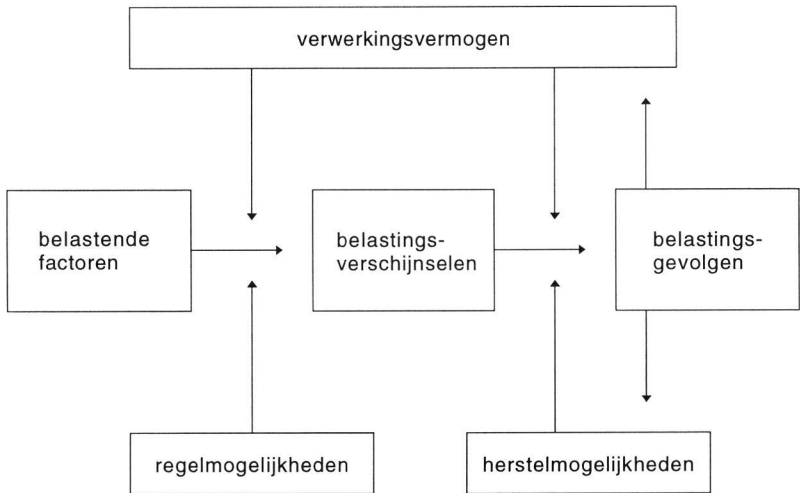
Warr ziet de arbeidstaak en arbeidssituatie niet als objectief meetbare gegevens, maar als individu-afhankelijke gegevens. Het zal derhalve niet verbazen dat hij vooral de vragenlijstmethode verkiest in zijn onderzoek (Warr, 1987; 1990a; 1990b).

2.7 Belasting-belastbaarheid benadering

In Nederland is er een onderzoekstraditie die we kunnen aanduiden als de belasting-belastbaarheid benadering. Oorspronkelijk werd deze benadering gestart in de arbeidsgeneeskunde door Burger (1959), gebaseerd op theorievorming in de arbeidsfysiologie. Het klassieke model belasting-belastbaarheid werd later helder geformuleerd door Ettema (1973).

Centraal hierin staat de uitwendige belasting door het werk (zowel taakgebonden als situatieve factoren), die leidt tot functionele belasting bij de werknemer (lichamelijke reacties). De belastbaarheid wordt gedefinieerd als de maximaal mogelijke uitwendige belasting die iemand kan verdragen, waarbij na afloop nog volledig herstel kan optreden.

Door van Dijk, van Dormolen, Kompier en Meijman (1990) is het klassieke model belasting-belastbaarheid herzien tot het zogenaamde model arbeidsbelasting, met als toepassingsgebied het terrein arbeid en gezondheid. Kompier (1991) publiceerde eenzelfde model als model voor het toepassingsgebied werkstress. Deze twee modellen baseren zich beide op eerder werk van Meijman (1989), die het belasting-herstel model ontwikkelde (zie ook: Meijman en Mulder, 1992). We verwijzen naar deze groep van modellen in dit proefschrift verder collectief onder de noemer "herziene belasting-belastbaarheid modellen". Het belasting-herstel model is ter illustratie van deze categorie modellen weergegeven in figuur 2.3.



Figuur 2.3: Het belasting-herstel model van Meijman (1989)

Waar men in het klassieke model uitging van een simpele balans tussen omgevingseisen en belastbaarheid van de werknemer, is in de herziene modellen het begrip sturing (afgeleid van het Engelse "control") toegevoegd. Het gaat hierbij om de mate waarin een werknemer invloed heeft op de arbeidstaak en op de omgeving waarbinnen de arbeidstaak wordt uitgeoefend. Om die invloed uit te

kunnen oefenen beschikt de werknemer over een bepaalde mate van regel mogelijkheden. Voor een nadere uiteenzetting over het begrip sturing zie Furda en Meijman (1992).

Een tweede herziening van het klassieke model belasting-belastbaarheid heeft betrekking op het belastbaarheidsbegrip. In het klassieke model is dit een normatief begrip (de maximale uitwendige belasting die men kan doorstaan zonder schade, of waarvan nog herstel mogelijk is). In de herziene modellen is het verwerkingsvermogen een veel dynamischer begrip. Niet zozeer wat iemand kan doorstaan (belastbaarheid), als wel wat iemand aan capaciteit wil leveren, staat daarbij centraal (Meijman en O'Hanlon, 1984; Meijman, 1989).

De derde herziening heeft betrekking op de verbinding van de belastingsverschijnselen met de belastingsgevolgen, die in het klassieke model eigenlijk volledig ontbrak. Dit aspect van het model is nader uitgewerkt door Kompier (1988) in zijn cumulatief procesmodel. Belastingsverschijnselen zijn tijdelijke reacties op de arbeidsbelasting. Bij normaal herstel zijn deze verschijnselen na verloop van tijd weer verdwenen (meestal dezelfde dag nog). Bij onvoldoende herstel treedt er een residu op. Bij herhaling van onvoldoende herstel cumuleert dit residu over de tijd. Hierdoor kunnen belastingsgevolgen ontstaan met een meer permanent karakter.

Het belasting-herstel model (Meijman, 1989), voegt aan het voorgaande nog een belangrijk element toe, namelijk: herstel mogelijkheden. Dit begrip heeft betrekking op de tijd en mogelijkheden die de werknemer heeft om van de belastingsverschijnselen ten gevolge van het werk te recupereren. Dit wordt, samen met het verwerkingsvermogen, gezien als één van de sleutelvariabelen bij het ontstaan van belastingsgevolgen uit belastingsverschijnselen.

Verband houdend met de plaatsing van de herstel mogelijkheden in het model, wordt ook het aspect regel mogelijkheden binnen het belasting-herstel model wat meer afzonderlijk geplaatst. Regel mogelijkheden worden gezien als sleutelvariabele in de relatie tussen belastende factoren en belastingsverschijnselen.

In welke versie het herziene model belasting-belastbaarheid ook wordt gepresenteerd, de centrale notie - waar het gaat om het voorkomen en bestrijden van werkstress - is het handhaven van een balans tussen belasting (al dan niet inclusief regel- en herstel mogelijkheden) en belastbaarheid (al dan niet dynamisch geformuleerd als verwerkingsvermogen). Het onderscheid in de tijd tussen belastingsverschijnselen en belastingsgevolgen is hierbij een graadmeter van de voortgang in het ontstaansproces van werkstress.

De belasting-belastbaarheid benadering heeft zowel veel onderzoek opgeleverd met objectieve metingen en registraties als met vragenlijsten. Er lijkt geen voorkeur te bestaan voor het één of het ander op theoretische of pragmatische gronden.

2.8 Integraal model Beehr en Newman

Een andere manier om een model te formuleren dan op basis van een theorie, werd toegepast door Beehr en Newman (1978). Zij voerden een kwalitatieve facet-analyse uit op de tot dan toe beschikbare literatuur over werkstress. Hierin onderscheidden zij zo'n 160 onderwerpen, die zij onderverdeelden in een zevental hoofdrubrieken, waarbinnen weer subrubrieken werden onderscheiden. Deze hoofd- en subrubrieken waren de volgende:

Rubriek 1: Omgeving

- * taakeisen en -kenmerken
- * roleisen en -verwachtingen
- * kenmerken van en condities binnen de organisatie
- * externe eisen aan en condities buiten de organisatie

Rubriek 2: Persoon

- * psychische conditie
- * fysieke conditie
- * kenmerken van de levensfase
- * demografische variabelen

Rubriek 3: Proces

- * psychologische processen
- * fysiologische processen

Rubriek 4: Gevolgen voor de werknemer

- * gevolgen voor de psychische gezondheid
- * gevolgen voor de fysieke gezondheid
- * gedragsconsequenties

Rubriek 5: Gevolgen voor de organisatie

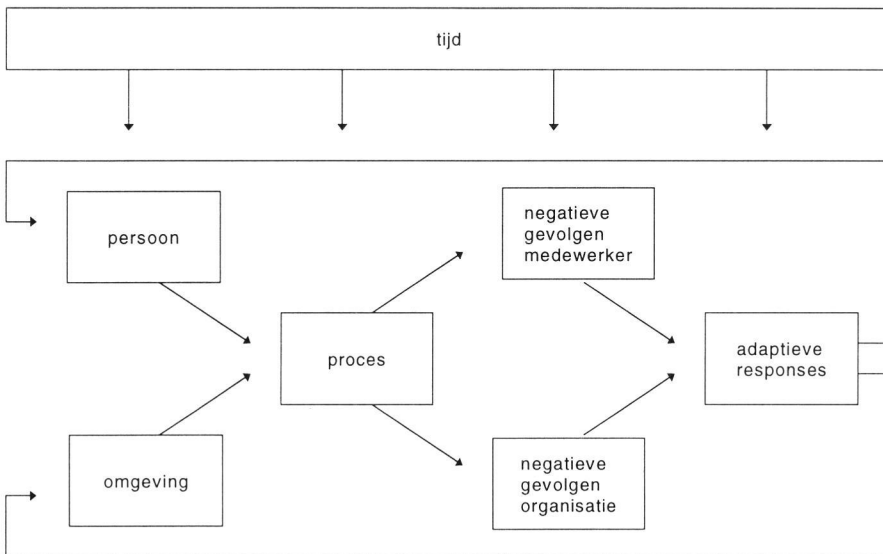
Rubriek 6: Aanpassingsreacties

- * aanpassingsreacties van de individuele werknemer
- * aanpassingsreacties van de organisatie
- * aanpassingsreacties van derden

Rubriek 7: Tijd

In woorden kunnen we de hierboven opgesomde elementen en hun onderlinge relaties (zie figuur 2.4) als volgt beschrijven. De omgeving bevat de elementen uit de situatie binnen en buiten het werk die een rol spelen bij psychosociale arbeidsbelasting. De persoonsgebonden eigenschappen bepalen de blootstelling aan en vatbaarheid voor deze factoren. Ten gevolge van de oorzaken in de omgeving

doen zich bij de werknemer voortdurend allerlei korte termijn aanpassingsprocessen voor (fysiologisch en psychologisch), gemodereerd door de persoonsgebonden eigenschappen van de werknemer. Deze aanpassingsprocessen resulteren, indien ze ineffectief blijken te zijn, in allerlei negatieve gevolgen voor de werknemer en de organisatie. Deze leiden op hun beurt weer tot adaptieve responses op langere termijn, zowel bij de werknemer, als bij de organisatie, als ook bij derden. Tenslotte onderstrepen Beehr en Newman het belang van het tijdsaspect in werkstress-onderzoek.



Figuur 2.4: Het integraal model van Beehr en Newman (1978)

Het model van Beehr en Newman kiest als uitgangspunt het proces van het ontstaan en aanpakken van werkstress. Het is daarmee vooral ook een praktijkgericht model. Toch kunnen we ook voor dit model de vragen beantwoorden over de gewenste strategie om werkstress te voorkomen en te meten.

Op de eerste vraag krijgen we ook in dit model als antwoord dat persoon en omgeving met elkaar in balans moeten zijn en blijven. Vooral wat betreft het meten laat dit model goed zien wat er allemaal te meten valt. Het model is in dit opzicht vele malen breder dan de andere gepresenteerde modellen.

Opvallend is dat Beehr en Newman ook het organisatie-niveau meenemen als te meten aspect: zowel als mede-oorzaak, als mede-slachtoffer, en als mede-acteur bij het oplossen van problemen. Wat type meting betreft wijzen Beehr en Newman in de richting van de volgende strategie: meet elk facet op de manier die daarvoor het meest geschikt is.

2.9 Vertrekpunt voor de vragenlijstconstructie

Aan het begin van hoofdstuk 2 formuleerden we een tweetal vragen, die we voor de acht verschillende benaderingen van PSA in het voorgaande hebben beantwoord. De eerste vraag had betrekking op de begrippen die een theorie voorstelt te meten, en eventuele voorkeuren over de manier waarop deze begrippen gemeten moeten worden. De tweede vraag had betrekking op de voorgestelde strategie voor het voorkomen c.q. bestrijden van werkstress.

Wanneer we eerst de antwoorden op de tweede vraag overzien, dan zijn er twee soorten van opstelling binnen de theorieën over PSA te onderscheiden. In de eerste plaats zijn er theorieën die zich uitsluitend richten op de arbeidssituatie. Dit zien we bij de sociotechniek en de Karasek-benadering. Beide zijn te karakteriseren als risico-benaderingen: elimineer de risico's en er treedt geen werkstress meer op. De overige modellen richten zich op de interactie tussen de persoon en de situatie. De strategie ter preventie van werkstress is telkens gericht op het bereiken van een balans, evenwicht of "fit" tussen de mogelijkheden van de persoon en de eisen van de situatie.

Wat betreft de eerste vraag betreft komen we tot de volgende conclusie. Alhoewel de te meten begrippen van theorie tot theorie verschillen, lijkt het wel zo te zijn dat er een grootste gemene deler is. Verder is het opvallend dat de keuze voor een al dan niet puur situationistische visie, niet samenvalt met een methodologische keuze voor of tegen vragenlijstonderzoek. Van de twee primair situatie-gerichte modellen werkt de sociotechniek vooral met (veronderstelde) geobjectiverde methoden, terwijl in de Karasek-benadering intensief gebruik wordt gemaakt van zelfbeoordelingsmethoden.

Van de andere zes modellen, die zich - althans in theoretisch uitgangspunt - richten op de interactie tussen persoon en omgeving, werken er twee hoofdzakelijk met vragenlijsten: het Michigan-model en het vitamine-model. De overige vier modellen integreren geobjectiverde methoden met zelfbeoordelingsmethoden: de taakkenmerkenbenadering, de handelingstheorie, de belasting-belastbaarheid benadering, en het model van Beehr en Newman.

In het tweede deel van dit proefschrift gaan we ons bezig houden met de constructie van een vragenlijst voor het meten van PSA en werkstress. Uit het

bovenstaande wordt duidelijk dat het werken met vragenlijsten niet theoriegebonden is. Vragenlijstinformatie kan nuttig zijn, zowel in een risico-benadering als in een interactionistische benadering. Het is dan ook niet makkelijk om voorafgaand aan de constructie van een vragenlijstmethode een fundamentele keuze voor een bepaald model te maken. Het lijkt ook niet verstandig om zo'n theoretische keuze vooraf te maken, omdat geen enkel model op dit moment algemene geldigheid en erkenning heeft (Ekkers en Sanders, 1987).

Er is behoefte om, op een andere manier dan via een a priori-keuze voor een bepaalde theorie, te komen tot een verzameling van aspecten die we in de te ontwikkelen vragenlijst willen meten. De facetanalyse methode, zoals gevolgd door Beehr en Newman, spreekt wat dit betreft aan. In paragraaf 5.2 zullen we een dergelijke facetanalyse gebruiken als vertrekpunt voor de vragenlijstconstructie.

2.10 Fundamentele methodologische verschillen tussen de theorieën

Er is een aantal methodologische kwesties, waar volgens de hier gepresenteerde theorieën over PSA zeer verschillend tegenaan wordt gekeken. De verschillen tussen de theorieën ontstaan niet zozeer alleen door verschillen in de te meten begrippen (hierin is een grootste gemene deler te ontdekken), maar vooral ook door verschillen in stellingname aangaande de volgende thema's:

Analyse-niveau

In de eerste plaats het analyse-niveau. In de meeste theorieën is dit het niveau van de individuele werknemer. Vooral de psychologische theorieën zitten op dit niveau. De sociotechniek en de handelingstheorie bekijken de zaak op het niveau van functies en functiegroepen.

Een opmerkelijke hybride vorm wat dit betreft is het model van Karasek: gegevens over de werkbeleving van individuele werknemers en gegevens over parameters van gezondheid en productiviteit bij groepen werknemers (bijvoorbeeld incidentie en prevalentie van coronaire hartziekten) worden hierin gekoppeld.

Dit verschil in analyse-eenheid is zeer fundamenteel, en heeft waarschijnlijk grote gevolgen voor de gevonden resultaten. Er bestaat slechts weinig onderzoek waarin deze materie op meerdere niveaus tegelijkertijd is onderzocht (Semmer, Zapf en Greif, 1995), mede doordat analyse-technieken hiervoor nog niet zolang beschikbaar zijn (Longford, 1993).

Aard van de verbanden

Een tweede onderwerp is de aard van de veronderstelde verbanden tussen PSA en werkstress. In het merendeel van de studies is gekeken naar de additieve effecten van diverse vormen van PSA op werkstress en andere uitkomstvariabelen.

Veelal wordt hierbij geredeneerd vanuit één algemeen stress-mechanisme: allerlei stressoren tellen bij elkaar op. Boven een bepaalde drempel ontstaan negatieve gevolgen.

Een mooi voorbeeld van een dergelijke redenering, dat in de Nederlandse praktijk ruime bekendheid geniet, is het verzuimmodel van Verbaan (1991). Hij beschrijft werknemers(-groepen) als een communicerend drukvat, waarop allerlei stressoren druk uitoefenen, en dat een bepaald vermogen heeft om deze druk te weerstaan. Wordt de druk te hoog dan kunnen allerlei soorten negatieve gevolgen ontstaan, afhankelijk van welke uitlaatklep wordt gebruikt. Alhoewel de manifestatie van de gevolgen hierdoor zeer verschillend kan zijn wordt er toch één algemeen, onderliggend mechanisme verondersteld.

Karasek (1979) doorbrak deze gangbare benadering door een interactie-effect te veronderstellen van een tweetal PSA-factoren op werkstress en andere uitkomsten, namelijk een interactie-effect van psychologische taakeisen en beslissingsruimte. In Karasek en Theorell (1990) is een grote hoeveelheid onderzoek verzameld dat het veronderstelde interactie-effect ondersteunt. In een aantal studies door andere auteurs is het interactie-effect echter niet gerepliceerd (Spector, 1987; Warr, 1990a; Parkes, 1990; Fletcher en Jones, 1993).

Een volgende mogelijkheid is dat het effect van (sommige) PSA-factoren op werkstress non-lineair is. Een dergelijke stellingname zien we bij het vitamine-model van Warr (1987). Warr (1990a) toonde aan dat kwadratische termen van psychologische taakeisen een uniek effect hebben op arbeidssatisfactie, boven op de bijdrage van de normale lineaire effecten van psychologische taakeisen. Warr veronderstelt verder dat er geen sprake is van één algemeen stress-mechanisme, maar dat er sprake is van differentiatie in een aantal, parallel werkzame mechanismen. Hij bedoelt hiermee dat specifieke PSA-factoren hun eigen specifieke werkstress-gevolgen hebben.

Oorzakelijkheidsketen

Een derde fundamentele kwestie in het onderzoek naar PSA is die van de oorzakelijkheidsrichting. In nogal wat van de besproken modellen ligt een keten-gedachte ten grondslag aan het model (JCM, Michigan-model, belasting-belastbaarheid, Beehr en Newman). Het is de vraag of deze basis-gedachte wel juist is. Met name vanuit recente theorieën over het ontstaan van psychologische stress, zoals bijvoorbeeld het symptoomperceptie-model (zie 3.3: Pennebaker, 1982; Watson en Pennebaker, 1989), is de ketengedachte in twijfel getrokken.

Spector, Dwyer en Jex (1988) vergeleken de correlaties tussen variabelen uit dezelfde gegevensbron (vragenlijsten x vragenlijsten, registratie x registratie, expert-beoordeling x expert-beoordeling) met de correlaties tussen variabelen uit meerdere gegevensbronnen (vragenlijsten x registraties x expert-beoordelingen).

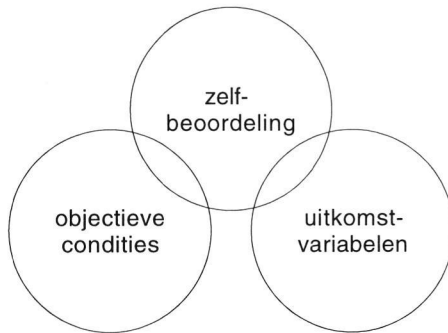
De correlaties binnen eenzelfde methode zijn systematisch hoger dan tussen methoden.

Spector (1992) betoogt dat deze situatie niet alleen ontstaat door methode-bias, maar dat er nogal wat alternatieve verklaringen voor mogelijk zijn. Al deze alternatieve verklaringen ondergraven een simpele keten-gedachte van stressor via strains naar outcomes.

James en Tetrick (1986) waren bij de eersten die met behulp van LISREL diverse mogelijke causale verbanden testten tussen taakkenmerken (volgens supervisors), taakbeleving (volgens werknemers), en arbeidssatisfactie (volgens werknemers). Zij onderzochten in hoeverre satisfactie een intermediaire variabele is tussen taakkenmerken en taakbeleving. Dit bleek niet het geval. Het model dat het beste op de resultaten paste, was een model waarin taakbeleving wordt verklaard vanuit taakkenmerken, en waarbij taakbeleving een wederkerig causaal verband heeft met arbeidssatisfactie. Het deel van de keten dat subjectieve stressoren oorzakelijk verbindt met subjectieve strains (denk aan het JCM en aan het Michigan-model) wordt hiermee dubieus.

Ook Frese (1985) onderzocht de relatie tussen stressoren en psychosomatische klachten. Hij bekeek de normale oorzakelijkheidsrichting en de omgekeerde oorzakelijkheidsrichting. Tevens onderzocht hij de invloed van een reeks van contaminerende variabelen. Hij concludeert dat: *"The results indicate that the correlation of stress and psychosomatic complaints is not spurious and cannot easily be explained by methodological artifacts; by variables such as income, job insecurity, and age; or by variables related to a stressful life of low SES groups and "political exaggeration". Finally, most data favor the interpretation that stress is a cause of psychosomatic complaints, not the reverse"* (p.326-327).

De onderzoeksresultaten over de oorzakelijkheidsketen spreken elkaar enigszins tegen. Een definitieve conclusie ontbreekt nog. Om de stand der techniek over het onderwerp "oorzakelijkheidsrichting" te beschrijven, tekent Spector (1992) een figuur van drie cirkels (zie figuur 2.5). Hierin overlapt de cirkel objectieve condities met zelfbeoordeling van condities. Ook de cirkel uitkomsten overlapt met die van zelfbeoordeling van condities. De cirkels objectieve condities en uitkomsten overlappen echter niet. Spector (1992) komt tot deze figuur op basis van de gemiddelde correlaties zoals die worden gevonden in de literatuur tussen de elementen in de cirkels. De correlaties tussen condities en uitkomsten zijn in de literatuur systematisch lager dan tussen condities en beleving enerzijds, en tussen beleving en uitkomsten anderzijds. Wat deze figuur aangeeft is, dat de aard van de verbanden tussen de drie cirkels redelijk bekend is, waarbij zelfbeoordeling een intermediaire rol heeft. De oorzakelijkheidsrichting wordt echter in het midden gelaten.



Figuur 2.5: Illustratie van een situatie waarin objectieve condities en uitkomstvariabelen beide een deel van de variatie verklaren in zelfbeoordelingen, en toch niet zijn gerelateerd aan elkaar. Ontleend aan Spector (1992)

Hoofdstuk 3: Het ontstaan van psychologische stress

In het vorige hoofdstuk is een aantal theorieën besproken waarin psychosociale arbeidsbelasting voorop staat, en waarin relaties worden aangegeven met werkstress. Wat hierbij onder werkstress wordt verstaan, hangt af van de theorie.

De werkstress-literatuur kan niet los worden gezien van de algemene stress-literatuur. In zekere zin is ze slechts een specifiek veld hierbinnen. We kunnen het probleem van het meten van psychosociale arbeidsbelasting en werkstress dan ook bezien vanaf de andere kant: wat zijn de hoofdstromingen in het benaderen van psychologische stress en wat kunnen we hieruit leren met betrekking tot het meten van PSA en werkstress met vragenlijsten?

In paragraaf 3.1 komen een zestal stress-benaderingen aan de orde. De opbouw van dit deel van het hoofdstuk is ontleend aan een overzichtartikel van Hobfoll (1989). In deze paragraaf zijn we op zoek naar wat de ontwikkeling in het algemene stress-onderzoek ons kan leren voor het terrein PSA en werkstress. Aan het eind van de paragraaf proberen we dit samen te vatten (3.1.7).

Verder besteden we in dit hoofdstuk extra aandacht aan twee specifieke theorieën die in het derde deel van dit proefschrift een rol zullen spelen: de sociale vergelijkingstheorie (3.2) en de symptoomperceptie-theorie (3.3). Helemaal aan het eind van dit hoofdstuk trachten we een conclusie te formuleren over de verwachtingen die we mogen hebben van de constructie van een vragenlijst voor het meten van PSA en werkstress, in het licht van de theorievorming over psychologische stress (3.4).

3.1 Zes theorieën over het ontstaan van psychologische stress

3.1.1 Stress als respons

Bij de introductie van de term stress in de psychologie werd deze term ontleend aan de technische wetenschappen. Cannon (1932) en Selye (1950) onderzochten de reacties van mensen en dieren op evidente stressoren uit de omgeving (koude, lawaai). In deze stroming werd stress bestudeerd als een standaardrespons. De stressrespons werd door Selye beschreven als een "general adaptation syndrome". Psychofysiologisch onderzoek naar stressreacties, onder andere in verband met PSA-kenmerken, is nog steeds een actief en relevant terrein van onderzoek. Met name in het kader van onderzoek naar de effecten van mentale en emotionele belasting komen we dergelijk onderzoek tegen (zie bijvoorbeeld Meijman, 1989; van Ouwkerk, Meijman en Mulder, 1994). Dit type onderzoek valt echter buiten het aandachtsgebied van dit proefschrift.

3.1.2 Stress als stimulus

Even eenzijdig als een pure respons-benadering is de benadering van stress-onderzoek die uitsluitend gericht is op een taxonomie van stressvolle gebeurtenissen. Een belangrijke exponent van deze richting is de "life events"-benadering. Voorbeelden van deze stroming zijn te vinden bij Holmes en Rahe (1967), Dohrenwend en Dohrenwend (1974), en Elliot en Eisdorfer (1982). De normale gebeurtenissen in een mensenleven (van geboortes tot sterfgevallen, van verhuizingen tot tentamens) worden hierin gerangschikt naar de mate van stress die ze zouden veroorzaken. Stress wordt hier bestudeerd als stimulus. De "life events"-benadering is ook toegepast op de werksituatie, met beperkt resultaat (House et al., 1986; House, 1987).

Hoe zinvol een dergelijke benadering ook is, het is gaande het onderzoek binnen de "life events"-benadering gebleken dat het hebben van een situatie-taxonomie slechts dient als startpunt om verschillen in reacties van individuele mensen op vergelijkbare situaties te kunnen leren begrijpen (Hobfoll, 1989).

Deze les uit de algemene psychologische stress-theorie staat haaks op de eerder besproken benaderingen van de sociotechniek en het demand-control model die puur gericht zijn op situatieve risico's. Het volgende citaat, dat betrekking heeft op ernstige gebeurtenissen als natuurrampen en oorlog, geeft precies aan waar de schoen wringt:

"Although everyday experience tells us that some immediate emotional disturbance, perhaps with accompanying minor physical symptoms, is common following such occurrences, there is a good deal of evidence that even after apparently major crises only a minority develop serious physical or psychiatric disturbance" (Brown, 1989; p. 4).

Niemand zal willen debatteren over de vraag of een oorlogs- of rampsituatie een risico vormt voor psychische en/of psychosomatische problemen. De wezenlijke vragen beginnen echter pas hierna. Hoeveel mensen zullen problemen hierdoor krijgen? Waarom krijgen sommige mensen wel problemen en andere niet? Kunnen we voorspellen wie het meeste risico lopen?

Als zelfs voor duidelijke en ernstige risico-situaties, zoals rampen en oorlogen, moeilijk is aan te geven hoeveel en welke personen hiervan negatieve gevolgen zullen ondervinden, hoe zou dat dan wel kunnen voor de arbeidssituatie? Hierin bestaan immers doorgaans geen acute risico's voor psychosociale problematiek (traumatische ervaringen tijdens het werk zijn hierop uiteraard een uitzondering). Als dergelijke informatie ontbreekt, hoe kunnen arbeidssituaties dan zodanig worden ontworpen, dat deze een acceptabel risico vormen?

Het is bovendien in de arbeidssituatie de vraag of risico-reductie voor de één (bijvoorbeeld minder inspanning door minder werk in dezelfde tijd; meer verantwoordelijkheid door meer regelmogelijkheden) voor de ander geen risico-toename is (bijvoorbeeld minder plezier in het werk door minder werk in dezelfde tijd; overbelasting door meer regelmogelijkheden). Met name het vitamine-model van Warr (1987) suggereert deze complicatie.

Aldus kan uit de "stress als stimulus"-literatuur een argument tegen een puur situatieve benadering van stress worden afgeleid. Om dit beeld te nuanceren is het echter van belang ook de keerzijde van de medaille te bespreken. Semmer (1992) beargumenteert op zorgvuldige wijze dat de inschatting van hoe stressvol bepaalde gebeurtenissen zijn binnen een bepaalde culturele context grote overeenstemming vertoont tussen individuele mensen. Stress-onderzoekers maken van deze kennis gebruik bij hun keuze van stressvolle situaties voor het bestuderen van stressreacties. Zijn opstelling geeft hij als volgt weer: *"Stress will in the end always be an individual experience but the individual processes involved are co-determined by the historical situation of a given cohort, relate to socially defined roles, and reflect socially shared beliefs and values"* (p. 157). Zo beredeneerd heeft het dus wel degelijk zin om te spreken van stressoren, d.w.z. factoren met een impact op mensen, die tot op zekere hoogte onafhankelijk is van de individuele kenmerken van deze mensen.

3.1.3 Individuele verschillen-benaderingen

Historisch gezien komen we dan bij een volgende stroming binnen de stress-literatuur, die zich tegelijkertijd bezighoudt met stressoren en individuele verschillen in de reacties hierop. Deze stroming kan worden begrepen als een logisch vervolg op de pure "stress als stimulus"-benadering.

Spielberger (1966) initieerde met zijn begrippen toestands-angst en dispositie-angst een hele lijn van onderzoek die het effect van allerlei stressoren onderzocht op individuen met veel dan wel weinig angst. Een van de meest onderzochte onderwerpen in dit verband is tentamenvrees (Sarason, 1975).

Ook in relatie tot arbeidsfactoren zijn nogal wat persoonlijkheidskenmerken onderzocht op hun effecten, zoals: type A-gedrag, hostiliteit, weerbaarheid, beheersingsoriëntatie, negatieve en positieve affectiviteit, prestatie-motivatie, zelfwaardering et cetera (Cooper en Payne, 1991).

Voor het gebied PSA en werkstress is de individuele verschillen-benadering zeker relevant. Qua uitgangspunten heeft de individuele verschillen-benadering verwantschap met het klassieke model belasting-belastbaarheid.

Hierin wordt immers ook de impact van specifieke belastende factoren bestudeerd, rekening houdend met de belastbaarheid van individuele werknemers. Net als bij een persoonlijkheidskenmerk wordt hierbij verondersteld dat de belastbaarheid een stabiel kenmerk is van een individuele werknemer.

Hoeveel aandacht aan individuele verschillen wordt gegeven, zal afhangen van praktische keuzes omtrent het type instrument dat men kiest en voor welke toepassing men dit wil inzetten. Een probleem is ook de onduidelijkheid over welke persoonskenmerken nu het meest belangrijk zijn om mee te nemen in werkstressonderzoek: de lijst met kandidaat-kenmerken neemt in de loop der tijd alleen maar toe. Geen van de kandidaten heeft op dit moment algemene erkenning.

3.1.4 Transactioneel model

Dit is op het moment waarschijnlijk het meest gangbare stressmodel. Stress wordt hierin beschouwd als een niet in balans zijn van de omgevingseisen en het vermogen van de persoon om hier mee om te gaan. Het transactionele model is vooral uitgewerkt door Lazarus en zijn collega's (Lazarus, 1966, 1990; Lazarus en Folkman, 1984).

Centraal in deze visie is het begrip "appraisal". Dit is een cognitief proces dat zich continue afspeelt in reactie op situaties die iemand meemaakt. Het appraisal-proces kent twee fasen: "primary" en "secondary appraisal". In de fase van primary appraisal gaat het om het herkennen van een situatie als bedreigend of anderszins actie vereisend. Bij de secondary appraisal gaat het om een beoordeling van de situatie in relatie tot de mogelijkheden van het individu om met deze situatie om te gaan. Het is vooral vanuit het transactionele model dat in de stress-literatuur interesse is ontstaan voor coping-mechanismen, dat wil zeggen: manieren om met stressvolle situaties om te gaan.

Het transactionele model is zeer expliciet in termen van consequenties voor de te volgen meetstrategie bij stress-onderzoek. Lazarus (1990) formuleert het zo:

"If one takes the position, as I have, that stress is always a product of appraisal, then some confounding between measures of stress and illness is inevitable" (p. 9). "There is no way to separate person and environment in the appraised person-environment relationship without destroying what is meant by psychological stress as conceived in a cognitive-relational theory. Even if a totally objective input or output measure of stress were possible, it would, perforce, mean an abandonment of a cognitive-relational conceptualization and, therefore, describe a totally different phenomenon" (p. 10).

Het transactionele model pleit sterk voor een individu-afhankelijke benadering van het meten van PSA en werkstress, en staat hiermee lijnrecht tegenover elke vorm van theorie die ervan uit gaat dat een risico-vrije arbeidstaak en arbeidssituatie zijn te ontwerpen, onafhankelijk van de eigenschappen van de daarin functionerende individuen.

Het transactionele model is duidelijk te herkennen in modellen als het Michigan-model, het vitamine-model, het JCM, het integraal model van Beehr en Newman, de handelingstheorie, en ook de herziene modellen belasting-belastbaarheid. Wat deze laatste groep van modellen betreft is het vooral de meer dynamische conceptualisatie van het verwerkingsvermogen, die maakt dat deze modellen als transactioneel beschouwd kunnen worden. Eén van de veronderstellingen is namelijk dat de belasting, belastingsverschijnselen en belastingsgevolgen het verwerkingsvermogen beïnvloeden. Hiermee wordt bedoeld dat de wil om een bepaalde arbeidsbelasting te dragen en de bijbehorende belastingsverschijnselen en belastingsgevolgen te tolereren, wordt beïnvloed door een inschatting van diezelfde belasting, belastingsverschijnselen en belastingsgevolgen. Er zit derhalve een duidelijk "appraisal"-element in de herziene belasting-belastbaarheid modellen.

3.1.5 Zelfregulatie-model

Net als het transactionele model is het zelfregulatie-model van Leventhal en zijn collega's een cognitief-psychologisch model (Leventhal en Mosbach, 1983; Leventhal, Nerenz en Steele, 1984). Dit model stelt dat mensen gemotiveerd zijn om bedreiging van hun gezondheid te reguleren c.q. minimaliseren, en dat zij op een manier zullen omgaan met bedreigingen van hun gezondheid die consistent is met hun perceptie van die bedreigingen. Het model veronderstelt drie stappen in het omgaan met bedreigende situaties: het vormen van een beeld van de dreiging (mentale representatie), het onderzoeken en uitvoeren van copingmogelijkheden, en het beoordelen van het effect van coping. Dit model veronderstelt verder dat deze drie stappen parallel aan elkaar worden uitgevoerd op het cognitieve vlak en op het emotionele vlak.

Het zelfregulatie-model is vooral toegepast bij allerlei patiënten-groepen (kanker, hart- en vaatziekten) en gezondheidsbedreigende handelingen (roken, medische ingrepen) (Leventhal, 1992). Het model is nog niet toegepast in de arbeidssituatie. Er is echter een verwantschap met de symptoomperceptie-theorie van Pennebaker (1982), die wel de nodige aandacht heeft gekregen (zie 3.3).

De relevantie van dit invloedrijke model in de gezondheidspsychologie voor het veld PSA en werkstress ligt hierin dat, net als bij het transactionele model, een centrale rol is toebedeeld aan het appraisal-proces. De meerwaarde van dit model zit in het begrip "mentale representaties".

Als we het model van Leventhal volgen is het voor het oplossen van werkstress-problematiek relevanter om deze mentale representaties te meten dan de objectieve situatie zelf.

Dit begrip van mentale representaties sluit goed aan bij de visie op nut en interpretatie van vragenlijstgegevens op het gebied van PSA en werkstress, zoals we die vinden bij Spector (1992).

3.1.6 Conservation of resources-model

Hobfoll (Hobfoll, 1988, 1989; Hobfoll en Stokes, 1988) formuleerde een stress-theorie die afwijkt van alle andere, door te stellen dat het omgaan met bedreigingen c.q. belastende factoren niet het centrale onderwerp is voor stress-onderzoek. In plaats hiervan stelt hij de mens voor als primair gericht op het opbouwen, beschermen en behouden van hulpbronnen (resources). Hobfoll onderscheidt vier soorten hulpbronnen: objecten, condities, persoonskenmerken, en energie.

Stress moet volgens Hobfoll vooral worden begrepen vanuit het dreigende dan wel feitelijke verlies van hulpbronnen. In tegenstelling tot de meeste andere modellen voorspelt dit model dat, in het geval van een bedreigende situatie, het gedrag niet primair zal worden gericht op het omgaan met de bedreigende factor, maar op het minimaliseren van het verlies van hulpbronnen.

De theorie van Hobfoll kan in tegenstelling tot andere stress-theorieën verklaren, waarom mensen in de afwezigheid van echte bedreigingen toch actief zullen zijn. Hun gedrag richt zich in dit geval namelijk op het vergroten van hun hulpbronnen.

De theorie van Hobfoll is te nieuw om in de context van de arbeidssituatie veelvuldig uitgeteerd te zijn. Hijzelf paste de theorie toe op het begrip burnout (Hobfoll en Freedy, 1993). Vanuit zijn redenering ziet het onderzoek binnen het veld PSA en werkstress er nogal beperkt en eenzijdig uit. Veel onderzoek richt zich uitsluitend op het verkennen van mogelijke stressoren, en nauwelijks op de doelen die mensen met het werk nastreven en/of de hulpbronnen die ze ermee opbouwen en beschermen. Een auteur die hier wel rechtstreeks op ingaat is Semmer (1995). Ook (de dreiging van) het verlies van hulpbronnen is binnen de context van de arbeidssituatie nog weinig bestudeerd. Een uitzondering hierop vormt het onderzoek naar de effecten van (dreigend) verlies van een baan (Warr, 1987).

3.1.7 Betekenis van de algemene stresstheorie voor het veld PSA en werkstress

Wanneer we de theorieën zoals besproken in hoofdstuk 2 bezien in het kader van de algemene psychologische stresstheorie, dan valt in de eerste plaats op dat er een zwaar accent ligt op de situatie-aspecten. Dit was gezien het toepassingsgebied uiteraard te verwachten.

Het sociotechnisch model en het Karasek-model maken een heldere keuze met betrekking tot de achterliggende algemene stresstheorie, namelijk: stress als stimulus. Bij het klassieke model belasting-belastbaarheid herkennen we een individuele verschillen-benadering. De overige modellen zijn te plaatsen binnen een transactionele benadering van stress. Daarmee is de transactionele benadering de meest dominante benadering in het veld PSA en werkstress van dit moment.

De gebruikelijke strategie bij onderzoek naar PSA en werkstress is, gezien vanuit de algemene psychologische stresstheorie, vrij smal. Het gaat meestal om een evenwicht- of balansmodel tussen persoon en situatie. De diverse modellen verschillen vooral in de manier en mate waarin ze het concept van evenwicht of balans expliciet in het model inbouwen.

Bij het Michigan-model, de herziene modellen belasting-belastbaarheid, en het integraal model van Beehr en Newman zit het evenwicht/balans-begrip expliciet in de theorie. Bij de taakkenmerken-benadering komt het evenwicht/balans-begrip binnen via de "critical psychological states".

Het vitamine-model werpt op deze materie een iets ander licht. Hierin wordt gesteld dat we inderdaad moeten kijken naar de interactie persoon-situatie, maar dat dit geen simpel additief of algemeen evenwicht/balans-verhaal is: sommige factoren kunnen een specifiek eigen "toxisch" effect hebben. Deze factoren vormen met andere woorden pas bij een bepaalde dosis een risico. Dit laat de mogelijkheid open, dat de drempels voor het optreden van "toxische" reacties individueel zijn bepaald.

Dit alles wekt de indruk dat het veld PSA en werkstress, met uitzondering van het werk van Warr, qua achterliggende psychologische stresstheorie sinds de jaren zeventig stilstaat bij het transactionele model. Dit blijkt ook hieruit dat recentere modellen, zoals die van Leventhal en Hobfoll, in het veld PSA en werkstress nog niet of nauwelijks worden toegepast.

Op deze stand van zaken bestaan twee uitzonderingen. Dit zijn de symptoomperceptie-theorie (Pennebaker, 1982) en een aantal theorieën op het gebied van de sociale vergelijking (Buunk en De Vries, 1991). Deze twee benaderingen zijn in de afgelopen jaren veelvuldig gebruikt in relatie tot PSA en werkstress. Ze komen bovendien terug in deel III van dit proefschrift. We gaan hier dan ook dieper op deze benaderingen in.

3.2 Sociale vergelijkingstheorie

De sociale vergelijkingstheorie vindt haar oorsprong in het werk van Festinger (1954). Volgens Festinger hebben mensen een motivatie om de juistheid van eigen meningen en het niveau van eigen bekwaamheden te evalueren. Hoe minder objectieve informatie beschikbaar is om een dergelijke evaluatie te maken, des te belangrijker sociale vergelijking wordt als bron van evaluatie.

Hoe meer iemand anders op een persoon lijkt qua meningen en bekwaamheden, des te groter de neiging van de persoon wordt om zich met zo iemand te vergelijken (gelijkheidshypothese). Festinger geeft een aantal verschillen aan in de werking van sociale vergelijkingsprincipes bij meningen versus bekwaamheden. Met zijn werk heeft Festinger de basis gelegd voor cognitieve sociaal-psychologische theorieën, die op het gebied van arbeid en gezondheid gebruikt beginnen te worden. Een voorbeeld hiervan is de theorie van gepland gedrag (Ajzen, 1991).

Sociale vergelijkingsprincipes zijn behalve op meningen en bekwaamheden ook toe te passen op emoties en op zaken als status en bezit. Aan sociale vergelijking op het gebied van emoties is de naam verbonden van Schachter (1959). Hij onderzocht de reactiepatronen van mensen in bedreigende situaties en constateerde dat mensen in dergelijke situaties sociale vergelijkingsprincipes gebruiken om te bepalen hoe zij zich zouden moeten voelen. Dit vormde de basis voor zijn cognitieve emotie-theorie, waarin hij stelt dat mensen in onzekere situaties zoeken naar een adequaat label voor hun gevoelens. Sociale vergelijking is in dergelijke situaties een mogelijke bron voor zo'n label, en wel meer naarmate de situatie minder duidelijk is.

Sociale vergelijking op het gebied van status en bezit heeft vooral gestalte gekregen in de vorm van de billijkheidstheorie van Adams (1965). Volgens Adams streven mensen naar "consonantie", d.w.z.: mensen streven naar een evenwicht tussen eigen inbreng en opbrengsten. Hierbij vergelijkt men zichzelf met relevante anderen. Het is makkelijk om de theorie van Adams toe te passen op de arbeidssituatie. Als vormen van inbreng kunnen hier bijvoorbeeld worden gezien: inzet, opleiding, kennis, vaardigheden, sociale steun voor anderen. Als opbrengsten kunnen in dit verband bijvoorbeeld worden gezien: beloning, waardering, erkenning. Is de ruilrelatie tussen inbreng en opbrengsten niet in balans (vergeleken met relevante anderen), dan wordt dit ervaren als onbillijkheid. In een dergelijk geval zal de werknemer proberen deze onbillijkheid op te heffen.

Met name op het gebied van de beloning vindt de billijkheidstheorie in Nederland reeds sinds geruime tijd navolging (Thierry, 1969; Von Grumbkow, 1980; Syroit, 1984). Het heeft langer geduurd alvorens onderzoek op gang kwam met betrekking tot werkstress.

De billijkheidstheorie mag zich in Nederland de laatste jaren echter verheugen in een grote belangstelling vanuit het onderzoeksveld arbeid en gezondheid (Van Yperen, 1995; Geurts, 1995). De theorie is toegepast op onderwerpen als burnout (Van Yperen, Buunk en Schaufeli, 1992; Buunk en Schaufeli, 1993; Van Gorp, Schaufeli, en Hopstaken, 1993; Van Dierendonck en Schaufeli, 1993; Schaufeli, Van den Eijnden, Brouwers, 1994), arbeidsongeschiktheid (Ybema en Buunk, 1992; Van Yperen en van Horn, 1993), ziekteverzuim (Geurts, Buunk en Schaufeli, 1991; Geurts, 1994; Van Yperen, Hagedoorn en Geurts, 1994) en personeelsverloop (Janssen en Buunk, 1990; Van Yperen, Hagedoorn en Geurts, 1994).

Als voorbeeld van dergelijk onderzoek bespreken we een onderzoek van Geurts (1994) over sociale vergelijking en ziekteverzuim. Zij verrichtte onderzoek bij medewerkers van twee metaalbedrijven. De werknemers vulden vragenlijsten in waarin vragen waren opgenomen over ervaren onbillijkheid (EO) in de ruilrelatie met het bedrijf, wrokgevoelens jegens het bedrijf, verwachtingen over de groepsnorm ten opzichte van ziekteverzuim, en persoonlijke normen ten aanzien van ziekteverzuim. Van deze werknemers werden tevens gegevens verzameld over de verzuimfrequentie in de twaalf maanden na het vragenlijstonderzoek.

Geurts ontwikkelde een LISREL-model op deze gegevens, waarbij de verzuimfrequentie wordt verklaard uit EO en de persoonlijke norm ten aanzien van verzuim. De persoonlijke norm wordt in dit model verklaard uit de verwachte groepsnorm en EO. Bij beide metaalbedrijven wordt EO als significante voorspeller gevonden van de verzuimfrequentie. Slechts bij één van de twee bedrijven is ook de persoonlijke verzuimnorm een significante voorspeller van verzuim.

Bij dezelfde populatie werden ook nog andere vragenlijsten afgenomen, namelijk over: ervaren gezondheidsklachten, attributie van de gezondheidsklachten aan het werk, en de mate van communicatie over problemen op het werk.

Op deze gegevens werd een tweede LISREL-model ontwikkeld waarin de verzuimfrequentie wordt voorspeld uit EO en ervaren gezondheidsklachten. EO wordt verklaard vanuit het toeschrijven van gezondheidsklachten aan het werk, wat weer verklaard wordt uit ervaren gezondheidsklachten en communicatie over werkproblemen. Ditmaal zijn zowel EO als ervaren gezondheid significante voorspellers van de verzuimfrequentie in beide metaalbedrijven.

Beide LISREL-modellen worden met elkaar gecombineerd tot een geïntegreerd model. Het wekt geen verbazing dat ook in dit model EO een significante voorspeller blijkt te zijn van de verzuimfrequentie. Uit de gerapporteerde regressiecoëfficiënten mogen we afleiden dat EO variantieniveaus kan voorspellen in verzuimgegevens van tussen de 5 à 10%.

Uit het onderzoek van Geurts (1994) komt naar voren dat EO een goede voorspeller is van ziekteverzuim. Het is jammer dat Geurts niet tegelijkertijd maten voor PSA meeneemt in haar onderzoek, of je zou EO als zodanig moeten opvatten. PSA (geoperationaliseerd met zeven schaa scores) verklaart namelijk eveneens zo'n 5 à 10% van de variantie in individuele verzuimcijfers (Smulders en De Winter, 1993). Omdat Geurts eenzelfde hoeveelheid variantie verklaart op basis van een enkele maat voor EO, terwijl Smulders en De Winter hiervoor zeven schaa scores nodig hebben, dringt zich de vraag op, of genuanceerd onderzoek naar PSA in verband met ziekteverzuim wel zin heeft.

In het derde deel van dit proefschrift zullen we proberen de vraag van de voorspelbaarheid van individuele verzuimgegevens te beantwoorden, waarbij EO en PSA-factoren gelijktijdig als determinanten worden onderzocht.

3.3 Symptoomperceptie-theorie

In de literatuur over PSA en werkstress is een sterk accent te vinden op factoren uit de omgeving die inwerken op de persoon, en die worden gezien als determinanten van werkstress. Opmerkelijk in dit verband is de symptoomperceptie-theorie van Pennebaker (1982; 1983; 1984), die stelt dat stimuli uit de omgeving en uit het lichaam van het individu parallel aan elkaar cognitief worden verwerkt, en derhalve dezelfde verwerkingsprincipes ondergaan. Twee belangrijke verwerkingsprincipes in dit verband zijn: de competitie van stimuli en het selectieve zoekproces.

Stimuli uit de omgeving en uit het lichaam zijn volgens Pennebaker (1982; 1983; 1984) voortdurend in competitie met elkaar om aandacht van het centrale cognitieve verwerkingsapparaat. De waarneming van lichamelijke symptomen is hierdoor afhankelijk van de omgevingsinformatie en vice versa. In hoeverre lichamelijke fenomenen als symptomen zullen worden waargenomen hangt echter ook af van een selectief zoekproces. Doelen en verwachtingen die de waarnemer heeft zullen leiden tot actieve manieren om binnenkomende stimuli te selecteren en aandacht te geven. Soms zal het selectieve zoekproces zelfs zo sturend zijn voor de waarneming dat kan worden gesproken van een zogenaamd "ziekteschema". Hiermee wordt bedoeld een cognitieve structuur, waarin lichamelijke symptomen, emotionele reacties, perceptuele herinneringen en beelden van ziekte zijn gecombineerd.

Het begrip "ziekteschema" komt dicht in de buurt van het begrip "mentale representaties", zoals dat wordt gehanteerd door Leventhal in zijn zelfregulatie-model. De theorieën van Leventhal en Pennebaker liggen in elkaars verlengde, zelfs zodanig dat ze in elkaar zijn geschoven tot één model door Van Vliet en Gijsbers van Wijk (1990).

De symptoomperceptie-theorie levert voor de vragenlijst-georiënteerde werkstressliteratuur een fundamentele complicatie op. Als de zelfbeoordelingen van PSA (bijvoorbeeld de relatie met collega's) en de zelfbeoordelingen van werkstress (bijvoorbeeld spanningsverschijnselen) tot stand komen op basis van hetzelfde centrale verwerkingsprincipe, dan is het mogelijk dat gevonden statistische verbanden tussen stress-determinanten en stress-verschijnselen zijn terug te voeren op gezamenlijke eigenschappen van het cognitieve verwerkingsproces. Er kan met andere woorden sprake zijn van schijnverbanden.

Watson en Pennebaker (1989) introduceren een kenmerk van het centrale verwerkingsproces, waarvan zij aantonen dat het stressor-strain relaties medieert: het persoonskenmerk "negatieve affectiviteit" (NA). Dit is vrij te vertalen als een voorkeur voor negatieve gemoedstoestanden (angst, afkeer, kwaadheid, verachting, schuld, depressie). Omdat NA correleert met subjectieve metingen van zowel stressoren als strains, maar niet of veel minder sterk met objectieve maten van gezondheid en welbevinden, betekent dit dat NA als een algemene, verstorende variabele werkt in het onderzoek naar de relatie tussen arbeid en gezondheid. NA is qua inhoud sterk gerelateerd aan het begrip neuroticisme, waarvoor een soortgelijk verstorend effect wordt verondersteld (McCrae, 1990).

Watson en Pennebaker (1989) rapporteren correlaties in de range van .30 tot .50 tussen NA en uiteenlopende maten van gezondheidsklachten in uiteenlopende onderzoekspopulaties. Vervolgens bieden zij een overzicht van onderzoek naar drie hypothesen die een verklaring kunnen bieden voor het gevonden fenomeen. In de eerste plaats de psychosomatische hypothese: hoge NA-scores als determinanten van gezondheidsklachten. Een tweede hypothese is de ongeschiktheidshypothese: hoge gezondheidsklachtenscores als determinanten van NA. Tenslotte onderzoeken zij de symptoomperceptie-hypothese, zoals die hierboven is weergegeven.

Er wordt niet veel evidentie gevonden die de eerste twee hypothesen kan bevestigen. De symptoomperceptie-hypothese blijkt het meest bruikbaar. Als belangrijkste argument hiervoor rapporteren Watson en Pennebaker hoge correlaties tussen stressoren en stress-uitkomsten (gemeten met vragenlijsten) versus lage correlaties tussen stressoren (gemeten met vragenlijsten) en objectieve indicatoren van gezondheid.

In de literatuur over PSA en werkstress is na het artikel van Watson en Pennebaker een reeks onderzoeken gedaan waarin hun claim is onderzocht. Hieronder bespreken we enkele hiervan:

Brief, Burke, Robinson en Webster (1988) deden vragenlijstonderzoek bij 497 medewerkers van een verzekeringsmaatschappij in de VS. Behalve NA werden vragenlijsten afgenomen over stressvolle gebeurtenissen op het werk en privé, arbeidssatisfactie, gezondheidsklachten op het werk, emoties op het werk, algemene tevredenheid, en depressiviteit.

We zijn hier vooral geïnteresseerd in de correlaties tussen werkstressoren en werkgerelateerde uitkomsten (arbeidssatisfactie, gezondheidsklachten op het werk en emoties op het werk). Zij berekenden eerste-orde correlaties tussen stressoren en stress-uitkomsten, en vervolgens partiële correlaties na correctie voor NA. De correlatie tussen werkstressoren en arbeidssatisfactie zakte hierdoor van gemiddeld $-.37$ naar gemiddeld $-.31$. De correlatie tussen werkstressoren en gezondheidsklachten op het werk liep terug van gemiddeld $.21$ naar gemiddeld $.03$. Tenslotte de correlatie tussen werkstressoren en emoties in het werk: deze ging terug van gemiddeld $.37$ naar gemiddeld $.22$.

Als titel van hun artikel stellen Brief et al. (1988) de vraag *"Should negative affectivity remain an unmeasured variable in the study of job stress?"*. Gezien deze resultaten geven zij hierop het volgende antwoord: *"The current results collectively demonstrate that NA should not remain an unmeasured variable in the study of job stress. Both its role as a methodological 'nuisance' factor as examined here and (we believe, more important) as a plausible cause of substantive stressful work events warrants additional attention"* (p. 197).

Een volgende studie waarin aandacht is besteed aan NA is die van Parkes (1990). Bij 157 werknemers uit een Britse onderwijs-populatie vonden twee meetmomenten plaats met vragenlijsten, met een tussenliggende periode van twee maanden. Bij moment 1 werden demografische gegevens verzameld. Bij moment 2 werden vragen gesteld over coping, eisen in het werk, steun in het werk en geestelijke gezondheid (met de General Health Questionnaire (GHQ)). De GHQ-scores werden als afhankelijke variabele genomen in een hiërarchische multiële regressie-analyse. Bij de eerste stap werden personalia en NA ingevoerd in de vergelijking: beide zijn goed voor 8% van de verklaarde variantie. In de tweede stap werden in eerste instantie eisen en steun in het werk ingevoerd. Deze voegen samen 12% verklaarde variantie toe. Daarna werden in de derde stap interactietermen ingevoerd van NA met eisen en steun in het werk, wat goed is voor 7% additionele verklaarde variantie. De vierde en laatste stap voegt coping toe aan de vergelijking en de interactie van coping met eisen en steun in het werk. Deze zijn beide goed voor 5% verklaarde variantie. In totaal verklaart de regressievergelijking hiermee 43% van de variantie in de GHQ-scores.

Voor ons is van belang dat PSA (eisen en steun in het werk) nog een flink deel additionele variantie verklaart bovenop het effect van personalia en NA. De resultaten van Parkes ondersteunen die van Brief et al. (1988). Het gegeven dat in haar analyses ook de interactieterm van NA en werkstressoren een significant effect heeft legt Parkes zo uit, dat NA behalve als "nuisance factor" voor alle vormen van werkstressoren in relatie tot stress-uitkomsten, tevens optreedt als factor die de reactiviteit van een individu bepaalt op bepaalde werkstressoren. In haar onderzoek treedt dit effect vooral op bij kwantitatieve taakeisen, niet zozeer bij interpersoonlijke aspecten.

Door Chen en Spector (1991) zijn resultaten gerapporteerd die de NA-hypothese minder sterk bevestigen. Een heterogeen samengestelde groep van 400 werknemers in Florida vulde vragenlijsten in over stressoren, stressverschijnselen (arbeidssatisfactie, emotionele reacties, gezondheidsklachten, verloopintentie, ziekteverzuim, artsbezoek), en NA. In tegenstelling tot Brief et al. (1988) vonden zij slechts kleine verschillen tussen eerste-orde en partiële correlaties van stressoren met stressverschijnselen. Hun conclusie is de volgende: *"Except for physical strains, there was little evidence to support the idea that NA accounts for a large amount of the relation between chronic work stressors and strains"* (p. 403). Chen en Spector schrijven hun afwijkende resultaten ten opzichte van Brief et al. (1988) voor een deel toe aan de kleinere mate van item-overlap tussen de schalen die zij gebruiken voor het meten van stressoren en strains.

De gedeelde variantie tussen stressoren en strains ligt in de range van 5% tot 30%. Chen en Spector bekeken hoe groot de reductie is in gedeelde variantie wanneer word gecorrigeerd voor NA (reduction rates). Als voorbeeld nemen we hier de reduction rates voor de stressor "rolonduidelijkheid". Deze reduction rates liggen voor emotionele reacties in de buurt van de 30% tot 40%, voor arbeidssatisfactie op 15%, voor gezondheidsklachten op 60% tot 100%, voor zelfgerapporteerd verzuim op 55% tot 75%, en voor verloopintentie op 25% à 30%. De reduction rate wisselt dus nogal per soort variabele en is duidelijk het grootst bij gezondheidsklachten en het kleinst bij arbeidssatisfactie. Dit is analoog aan de hierboven beschreven resultaten van Parkes (1990) en Brief et al. (1988). Opmerkelijk is ook dat de reduction rate samenhangt met de hoogte van de correlatie tussen rolonduidelijkheid en de diverse uitkomstmaten. Dit leidt tot de volgende slotconclusie: *"These results suggest that NA tends to account for a relatively constant amount of variance, which becomes proportionally smaller as the relations between stressors and strains increase"* (p. 403).

Parkes, Mendham en Von Rabenau (1994) onderzochten personeel werkzaam in de Britse psychiatrie (N=145). Zij voorspelden arbeidssatisfactie en gezondheidsklachten met een multiële hiërarchische regressietechniek. Stap 1 bracht leeftijd, geslacht en functiotype in de vergelijking in, stap 2 NA, stap 3 psychologische taakeisen, taakautonomie, en sociale ondersteuning, en stap 4 de interactie van psychologische taakeisen, taakautonomie, en sociale ondersteuning (cf. Karasek en Theorell, 1990).

Voor arbeidssatisfactie zijn de unieke bijdragen van deze vier stappen aan de verklaarde variantie respectievelijk: 9%, 9%, 18%, en 2%. Voor gezondheidsklachten zijn deze getallen respectievelijk: 8%, 19%, 3%, en 4%. Haar conclusie is dan ook de volgende: *"These results suggest that symptom reports are most strongly predicted by individual differences in negative affectivity, with the work environment making a smaller contribution. Conversely satisfaction was more strongly associated with the perceived work environment"* (p. 109). Net als in de eerder besproken onderzoeken is dus ook in deze studie een verschil gevonden van het effect van NA op gezondheidsklachten versus arbeidssatisfactie.

Tenslotte bespreken we het onderzoek van Elliott, Chartrand en Harkins (1994). Zij onderzochten 127 leraren en 126 journalisten in de VS met een vragenlijst. De meest belangrijke werkstressor werd op individuele basis geïdentificeerd met een open vraag. Vervolgens werden voor deze stressor de volgende vragen beantwoord: "hoe moeilijk is deze stressor door u te tolereren?", "hoeveel interfereert deze stressor met uw arbeidsprestatie?" en "hoeveel controle heeft u over deze stressor?". Hiernaast werd NA gemeten en emotionele stressverschijnselen.

Als voorbeeld van de resultaten bespreken we de relatie tussen de moeilijkheid om de stressor te tolereren en emotionele stress-verschijnselen. De gedeelde variantie hiertussen is bij de leraren 26%, bij de journalisten 25%. Wanneer eerst NA in een regressievergelijking wordt ingevoerd en vervolgens de moeilijkheid om de stressor te tolereren dan blijkt NA een bijdrage te leveren aan emotionele stress van 26% bij leraren en 17% bij journalisten.

Hier bovenop draagt de moeilijkheid om de stressor te tolereren nog eens 12% (leraren) c.q. 14% (journalisten) bij. De verklaarde variantie door de stressor wordt dus ongeveer gehalveerd, maar blijft een significante eigen bijdrage behouden bovenop het effect van NA. Elliott et al. (1994) concluderen dan ook: *"While trait negative affectivity may confound the prediction of stress-related outcomes, the extent of this contaminating effect on individual difference constructs may be greatly exaggerated"* (p. 197).

Als hoofdlijn kunnen we uit het onderzoek met NA halen dat psychosociale arbeidsbelasting ook na correctie voor NA een significant effect zal hebben op werkstress. Het is aannemelijk dat NA meer vertekenend zal werken op de relatie tussen PSA en spanning (verwachte verklaarde variantie door NA: 15% à 20%, dan op de relatie tussen PSA en welbevinden (verwachte verklaarde variantie door NA: 5 à 10%).

3.4 Betekenis van de algemene stresstheorie voor de vragenlijstconstructie

In deel II van dit proefschrift wordt een vragenlijst ontwikkeld voor de meting van PSA en werkstress. Zowel EO als negatieve affectiviteit krijgen de laatste tijd veel aandacht in het onderzoeksveld arbeid en gezondheid, en doen de claim dat zij werkstress-uitkomsten kunnen verklaren op basis van cognitieve principes, die voor een deel onafhankelijk opereren van de psychosociale arbeidsbelasting.

Het is voor het meten van PSA en werkstress dan ook van belang om vast te stellen of de volgende stelling waar is: EO en NA verklaren grote delen van de variantie in de beoordeling van PSA en werkstress. Voor zover bekend zijn deze twee aspecten nog nooit eerder gelijktijdig onderzocht qua effecten op PSA en werkstress.

Verder is het van belang om te onderzoeken of, na correctie voor personalia, NA, en sociale vergelijking, PSA nog een additionele bijdrage levert aan de verklaring van werkstress. Als dit niet zo is, heeft vragenlijstonderzoek naar PSA in relatie tot werkstress weinig zin. Als PSA een eigen bijdrage blijkt te leveren, is deze dan voor de diverse werkstress-verschijnselen even groot, of is het contaminerende effect van NA en sociale vergelijking voor sommige werkstress-verschijnselen groter dan voor andere? In deel III van dit proefschrift zullen we ons met de hier geformuleerde vragen bezighouden.

Een laatste opmerking betreft het vertrekpunt voor de vragenlijstconstructie, zoals dat is geformuleerd in paragraaf 2.9. Als we de keuze van de aspecten die in de vragenlijst gemeten moeten worden gaan bepalen op grond van een facetanalyse van de literatuur, dan is de kans groot dat we uitkomen bij een operationalisatie in de traditie van het transactionele model van stress. We zagen immers in paragraaf 3.1.7, dat dit de achterliggende psychologische stresstheorie is van de meeste theorieën op het gebied van PSA en werkstress. Van de te construeren vragenlijst hoeven we qua inhoud dan ook geen theoretische vernieuwing te verwachten: dit zit ingebakken in het gekozen vertrekpunt.

Hoofdstuk 4: Het meten van PSA en werkstress met vragenlijsten

In de vorige twee hoofdstukken zijn theorieën besproken met betrekking tot PSA en (werk-)stress. Bij het bespreken van deze theorieën zijn we de voorkeuren van bepaalde theorieën voor bepaalde methoden tegengekomen. Ook is al een aantal opmerkingen gemaakt over de vragenlijstmethode in het bijzonder, gezien vanuit de besproken theorieën.

Alvorens we ons in het tweede deel van dit proefschrift bezig gaan houden met de constructie van een vragenlijst voor het meten van PSA en werkstress, gaan we in dit hoofdstuk in op de algemene voor- en nadelen van een vragenlijstmethode tegenover andere methoden (4.1), de mogelijkheden en beperkingen van zo'n methode op het gebied van PSA en werkstress (4.2), en op de eisen waaraan een goede vragenlijst moet voldoen ten behoeve van gebruik in de praktijk (4.3). Vooraf moge reeds duidelijk zijn dat wat de auteur betreft de voordelen van een vragenlijstmethode duidelijk opwegen tegen de nadelen, en dat - mits oordeelkundig gebruikt - de vragenlijstmethode groot belang heeft voor de praktijk.

Een tweede opmerking vooraf: naast, op het oordeel van de werknemer zelf gebaseerde, vragenlijstmethoden wordt er in het onderzoek van psychosociale aspecten van arbeidsbelasting ook gebruik gemaakt van andere methoden, zoals observaties (al dan niet met behulp van daartoe geëigende instrumenten) en oordelen van experts, supervisors, of collega's. Aan de methode van vragenlijst-onderzoek die zich baseert op het oordeel van de werknemer zelf, zullen we in dit hoofdstuk refereren als de "self-report" methode. De vragenlijstmethode is één van de mogelijke uitwerkingen van een self-report methode (andere uitwerkingen zijn bijvoorbeeld rating-scales en interesse-tests). Aan de andere methoden zullen wij refereren als "geobjectiveerde" methoden. Hoewel het woord "geobjectiveerd" vaak meer onpartijdigheid en/of werkelijkheid suggereert dan in de praktijk kan worden waargemaakt, conformeren wij ons aan het gebruik van deze term in de literatuur.

4.1 De vragenlijstmethode

4.1.1 Voor- en nadelen van de vragenlijstmethode

Zapf (1989) somt de volgende mogelijke nadelen op van een vragenlijstmethode:

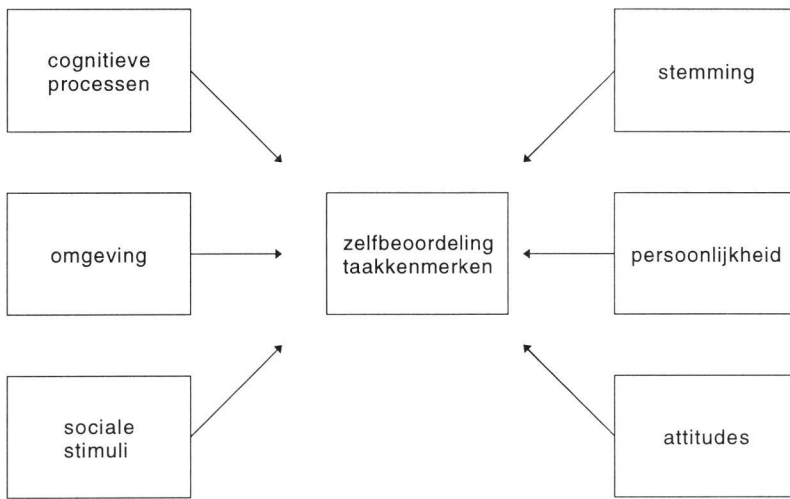
- de neiging om antwoorden rond het gemiddelde te geven
- vertekening door persoonskenmerken van de werknemer (bijvoorbeeld klagerigheid)
- beantwoording vanuit een impliciete theorie die de werknemer heeft over doel en inhoud van het onderzoek

- beantwoording van een lijst vanuit een eerste indruk die een werknemer heeft van de vragen (halo-effect)
- hoe eenvoudiger en minder autonoom de arbeidstaak, des te groter de kans op vertekening door gebrek aan vergelijkingsmateriaal bij de werknemer
- de ambiguïteit van de taal, waardoor werknemers dezelfde vraag op verschillende manieren interpreteren
- het gebrek aan concreetheid van vragen leidt tot verschillen in de interpretatie door werknemers
- gevoeligheid voor sociale wenselijkheid
- de antwoorden worden meestal retrospectief gegeven, wat tot een beperkte relevantie kan leiden naar de huidige situatie toe
- de antwoordpolen worden door de werknemers verschillend geïnterpreteerd
- de beantwoording is afhankelijk van de mate van ervaring van de werknemer met een bepaald onderwerp
- wanneer een arbeidstaak weinig beroep doet op het bewustzijn van een werknemer, dan levert dit vertekende antwoorden op bij vragen over die arbeidstaak
- gevoeligheid voor verdringingsprocessen (ontkenning)

Samengevat is de belangrijkste zwakte van de vragenlijstmethode, dat onduidelijk is in hoeverre de beoordeling van een situatie door een respondent overeenkomt met de objectieve situatie.

Spector (1992) vat desalniettemin de stand van zaken van onderzoek over deze kwestie als volgt samen: *"There is considerable evidence that perceptual measures reflect the objective environment. There is also considerable evidence that they are affected by additional factors, such as attitudes, moods, cognitive processes and dispositions. Although many have criticized the self-report approach as being flawed by method effects, there is surprisingly little evidence that observed relations are attributable to the self-report method"* (p. 143).

Door voorstanders van een geobjectiveerde meting van PSA wordt de self-report methode beschouwd als een zwakke manier van meten: een minimum-strategie. De draai die Spector (1992) hieraan geeft, is dat een vragenlijstmeting juist iets toevoegt aan een geobjectiveerde meting: de impact van een situatie op mensen. Belangrijk is niet of een self-report meting al dan niet iets zegt over de objectieve situatie. Dat is naar zijn mening ondertussen voldoende onderbouwd. Interessanter (want informatiever) is het om te verklaren waarom een self-report methode tot andere conclusies leidt dan een geobjectiveerde meting. Hiervoor zou eigenlijk meer bekend moeten zijn over de manier waarop een self-report tot stand komt, en over de invloed die factoren als belastbaarheid, persoonlijkheid en loopbaan-fase hierop uitoefenen. Figuur 4.1 illustreert welke factoren Spector (1992) van belang acht bij de totstandkoming van een zelfbeoordeling van taakkenmerken.



Figuur 4.1: Mogelijke oorzaken van zelfbeoordelingen van taakkenmerken. Ontleend aan Spector (1992)

Uit onderzoek zoals uitgevoerd door Zapf (1989) blijkt dat een vragenlijstmeting, afhankelijk van het gemeten begrip, al dan niet goed overeenkomt met een geobjectiveerde meting. Zo geven de antwoorden op een vragenlijstschaal over taakvariatie een goede correspondentie te zien met taakvariatie en -complexiteit, zoals gescoord door getrainde beoordelaars. Voor een vragenlijstschaal werkdruk blijkt een dergelijke correspondentie echter nagenoeg afwezig te zijn.

Op grond van onderzoek zoals dat van Zapf (1989) mogen we concluderen dat het radicaal afwijzen van self-report methoden onverstandig is. Het is beter om aan te geven welke begrippen het beste met een self-report methode kunnen worden gemeten, en welke begrippen het beste met andere methoden.

Zoals wel vaker het geval is, schuilt in de gesignaleerde zwakte van de vragenlijst-methode dus ook een sterkte. De vragenlijstmethode is uitermate geschikt waar het gaat om het vaststellen van de mening van mensen. Ook de beleving van symptomen en gevoelens door een werknemer leent zich goed voor een self-report methode, omdat die alleen via introspectie kan worden waargenomen.

Soms kan hierdoor een probleem eerder worden gesignaleerd dan wanneer we zouden wachten tot het manifest wordt, wat zeker op het terrein van preventie van psychische en psychosomatische klachten belangrijk is.

4.1.2 Alternatieven

Het is, gezien de genoemde beperkingen van een self-report methode, voor de hand liggend om naar alternatieven te kijken. Deze zijn grofweg te onderscheiden in al dan niet geïnstrumenteerde observaties en registraties, en beoordelingsmethoden door experts, supervisoren en collega's. Deze methoden zijn echter ook niet zonder problemen.

Observaties en registraties zijn vaak omslachtig, tijdrovend, en dus duur. Dit geldt bijvoorbeeld voor medische onderzoeken als biometrie, audiometrie, ECG, bloedproeven en dergelijke. Het ontwikkelen van self-report methoden is vaak een rechtstreeks gevolg van het feit dat meer geobjectiverde methoden te duur c.q. te moeilijk realiseerbaar zijn bij grote onderzoekspopulaties. Bovendien worden geobjectiverde methoden in de praktijk vaak zo weinig gestandaardiseerd toegepast, dat de resultaten op populatieniveau slechts beperkt bruikbaar zijn.

Anders ligt de situatie voor registraties van ziekteverzuim, personalia, functiekenmerken en dergelijke, die bij bedrijven beschikbaar zijn. Deze zijn (op bedrijfsniveau) meestal betrouwbaar en worden dan ook veelvuldig gebruikt bij onderzoek naar PSA. Hier loopt de onderzoeker echter regelmatig tegen privacy-aspecten aan. Het is soms niet wenselijk of acceptabel vanuit privacy-overwegingen om gegevens die vanuit zeer verschillende doelstellingen zijn verzameld en door werknemers ter beschikking zijn gesteld, te koppelen ten behoeve van onderzoek.

Voor externe beoordelingsmethoden (door experts, supervisoren, of collega's) geldt, dat de kosten hiervan meestal aanzienlijk zijn. Verder zijn ook bij het werken met geobjectiverde beoordelingsmethoden allerlei meetproblemen te verwachten. Zapf (1989) noemt de volgende:

- de neiging om oordelen rond het gemiddelde te geven
- vertekening door de persoonlijkheid van de beoordelaar (bijvoorbeeld meegaandheid)
- beoordeling vanuit een impliciete theorie die de beoordelaar heeft over doel en inhoud van het onderzoek
- beoordeling van een werkplek/arbeidstaak vanuit een eerste indruk bij de beoordelaar (halo-effect)
- hoe eenvoudiger en minder autonoom de arbeidstaak, hoe groter de kans op vertekening in de beoordeling door gebrek aan vergelijkingsmateriaal over een bepaalde werknemer die dit werk verricht

- vertekening in de beoordeling door te korte beoordelingstijd
- vertekening in de beoordeling door gebrek aan observeerbaarheid, bijvoorbeeld bij mentale arbeidstaken
- gebrek aan representativiteit van beoordeelde werkplekken
- vertekening doordat van pure observatie geen sprake is, maar van participerende observatie

We kunnen de problemen van externe beoordelingsmethoden illustreren aan de hand van de methode Welzijn bij de Arbeid (WEBA), een bekende Nederlandse expert-beoordelingsmethode op het gebied van PSA. Door Dhondt (1993) is onderzoek gedaan naar de interbeoordelaarsbetrouwbaarheid van de WEBA. Voor de functiematrix van de WEBA was de gemiddelde overeenstemming tussen beoordelaars en een norm-beoordelaar 57% (Cohen's Kappa). Voor de regelproblemen van de WEBA was dit cijfer 68%, en voor de welzijnscondities (het welzijnsprofiel dat het feitelijke resultaat is van de WEBA) 57%. Dhondt (1993) komt tot de volgende conclusie: *"De voornaamste conclusie is dat de interbeoordelaarsbetrouwbaarheid van de WEBA-methodiek op grond van dit onderzoek laag moet worden genoemd"* (p. 21). Dit resultaat werd verkregen ondanks een verregaande standaardisatie van het materiaal (op video) dat de experts ter beoordeling kregen voorgelegd. Tenslotte bleek dat een grotere mate van ervaring met de WEBA-methodiek niet leidde tot een betere overeenstemming tussen beoordelaars.

De situatie aangaande het meten van PSA en werkstress, met de wijd verbreide kritiek op zelfbeoordelingsmethoden enerzijds, en het blinde vertrouwen in de betere kwaliteit van geobjectiveerde methoden anderzijds, wordt aardig verwoord in het volgende citaat uit Semmer, Zapf en Greif (1995): *"It is somewhat astonishing that ratings by observers or supervisors so often are regarded as "objective" in contrast to the "error-prone" information contained in self-reports. In addition, it is astonishing that the discussion often is quite global, that is, information from incumbents is contrasted with information which is independent of the incumbent in general, without differentiating between types of information"* (p. 5). Discussies over de keuze van instrumenten leiden in de praktijk vaak tot zwart-wit discussies tussen voorstanders van vragenlijstonderzoek versus voorstanders van beoordelingen door deskundigen en/of leidinggevenden. Het is verstandiger om, afhankelijk van het onderzoeksdoel en het type te verzamelen informatie, een keuze te maken voor één van beide of een combinatie van beide. Om deze beslissing te kunnen nemen zou gekeken kunnen worden naar een aantal dimensies (Zapf, 1989): observeerbaarheid, stabiliteit, en afhankelijkheid van de eigen affectieve reacties.

Een verbeterde situatie ontstaat door het combineren van verschillende meetmethoden, bijvoorbeeld self-reports, medisch onderzoek, verzuimregistratie, en beoordelingsinstrumenten.

De situatie verbetert nog verder wanneer deze informatie beschikbaar is over meerdere meetmomenten (Frese en Zapf, 1988; Spector, 1992). In Duitsland ontwikkelt zich een onderzoekstraditie op dit gebied. Frese (1985), Zapf (1989), en Leitner (1993), tonen aan dat een dergelijke multitrait-multimethod-multimoment aanpak fundamentele verbeteringen geeft in de kwaliteit van de meting van PSA. De mogelijkheden om conclusies te trekken over het causale effect van PSA op de gezondheid nemen eveneens toe. Toch lijkt het erop dat, ook in dergelijke studies met vrij ideale onderzoeksinstrumenten en -designs, een limiet bestaat voor de hoeveelheid variantie die door PSA verklaard kan worden in werkstress-verschijnselen. Deze limiet ligt bij 10% à 15% van de variantie (Semmer, Zapf en Greif, 1995). We komen op deze materie terug in hoofdstuk 8, wanneer we ons gaan bezighouden met de relatie tussen PSA en werkstress.

In veel praktijksituaties waar uit pragmatische- en tijdsoverwegingen wordt gekozen voor een self-report methode, is een dergelijke combinatie van meer meetmethoden en meetmomenten niet haalbaar. Dit kost immers dubbel tijd en geld. Vaak wordt derhalve gekozen voor een tweetraps-raket. Eerst wordt een self-report methode toegepast. Hieruit worden dan op allerlei manieren "signalen" gehaald, die vervolgens aanleiding kunnen zijn tot een meting en/of een geobjectiveerde beoordeling. De kwaliteit van dergelijke signaleringsmethoden varieert van simpele vuistregels tot statistische methoden (Broersen, Weel en van Dijk, 1989; Weel en Broersen, 1992).

Ook in het onderzoek waarover in dit proefschrift wordt gerapporteerd hebben we ons moeten schikken naar de beperkingen van de praktijk: het gaat om cross-sectionele vragenlijstgegevens met weinig meer additionele data dan individuele verzuimcijfers. Anderzijds betekent dit wel dat het onderzoek in een praktijkrelevante setting is uitgevoerd, wat de generaliseerbaarheid en bruikbaarheid van de resultaten ten goede komt.

4.1.3 De triviality trap

Sommige auteurs pleiten er sterk voor om bij de meting van PSA uit te gaan van een individu-afhankelijke meting (Lazarus et al., 1985; Warr, 1990b). Indien we dit uitgangspunt accepteren, biedt een self-report methode de enig juiste informatie over dit onderwerp (ongeacht de feitelijke situatie). Andere auteurs (Kasl, 1978, 1986; Christis, 1988) stellen zich negatief tegenover een dergelijke individu-afhankelijke meting op, omdat dit een meer objectieve definitie en vaststelling van risico's in werk en werkomstandigheden onmogelijk zou maken, en een aantal meetproblemen in de hand zou werken.

Kasl (1978) geeft in het volgende citaat de kern van de problematiek van een individu-afhankelijke meting weer: *"Unfortunately, this convergence of theoretical formulations (about the role of individual differences in appraisal) has led to a self-serving methodological trap which has tended to trivialize a good deal of research on work stress or role stress: the measurement of the "independent" variable (e.g., role ambiguity, role conflict, quantitative overload, etc.) and the measurement of the "dependent" variable (work strain, distress, dissatisfaction) are sometimes so close operationally that they appear to be simply two similar measures of a single concept"* (p. 13).

Met andere woorden: de gevonden relaties tussen de onderzochte begrippen komen wellicht meer voort uit het feit dat de begrippen allemaal op dezelfde manier zijn gemeten (namelijk met een vragenlijstmethode, in ongeveer gelijke bewoordingen), dan dat ze iets zeggen over de daadwerkelijke interactie tussen oorzaken en gevolgen van psychosociale arbeidsbelasting. Dit wordt, in navolging van Kasl, wel de "triviality trap" genoemd. Neem als voorbeeld de vraag "is uw werk geestelijk inspannend?", die geacht wordt mentale belasting te bevragen. Neem anderzijds de vraag "heeft u last van spanningsklachten?", die geacht wordt belastingsgevolgen te meten. Een correlatie tussen deze twee vragen ligt voor de hand, alleen al door het woordgebruik.

Een aantal opmerkingen kan naar aanleiding van de kritiek van de triviality trap worden gemaakt.

In de eerste plaats is het noodzakelijk om een vragenlijst zo te construeren dat de daaruit berekende scores betrouwbaar zijn en verwijzen naar eenduidige begrippen. Op deze manier levert een vragenlijst een helder en bruikbaar begrippenkader op. Ook alleen op deze manier is het mogelijk om ooit valideringsonderzoek te doen naar de mate waarin signalen vastgesteld aan de hand van een vragenlijstmethode corresponderen met signalen vastgesteld aan de hand van geobjectiverde meetmethoden. Daarbij moet het uiteraard zo zijn dat ook deze andere methoden het bedoelde begrip betrouwbaar en valide meten!

In de tweede plaats is de genoemde kritiek aanleiding om de doelstelling van een vragenlijstonderzoek over PSA en werkdruk op groepsniveau opnieuw te bezien, en realistische verwachtingen hierover te formuleren. Wellicht ligt de oplossing niet zover verwijderd van de manier waarop op het niveau van de individuele werknemer met deze materie wordt omgegaan. Daar levert de lijst immers niets meer of minder op dan een aantal gespreksonderwerpen tussen werknemer en deskundige. Hierdoor neemt de efficiency van het gesprek toe.

Op een analoge manier levert een groepsgewijs vragenlijstonderzoek een aantal gespreksonderwerpen op tussen werknemers(-vertegenwoordiging), werkgever, en deskundige. Hierdoor kan de efficiency van dit gesprek en het daaruit voortvloeiende beleidsproces toenemen.

Deze onderzoeksstrategie wordt beter naarmate behalve vragenlijstgegevens ook tegelijkertijd andere informatie (verzuimcijfers, verloopcijfers, productiviteitscijfers, medische gegevens) in het onderzoek wordt betrokken. Voor voorbeelden van creatief gebruik van vragenlijsten in een dergelijke onderzoeksopzet zie het boek van Karasek en Theorell (1990).

Een laatste opmerking heeft betrekking op het analyse-niveau bij vragenlijstonderzoek. De self-report methode biedt de meest betrouwbare resultaten wanneer door een goede onderzoeksopzet en groepskeuze het effect van de individuele persoonlijkheid zoveel mogelijk "buiten haakjes" kan worden gehaald. Hiermee wordt het volgende bedoeld:

Wanneer we het oordeel over tijdsdruk van 400 werknemers van ziekenhuis A vergelijken met een op dezelfde wijze gemeten oordeel van 2000 werknemers van andere ziekenhuizen in een regio, hoe aannemelijk is het dan nog dat een significant verschillend oordeel over tijdsdruk tussen deze twee groepen kan worden verklaard door interindividuele verschillen wat betreft assertiviteit, neuroticisme, type A-gedrag, locus of control, of sterfgevallen in de familie? Het is aannemelijk dat de invloed van diverse factoren op het niveau van de individuele medewerker (bijvoorbeeld persoonlijkheid, coping, life events, en verwerkingsvermogen) zich bij een dergelijke vergelijking zullen uitmiddelen. Dit sluit aan bij de stellingname van Semmer (1992), dat er sprake is van een redelijke overeenkomst in de mate waarin bepaalde situaties door mensen als stressvol worden beoordeeld, zolang er een overeenkomst is in de sociaal-culturele context waarin deze mensen zich bevinden.

4.1.4 Onderzoeksdesign en vragenlijsten

Veel van de nadelen van vragenlijsten (vragenlijsten zijn subjectief) zijn - zo hebben we gezien - evenzeer van toepassing op (zogenaamde) geobjectiverde methoden: hierin ligt dus niet het verschil. Een groot deel van de resterende kritiek op de vragenlijstmethode heeft niet zozeer betrekking op de vragenlijstmethode zelf, als wel op het onderzoeksdesign dat meestal in de praktijk bij vragenlijstonderzoek wordt toegepast (cross-sectionele meting zonder additionele gegevens).

Door de cross-sectionele meting is het trekken van conclusies over oorzakelijkheid onmogelijk. Het ontbreken van gegevens uit andere bronnen vergroot de gevoeligheid voor methode-bias. Dit soort beperkingen vloeit voort uit het onderzoeksdesign, niet uit de vragenlijstmethode.

Voor toepassing als signaleringsinstrument is vragenlijstonderzoek één van de weinige methoden die geschikt is om op grote schaal te worden toegepast. Het "geprovoceerde oordeel" van de werknemer zelf is in de praktijk een van de meest

efficiënte bronnen van informatie. Dit geldt met name voor onderzoek bij grote bedrijven en instellingen, bij bedrijfstakken en branches, en bij de Nederlandse beroepsbevolking als geheel.

De keuze voor vragenlijsten wordt in de praktijk meestal gemaakt vanuit overwegingen van kosten versus opbrengsten, in het licht van een specifieke toepassing. Het design van praktijkonderzoek is vaak gebrekkig om dezelfde reden als waarom voor vragenlijsten als instrument wordt gekozen: de toepassing vraagt om informatie over een grote doelgroep tegen relatief geringe kosten. Het is niet fair om achteraf de vragenlijstmethode negatief te evalueren ten gevolge van een gebrekkig onderzoeksdesign. Een design staat los van de gekozen methode voor gegevensverzameling.

Door een vragenlijstonderzoek op systematische wijze te herhalen, waarbij veranderingen in werk en werksituatie worden geregistreerd, en door het gebruik van controlegroepen van werknemers die niet aan deze veranderingen zijn blootgesteld, is wel degelijk te onderzoeken in hoeverre PSA van invloed is op het ontstaansproces van werkstress. Misschien is het zelfs mogelijk om op deze manier met vragenlijstonderzoek oorzakelijke relaties aan te tonen tussen PSA en werkstress. Voor voorbeelden van longitudinale analyses zie Furda (1995) en De Jonge (1995).

4.2 Mogelijkheden en beperkingen van de vragenlijstmethode voor PSA en werkstress

Het kiezen voor een vragenlijstmethode om PSA en werkstress te meten geeft a priori een aantal beperkingen wat betreft de mogelijk te meten begrippen. Om te illustreren waar deze beperkingen liggen, gebruiken we het model van Beehr en Newman uit 1978 (zie paragraaf 2.8).

Een nadere bestudering van de onderwerpen die onder de diverse rubrieken worden geplaatst, en van de inhoudelijke beschrijving die Beehr en Newman van de rubrieken geven, leert dat een aantal rubrieken min of meer buiten de mogelijkheden valt van een vragenlijstonderzoek:

- de objectieve situatie
- biomedische parameters voor de fysieke conditie van een werknemer
- fysiologische en psychologische processen bij de werknemer die zich niet lenen voor zelfobservatie
- gevolgen van stress voor de organisatie, gezien vanuit de werkgever en de hem ter beschikking staande management-informatiesystemen
- aanpassingsreacties van de organisatie, gezien vanuit de werkgever

De invulling van deze elementen zal in werkstressonderzoek moeten worden opgevangen met andere meetmethoden en registraties.

Als grote blokken met betrekking tot PSA en werkstress, die met behulp van een vragenlijstmethode te meten zijn, blijven derhalve over:

- kenmerken van de omgeving volgens het oordeel van de werknemer
- kenmerken van de werknemer die zich lenen voor zelfobservatie
- kenmerken van het verwerkingsproces die zich lenen voor zelfobservatie
- negatieve gevolgen voor de werknemer volgens het oordeel van de werknemer
- negatieve gevolgen voor de organisatie volgens het oordeel van de werknemer
- aanpassingsreacties volgens het oordeel van de werknemer

We gebruiken deze zes-deling in hoofdstuk 5 om te bekijken wat de stand der techniek is op het gebied van instrumenten die PSA en werkstress pretenderen te meten.

4.3 Praktijkeisen

Door Zijlstra en Mulder (1989) en Van Ouwerkerk, Meijman en Mulder (1994) is een aantal criteria aangereikt waaraan een instrument voor de arbo-praktijk dient te voldoen. We bespreken deze criteria hieronder, met enkele aanvullingen.

Technische kwaliteit

Voor een vragenlijstmethode kan de technische kwaliteit in een aantal psychometrische criteria worden uitgedrukt. Veelal komen we in de literatuur alleen de betrouwbaarheid tegen, en dan nog meestal uitgedrukt in één maat: de interne consistentie (Cronbach's Alpha).

Nu is er binnen de psychometrie een verandering gaande van de klassieke testtheorie naar een nieuwe categorie van theorieën, die samen worden gevat onder de noemer van de "item-respons theorie". Voorbeelden van dit soort theorieën zijn die van Rasch, Birnbaum, en Mokken (Drenth en Sijtsma, 1990).

In de item-respons theorie is het doel om na te gaan welke relatie een reeks van antwoorden op vragenlijstitems vertoont met een veronderstelde, achter deze items liggende, gezamenlijke dimensie (latente trek). Voor de aard van deze relatie zijn per model wiskundige voorwaarden geformuleerd. Als de relatie tussen antwoorden op de items en de latente trek voldoet aan de wiskundige voorwaarden behorend bij een bepaald model (bijvoorbeeld Mokken, Birnbaum, Rasch), mogen we hieruit bepaalde eigenschappen van de vragenlijstschalen afleiden die door deze items wordt gevormd. Betrouwbaarheid (samenhang van de items), unidimensionaliteit (eenduidigheid van het gemeten begrip), en meetniveau (nominaal-ordinaal-interval-ratio) zijn voor de hier te ontwikkelen vragenlijst de belangrijkste van dergelijke eigenschappen.

Voor de praktijk van advisering over PSA en werkstress is het van belang dat de met vragenlijsten gemeten begrippen betrouwbaar en eenduidig zijn, en minimaal een ordinaal meetniveau hebben. Dit laatste wil zeggen dat we in staat moeten zijn om werknemers te ordenen van laag tot hoog op het met de vragenlijst gemeten begrip.

Een volgende technische eigenschap van een vragenlijst, wil die voldoen aan praktischeisen, is de validiteit. Dit houdt in dat de meting met deze vragenlijst samenhangt met andere (gerenommeerde) vragenlijsten of fenomenen, op een manier die klopt met theorievorming en eerder onderzoek in het vakgebied. Dit kan zijn in de zin dat veronderstelde samenhangen worden gevonden (convergente validiteit). Het kan ook zijn dat de veronderstelde afwezigheid van samenhang wordt bevestigd (divergente validiteit).

Een laatste en waarschijnlijk belangrijkste vorm van validiteit voor de praktijk is die van criterium-validiteit. Hiermee wordt bijvoorbeeld bedoeld dat de vragenlijstscore iets zegt over te verwachten uitkomsten, met andere woorden: dat een relevante gebeurtenis voorspeld kan worden op basis van de vragenlijstscore (predictieve validiteit). Voor de arbo-praktijk is het niet zo moeilijk om aan te geven wat men zou willen voorspellen met een vragenlijstscore, namelijk: ziekteverzuim, arbeidsongeschiktheid, verloop, lage productiviteit et cetera. Een ander voorbeeld van criteriumvaliditeit is een samenhang van een vragenlijstscore met een gelijktijdig gemeten criterium (concurrente validiteit). Soms kan dit worden gedaan in termen van samenhangen tussen onderdelen in een vragenlijst, bijvoorbeeld tussen veronderstelde determinanten en veronderstelde gevolgen.

Behalve de hier besproken vormen van betrouwbaarheid en validiteit zijn er nog enkele technische eigenschappen waaraan een goed instrument dient te voldoen. We bespreken deze hier kort. In de eerste plaats de sensitiviteit. Dit geeft weer hoe gevoelig het instrument is voor veranderingen in datgene wat ermee wordt gemeten. In de tweede plaats de stabiliteit. Dit kenmerk van een meetinstrument hangt samen met de mate waarin datgene wat wordt gemeten stabiel is in de tijd. Ten derde de diagnosticiteit. Dit heeft betrekking op de vraag hoe makkelijk het is om een meetresultaat te vertalen in een of andere vorm van vervolgactie. Ten vierde de selectiviteit. Hiermee wordt bedoeld dat een instrument gevoelig is voor veranderingen in een bepaalde parameter, en niet voor veranderingen in andere parameters.

Strategische waarde

Een volgend criterium voor de arbo-praktijk is de strategische waarde. In dit werkveld is het van groot belang om aan te sluiten bij de planmatige manier van werken die bedrijven, instellingen en koepelorganisaties volgen om tot verbeteringen te komen van de kwaliteit van de arbeid, de arbeidsorganisatie en het arbeidsresultaat. Als een instrument hier niet bij aansluit, kan het technisch nog zo goed zijn, men zal het dan niet gebruiken.

Van bijzonder belang op dit moment is dat een vragenlijstinstrument aansluit bij twee kaders die vanuit de nieuwe Arbowet worden aangereikt: de risico-inventarisatie en -evaluatie (RI&E) en het periodiek arbeidsgezondheidkundig onderzoek (PAGO). Dit zijn twee hoofdlijnen waarlangs beleid en dienstverlening op het terrein van arbeid en gezondheid in de komende jaren zullen worden ingevuld richting bedrijven en instellingen.

Acceptatie en gebruiksgemak

Met name voor een vragenlijstmethode die door grote groepen werknemers zal worden ingevuld, is het van belang dat dit instrument door werknemers wordt geaccepteerd en als gemakkelijk te gebruiken wordt beoordeeld. Als aan deze voorwaarde niet wordt voldaan zakt de deelname bij onderzoek, wat de representativiteit van de onderzoeksresultaten ondergraaft. Ook kan een laag gebruiksgemak leiden tot een slechte kwaliteit van invullen wat eveneens leidt tot verlies van bruikbare onderzoeksgegevens.

Behalve voor de werknemers, moet het instrument ook voor opdrachtgevers en ondersteunende diensten acceptabel zijn en makkelijk te gebruiken. Dit stelt onder andere eisen aan de presentatie van handleiding en normen, en aan het gemak van scoring, verwerking, en interpretatie van de vragenlijst.

Een laatste aspect dat in deze categorie valt is de mate waarin de toepassing van een instrument de normale taakuitoefening verstoort. Aangezien vragenlijsten vaak buiten de werksituatie worden ingevuld is dit punt van minder belang.

Kosten/baten

Een laatste criterium is dat van kosten versus opbrengsten. Dit is voornamelijk afhankelijk van de toepassing (RI&E, PAGO, werkbelevingsonderzoek, ziekteverzuimonderzoek) en de doelgroep (omvang, complexiteit, gemiddeld opleidingsniveau, soort werk).

De ervaring leert dat de kosten/baten-verhouding van een vragenlijstsonderzoek op het gebied van PSA en werkstress toeneemt met de omvang van de populatie. In kleinere bedrijven vergt een vragenlijstsonderzoek vaak meer voorbereiding, introductie, en organisatie dan het bedrijf kan opbrengen. Tevens is de meerwaarde van het onderzoek kleiner, waar het gaat om het localiseren van problemen binnen de organisatie. In kleinere bedrijven heeft het management hier toch vaak een redelijk overzicht van. In grote organisaties mag men echter nauwelijks verwachten dat een dergelijk overzicht tot op het niveau van afdelingen bestaat bij het topmanagement, zonder een gestructureerde vorm van management-informatie, bijvoorbeeld middels een vragenlijstsonderzoek.

De aspecten die in deze paragraaf zijn genoemd met betrekking tot de item-respons theorie en de validiteit, komen in dit proefschrift uitgebreid aan bod. De overige aspecten worden hier niet verder onderzocht, maar hebben deels wel meegespeeld bij de vragenlijstconstructie.

4.4 Slot deel I

In het eerste deel van dit proefschrift hebben we aandacht besteed aan theorie en literatuur die relevant is voor het meten van PSA en werkstress met vragenlijsten. Deze theorie en literatuur vormt de basis voor de beantwoording van de vragen die in het eerste hoofdstuk zijn gesteld. We herhalen hier nogmaals deze vragen, maar koppelen dit nu met de besproken theorie en literatuur.

Wat zijn de te onderscheiden dimensies bij onderzoek naar PSA en werkstress met vragenlijsten? We zijn met name geïnteresseerd in de operationalisatie van het begrip decision latitude zoals dat door Karasek (1979) is gepostuleerd. Hebben we aparte schalen nodig voor skill discretion en task authority, of kunnen we met een globale schaal voor decision latitude volstaan?

In hoeverre wordt het onderzoek naar relaties tussen PSA en werkstress met zelfbeoordelingsmethoden bemoeilijkt door verstorende variabelen, met name door negatieve affectiviteit (Watson en Pennebaker, 1989)? Blijft er na correctie voor NA nog een eigen effect van PSA op werkstress over? Is de verstorende invloed van NA voor alle typen werkstress-uitkomsten gelijk of zijn er verschillen?

In hoeverre spelen sociale vergelijkingsprocessen een rol bij onderzoek naar PSA en werkstress? We zijn hierbij vooral geïnteresseerd in de rol van ervaren onbillijkheid (Adams, 1965). Treedt EO op als confounder, vergelijkbaar met het effect van NA?

Als we proberen om individuele symptomen van werkstress te verklaren op basis van PSA, hoeveel verklaring ligt dan op het niveau van de individuele werknemer, hoeveel op het niveau van de afdeling of de functie, en hoeveel op het niveau van de organisatie? Dit deel van het proefschrift wordt een verkenning naar de mogelijkheden van multi-level analyse voor onderzoek naar PSA en werkstress (Bosker en Snijders, 1990; Longford, 1993). Lopen we bij het verklaren van werkstress-verschijnselen op basis van PSA inderdaad aan tegen de limiet van 10% á 15% verklaarde variantie? (Semmer, Zapf en Greif, 1995).

Verder onderzoeken we de relatie tussen PSA en werkstress zelf. We vragen ons hierbij af of er sprake is van één algemeen stress-mechanisme dan wel van een aantal parallel opererende mechanismen.

Deze vraagstelling wordt beantwoord door twee theorieën tegenover elkaar te plaatsen die prototypisch zijn voor de genoemde mogelijkheden. Dit zijn de theorieën van Karasek (1979) en van Warr (1987). Waar we voor de Karasek-benadering verwachten dat PSA op diverse typen van werkstress-uitkomsten gelijksoortige effecten zal hebben, voorspelt de theorie van Warr dat verschillende werkstress-uitkomsten verschillende PSA-predictoren hebben. Ook wordt een analyse uitgevoerd om te bekijken in hoeverre de relatie tussen PSA en werkstress lineair of non-lineair is.

Tenslotte onderzoeken we de relatie van PSA en werkstress met ziekteverzuim. Tegelijkertijd onderzoeken we of EO inderdaad een goede voorspeller is van verzuim (Geurts, 1994).

De kern van dit proefschrift is het aantonen van de waarde van een zelfbeoordelingsmethode voor het meten van PSA en werkstress. In deze paragraaf is een aantal methodologische en theoretische vragen gesteld die hiervoor van belang zijn. Om aan de beantwoording van deze vragen toe te komen hebben we een goede vragenlijst nodig. Deel II van dit proefschrift gaat over de ontwikkeling van een reeks vragenlijstschalen over PSA en werkstress, volgens technische criteria gebaseerd op de item-respons theorie, meer in het bijzonder op het item-respons model van Mokken (1971).

Deel II: Ontwikkeling van een vragenlijst voor het meten van PSA en werkstress

Hoofdstuk 5: Achtergrond en ontwerp

5.1 Achtergrond

In de afgelopen tijd is een aantal interessante overzichten verschenen van instrumenten op het gebied van PSA en werkstress (Kompier en Marcelissen, 1990; Roe en Zijlstra, 1991; Meulenbeld en Van Lingen, 1993; Schabracq en Winnubst, 1993; Van Ouwerkerk, Meijman en Mulder, 1994; Evers, 1995). Hieruit wordt enerzijds duidelijk dat al het nodige beschikbaar is voor het Nederlands taalgebied. Anderzijds blijkt hoe weinig op het gebied van PSA en werkstress nog is ontwikkeld voor toepassing door arbodiensten, en hoe weinig consensus er is over wat en hoe er gemeten moet worden.

De kaders waarbinnen arbodiensten gebruik maken van instrumenten voor het meten van PSA en werkstress zijn sterk aan verandering onderhevig. Was het Periodiek Bedrijfsgezondheidkundig Onderzoek (PBGO) een aantal jaren geleden nog nagenoeg de enige plaats waar dit soort factoren onderzoeksmatige aandacht kreeg, momenteel is er een belangrijke vraag naar integratie van onderzoeksmethoden op het gebied van PSA en werkstress in instrumenten voor Risico-Inventarisatie en -Evaluatie (RI&E), Periodiek ArbeidsGezondheidkundig Onderzoek (PAGO), en diagnostiek bij individuele Sociaal-Medische Begeleiding (SMB).

Alhoewel niet de enige oplossing en niet in alle gevallen de beste, kan vragenlijst-onderzoek in veel van de hierboven genoemde situaties een nuttige rol vervullen. Tegen deze achtergrond is in 1992 gestart met de ontwikkeling van een vragenlijst voor het meten van PSA en werkstress in de setting van een arbodienst. Deze vragenlijst werd opgeleverd in 1994, en kreeg de naam "Vragenlijst Beleving en Beoordeling van de Arbeid" (VBBA) mee (Van Veldhoven en Meijman, 1994). In 1995 werd de lijst op enkele kleine punten bijgesteld. Voor een exemplaar van de VBBA: zie bijlage 3 in Van Veldhoven en Meijman (1994).

Voor gedetailleerde informatie over de ontwikkeling van de VBBA verwijzen we de lezer naar een drietal rapporten (Van Veldhoven, 1993a; 1993b; Weel en Broersen, 1995) en naar een boek (Van Veldhoven en Meijman, 1994).

In het kader van dit proefschrift is het niet zozeer van belang om de ontwikkeling van de VBBA volledig te beschrijven, als wel om dieper in te gaan op de wijze waarop het meten van PSA en werkstress met behulp van een zelfbeoordelingsmethode is geoperationaliseerd.

In deel II van dit proefschrift concentreren we ons derhalve op de constructie van een verzameling van vragenlijstschalen voor het meten van PSA en werkstress, voor zover dit van belang is voor het beantwoorden van de in deel I geformuleerde vraagstellingen. Hoofdstuk 5 gaat in op de achtergrond en het ontwerp van de vragenlijst. Hoofdstuk 6 behandelt het schaalconstructie-onderzoek.

5.2 Facetanalyse

In deel I van dit proefschrift hebben we gezien dat verschillen in theoretische uitgangspunten hebben geleid tot een grote diversiteit aan instrumenten op het terrein van PSA en werkstress.

Veel instrumenten kiezen om principiële redenen voor een beperkte operationalisatie. Als voorbeelden van extremen kunnen genoemd worden: een vragenlijstmethode die uitsluitend is gericht op het signaleren van stress-verschijnselen bij individuen (Vragenlijst Onderzoek Ervaren Gezondheid (VOEG): Dirken, 1967) en een beoordelingsmethode die uitsluitend is gericht op het onderzoeken van condities binnen organisaties, die het ontstaan van werkstress bij werknemers kunnen bevorderen (methode Welzijn bij de Arbeid (WEBA): Pot et al., 1989).

De meeste instrumenten zijn mengvormen van deze twee extremen. Deze diversiteit aan instrumenten weerspiegelt de problemen die het werkstress-onderzoek in zijn algemeenheid kent, namelijk het ontbreken van algemeen aanvaarde definities van sleutelbegrippen en theoretische kaders (Ekkers en Sanders, 1987).

Door de grote diversiteit aan theorieën en instrumenten, is het niet a priori duidelijk welke onderwerpen een nieuw te ontwikkelen vragenlijst over PSA en werkstress zou moeten omvatten. Er bestaat in eerste instantie behoefte aan een overzicht van te meten begrippen. Om de stand der techniek van instrumenten op het gebied van PSA en werkstress te inventariseren is een aantal bestaande instrumenten gescreend op de daarin onderscheiden onderwerpen. We noemen dit een facetanalyse in navolging van de manier waarop Beehr en Newman in 1978 kwamen tot hun model (zie 2.8).

5.2.1 Opzet en werkwijze van de facetanalyse

Bij het verzamelen van instrumenten voor de facetanalyse is gewerkt volgens de sneeuwbalmethode. Dit wil zeggen dat via de referenties van het ene instrument, andere instrumenten zijn opgespoord. Hierbij is zowel binnen als buiten het Nederlands taalgebied gezocht. Er is niet gewerkt met een systematische, laat staan uitputtende screening van de literatuur. Gestart is vanuit een reeks bekende instrumenten: de Job Diagnostic Survey (JDS; Hackman en Oldham, 1974), de Vragenlijst Periodiek Functioneren in de Arbeidssituatie (VPFA; Dijkstra et al.,

1981), de Job Content Instrument (JCI; Karasek, 1985), en de Vragenlijst Organisatie-Stress Doetinchem (VOS-D; Bergers, Marcelissen, en De Wolff, 1986). De facetanalyse heeft ten doel om de in deze instrumenten gekozen onderwerpen in bredere kring te toetsen en zonodig aan te vullen.

De hoeveelheid instrumenten die op deze manier werd verzameld om voor de facetanalyse in aanmerking te komen, was aanzienlijk. Het was dan ook noodzakelijk om een aantal selectiecriteria te formuleren waaraan instrumenten moesten voldoen om voor opname in de facetanalyse in aanmerking te komen. Ook werd vooraf nog een aantal andere keuzes ten behoeve van de facetanalyse gemaakt.

1. Het instrument moest specifiek zijn ontwikkeld ten behoeve van onderzoek naar de beleving van de arbeid, de arbeidssituatie, of effecten daarvan; dus geen instrumenten die primair zijn ontwikkeld voor toepassing in de algemene gezondheidszorg, maar die incidenteel in een bedrijfssetting worden gebruikt.
2. Het instrument moest ontwikkeld zijn voor toepassing bij uiteenlopende beroepen, afdelingen, en bedrijven; dus geen instrumenten die specifiek zijn ontworpen voor een bepaalde doelgroep en/of situatie.
3. Besloten werd om niet meer dan 50 instrumenten in de facetanalyse te betrekken. Dit enerzijds omdat het onmogelijk was om bij deze exercitie uitputtend te werk te gaan, anderzijds omdat werd aangenomen dat met deze 50 instrumenten de belangrijkste oorspronkelijke instrumenten in de exercitie betrokken zijn.
4. Instrumenten mochten eventueel breder zijn dan alleen gericht op PSA en werkstress.
5. Het instrument mocht berusten op een andere methode dan een vragenlijstmethode. Het mocht bijvoorbeeld ook een interviewmethode, beoordelingsmethode, checklist, of rating-scale zijn.
6. Alhoewel er een voorkeur bestond voor instrumenten die diverse facetten meten, werden ook instrumenten die specifiek op een bepaald facet zijn gericht meegenomen.
7. Zowel Nederlandse als buitenlandse instrumenten werden geïnventariseerd. Voor het Nederlands taalgebied is een wat grotere volledigheid van de facetanalyse nagestreefd, dan voor het buitenlandse taalgebied.

Van Veldhoven en Meijman (1994) hebben een alfabetische lijst opgenomen van instrumenten die in de facetanalyse zijn betrokken, inclusief korte literatuurverwijzingen (zie bijlage 1 van die publicatie).

Deze lijst is gesplitst in een Nederlands en een buitenlands gedeelte. De buitenlandse instrumenten zijn vooral afkomstig uit de VS, Groot-Brittannië, Scandinavië, en Duitsland). De facetanalyse is uitgevoerd eind 1992.

Bij het vastleggen van de facetten is zoveel mogelijk gebruik gemaakt van de benamingen van schalen of dimensies die voorkomen in de geselecteerde instrumenten. Er is bij het merendeel van de instrumenten derhalve niet op item-niveau geïnventariseerd wat de onderwerpen waren. Op deze manier werd gebruik gemaakt van de clustering die de diverse instrumentontwerpers zelf reeds aan hebben gebracht in hun instrumenten. Dit zou het indikken van de facetanalyse achteraf moeten vergemakkelijken. Het is belangrijk om op te merken dat de geselecteerde instrumenten niet onafhankelijk van elkaar zijn. Er zitten gedeeltelijke doublures in het geselecteerde materiaal.

5.2.2 Resultaat facetanalyse

De facetanalyse resulteerde in een lijst van ongeveer 100 onderwerpen, die in de geselecteerde instrumenten aan de orde komen. Vervolgens is een eerste poging ondernomen om de onderwerpen uit de facetanalyse in te dikken. Dit heeft geleid tot het verder combineren van onderwerpen tot één facet. Daarbij werd ernaar gestreefd om alleen facetten te benoemen die in minstens vijf instrumenten aan de orde werden gesteld. Het inbouwen van een dergelijke drempel was een voorwaarde om tot enige vorm van indikking te komen.

Er is getracht om de facetten te ordenen onder de zes kopjes die we eerder hebben afgeleid van het model van Beehr en Newman als mogelijkheden voor een meting met behulp van een vragenlijst. Het resultaat van deze ordening is opgenomen in tabel 5.1. Per facet wordt hierin aangegeven in hoeveel van de 50 instrumenten het facet is aangetroffen.

Wanneer we de gevonden facetten afzetten tegen onze verwachtingen op grond van het model van Beehr en Newman (1978), dan zien we het volgende: maar liefst 29 facetten blijken onder de omgevingskenmerken te vallen. We kunnen deze in twee groepen verdelen: beoordeling van taakkenmerken (15 facetten) en beoordeling van de context van de taak (14 facetten). De eerste groep heeft betrekking op de rechtstreekse interactie tussen persoon en arbeid. De tweede groep heeft betrekking op de interactie tussen de persoon en de groep c.q. de organisatie waarbinnen de werkzaamheden worden uitgeoefend. Dit bevestigt het accent op situatiekenmerken dat we eerder constateerden in paragraaf 3.1.7.

Tabel 5.1: *Lijst van onderwerpen uit de facetanalyse*

1. Beoordeling van de omgeving	Aantal instrumenten
A: Taak	
- autonomie qua tempo, volgorde en methode van werken	24
- kortcyclisch werk/weinig afwisseling in het werk	23
- werkdruk/tijdsdruk/variabiliteit hiervan	22
- kwalitatieve over- en onderbelasting/complexiteit	20
- perceptief-mentale belasting	17
- informatievoorziening/feedback taakrelevante informatie	17
- functievolledigheid/taakuitgebreidheid/compleet gebruik vaardigheden	16
- fysieke belasting	15
- verantwoordelijkheid voor mensen/middelen/beleid	14
- kwantitatieve over- en onderbelasting/overwerk	14
- (technische) storingen arbeidsproces/regelmogelijkheden hierin	11
- plaatsgebondenheid/mobiliteit	9
- emotionele belasting/aangrijpende situaties/cliëntcontacten	8
- werken met (nieuwe) technologie	7
- zinvolheid/relevantie van de taak/taakidentiteit	7
B: Context van de taak	
- arbeidsverhoudingen/sociale ondersteuning collega's/ discriminatie/ongewenste intimiteiten/sociale controle	27
- arbeidsverhoudingen/sociale ondersteuning leiding/ stijl van leidinggeven	27
- onveilige/onplezierige fysisch-chemische arbeidsomstandigheden	22
- functionele contacten/sociale contactmogelijkheden tijdens en naast het werk/sociaal isolement	18
- sociale beslissingsruimte/inspraak/participatie/werkoverleg	18
- leer- en groeimogelijkheden/loopbaanproblematiek	18
- onzekerheid/negatieve verwachtingen baan, positie en toekomst bedrijf	17
- waardering/beloning	14
- samenwerking	10
- arbeidsvoorwaarden/regeling werk- en rusttijden/dienstroosters/ verlof en ADV-regeling/werken met uitzendkrachten/vervanging bij ziekte/vacaturebeleid/voorzieningen bedrijf	10
- rolambiguititeit/rolconflict	9
- communicatie/organisatieklimaat en -cultuur/bureaucratie	9
- werkorganisatie	9
- relatie werk/privé	5

Tabel 5.1: Lijst van onderwerpen uit de facetanalyse (vervolg)

2. Zelfobservatie persoon van de werknemer	
geen	
3. Zelfobservatie verwerkingsproces	
- vermoeidheidsgevoelens/uitputting/burnout/herstelproblemen	10
- emotionele reacties (angst, depressie, agressie, irritatie)	9
- gebruik/ misbruik alcohol, nicotine, cafeïne, drugs	6
- (mentale) belastingsgevoelens/ concentratieverlies	6
- inspanningsverschijnselen	5
4. Beoordeling negatieve gevolgen voor werknemer	
- psychische/psychosomatische/psychiatrische klachten	18
- ontevredenheid met arbeid en werkgever	16
- gezondheidsklachten	12
- ontevredenheid met prestatie/frustratie/negatief zelfbeeld/geringe ervaren competentie/piekeren	8
- storing van lichaamsfuncties (slapen, eten, sex)	8
- medische consumptie	7
- ontevredenheid met mogelijkheden voor groei en ontwikkeling/ verlies interne motivatie/verlies betrokkenheid/aspiraties/ intrapersoonlijk conflict	7
- minder plezier in het werk/verveling/taakweerstand/gevoelens	7
- verminderde kwaliteit privé-leven/vrije tijd	5
5. Inschatting negatieve gevolgen voor organisatie	
- verzuim/verloop	6
- prestatievermindering/fouten/traagheid/verlies produktiviteit/ verminderd interpersoonlijk functioneren	6
6. Beoordeling aanpassingsreacties	
geen	

Verder valt op dat er op twee gebieden géén facetten werden gevonden die voldeden aan de drempel van het genoemd worden in minstens vijf instrumenten, namelijk de gebieden "persoon" en "aanpassingsreacties".

Ondanks de grote interesse in de literatuur voor onderwerpen als individuele verschillen en coping komen deze onderwerpen in de facetanalyse niet als hoofdlijn naar voren.

Dit is ongetwijfeld voor een deel te wijten aan de werkwijze bij de facetanalyse. De keuze van de instrumenten kwam immers tot stand via een sneeuwbalmethode, en niet via een uitputtende inventarisatie. Ook het hanteren van het selectie criterium "specifiek ontwikkeld voor de arbeidssituatie" resulteerde erin dat een aantal instrumenten op deze terreinen is afgefallen.

Misschien gaat het bij "persoonskenmerken in relatie tot de gezondheid" en "aanpassingsreacties naar aanleiding van stressvolle gebeurtenissen" om onderwerpen die toch meer in een ander onderzoeksgebied thuishoren, met name in de gezondheidspsychologie (Taylor, 1986; Maes en Van Veldhoven, 1990). Een andere mogelijkheid is, dat uit de onderzoekstraditie op het gebied van PSA en werkstress nog geen echte consensus naar voren is gekomen over de wenselijkheid en wijze van meting van persoonskenmerken en coping. Een grootste gemene deler ontbreekt nog (Cooper en Payne, 1991).

Verder worden op het gebied van "negatieve consequenties voor de organisatie" slechts twee facetten gevonden die in vijf of meer instrumenten voorkomen. Dit onderwerp krijgt derhalve geen accent in de instrumentliteratuur op het gebied van PSA en werkstress.

Een volgende opmerking bij het indikken tot hoofdfacetten heeft betrekking op de plaats van het begrip "arbeidssatisfactie". Er is voor gekozen om de term satisfactie te interpreteren als "algemene tevredenheid met...". Dit in tegenstelling tot de zogenaamde facet-satisfactie (zie bijvoorbeeld Biessen, 1992). Op deze manier denken we dat het mogelijk is om satisfactie te plaatsen onder het kopje "mogelijke gevolgen van psychosociale arbeidsbelasting". Hierdoor wordt de conceptuele overlap beperkt tussen facetten die betrekking hebben op bronnen van PSA in het werk en in de werkomgeving enerzijds, en facetten die betrekking hebben op mogelijke werkstressgevolgen van PSA anderzijds. Dit is één van de manieren waarop de in paragraaf 4.1.3 beschreven "triviality trap" deels kan worden vermeden.

Tenslotte bleek bij het ordenen van de facetten, dat het soms moeilijk is een streep te trekken tussen facetten die duiden op het verwerkingsproces (bijvoorbeeld vermoeidheid en emoties), en facetten die duiden op negatieve gevolgen voor de werknemer (bijvoorbeeld gezondheidsklachten en psychische klachten). Dit onderscheid lijkt meer een splitsing te zijn tussen mogelijke gevolgen van psychosociale arbeidsbelasting op kortere en langere termijn. Dit is geheel in lijn met modellen uit de belasting-belastbaarheid benadering (Meijman, 1989; Van Dijk et al., 1990; Kompier, 1991).

In paragraaf 2.9 en 2.10 spraken we de verwachting uit een grootste gemene deler te kunnen vinden in de begrippen, die volgens verschillende theorieën gemeten dienen te worden met betrekking tot PSA en werkstress.

Kijkend naar de frequenties waarmee begrippen in de 50 instrumenten voorkomen, dan zien we dat maar liefst veertien begrippen in minstens één op de drie instrumenten aan de orde komen.

Uit de facetanalyse is verder de conclusie te trekken dat het merendeel van de instrumenten betrekking heeft op oorzakelijke taak- en omgevingsaspecten, negatieve gevolgen hiervan voor werknemers, ofwel een combinatie van deze twee. Bij de oorzakelijke omgevingsaspecten dient te worden opgemerkt, dat in de gescreende instrumenten enerzijds belastende factoren worden onderscheiden, maar dat anderzijds ook het belang van regelmogelijkheden wordt benadrukt. Bij de gevolgen kunnen we in de facetanalyse een onderscheid herkennen naar gevolgen op kortere (verwerkingsproces) en langere termijn (gezondheidsklachten, ontevredenheid). Het is niet zo moeilijk om in de resultaten van de facetanalyse de herziene modellen belasting-belastbaarheid te herkennen (Meijman, 1989; Van Dijk et al., 1990; Kompier, 1991: zie 2.7). Wat in de facetanalyse niet wordt teruggevonden is het verwerkingsvermogen, zoals de herziene modellen belasting-belastbaarheid dit poneren. Hier ligt een parallel met het eerder geconstateerde gebrek aan consensus in de instrumentliteratuur over het meten van persoonskenmerken en aanpassingsreacties (coping).

Uit de facetanalyse wordt duidelijk dat het conceptuele model van Beehr en Newman (1978) en de stand der techniek op het gebied van onderzoeksinstrumenten nogal van elkaar verschillen. Slechts een klein gedeelte van het model van Beehr en Newman is ingevuld op een redelijk algemeen aanvaarde manier.

De open plekken in het model van Beehr en Newman (1978) hebben betrekking op het proces dat taak- en omgevingsaspecten en mogelijke gevolgen hiervan met elkaar verbindt, de rol van de factor tijd bij het ontstaan van de gevolgen, de dynamiek die ontstaat doordat zowel werknemer als organisatie zich proberen aan te passen aan de gevolgen, de mate waarin kenmerken van de persoon de relatie tussen oorzakelijke factoren en gevolgen modereren, en de mate waarin ook de organisatie met negatieve gevolgen te maken krijgt. De herziene modellen belasting-belastbaarheid sluiten beter aan bij de resultaten van de facetanalyse dan het model van Beehr en Newman (1978). Met uitzondering van het verwerkingsvermogen komen alle elementen uit deze modellen in de facetanalyse naar voren.

Bij de ontwikkeling van een vragenlijst voor het meten van PSA en werkstress wordt uitgegaan van een facetanalyse voor het bepalen van de inhoud. Dit in tegenstelling tot een a priori-keuze voor een bepaalde theorie, die voorschrijft welke concepten gemeten moeten worden. Het bovenstaande geeft echter aan, dat we ons bij de vragenlijstconstructie toch theoretisch gesteund kunnen voelen door de herziene modellen belasting-belastbaarheid.

5.3 Constructie van de itempool

De facetanalyse zoals hierboven besproken is als uitgangspunt genomen voor de vragenlijstconstructie. Bij de gevonden facetten werden items uit bestaande vragenlijstinstrumenten gezocht, die deze begrippen meten. Waar die niet te vinden waren werden nieuwe items geformuleerd.

Heel expliciet is ervoor gekozen om bij de "mogelijke gevolgen" geen schalen te ontwikkelen voor het meten van (psychische en/of psychosomatische) gezondheidsklachten, en hiermee verband houdende medische consumptie. Hiervoor bestaan binnen en buiten het arbo-veld reeds goede alternatieven. Voor een uitgebreid overzicht van schalen voor het meten van de gezondheidstoestand zie de driedelige reeks van König-Zahn, Furer en Tax (1993-1995) hierover.

In Van Veldhoven en Meijman (1994) wordt een tabel gepresenteerd met een overzicht van welke vragenlijsten als bronnen zijn gebruikt voor het formuleren van items in de te construeren schalen (zie bijlage 2/tabel 1 in die publicatie). We bespreken hier alleen die facetten, die problemen opleverden bij het komen tot een itempool.

Een eerste probleem bij het kiezen van items voor de itempool was de relatie tussen werk en privé-leven. Dit facet komt zowel voor onder "context van de taak" (de invloed van het privé-leven op de taakuitoefening) als onder "mogelijke gevolgen" (de invloed van het werk op het privé-leven). Besloten is om dit facet uitsluitend aan de gevolgenkant te plaatsen. Op deze manier wordt de "context van de taak" beperkt tot de groep en de organisatie waarbinnen het werk wordt uitgeoefend. Hierop dient zich naar onze mening de aandacht te concentreren wanneer bij een onderzoek speciale aandacht wordt besteed aan het gebied PSA en werkstress. Hierbij vinden we steun in de conclusie van Gutek et al. (1988): *"In fact, there seems to be some consensus that work probably impacts family life more than family does work"* (p. 163).

Het lijkt inderdaad aannemelijker dat zaken als ploegendiensten en overwerk effect hebben op de privé-situatie van groepen werknemers, dan dat organisaties verschillen in de mate waarin ze groepen werknemers aantrekken waarbij het werk te lijden heeft van privé-omstandigheden.

Een tweede probleem bij het kiezen van bestaande schalen, deed zich voor bij de facetten die betrekking hebben op "mogelijke gevolgen" op het fysieke en gedragsmatige vlak. Deze zijn voor een deel goed te meten met een aantal specifieke schalen, die in het onderzoek van het Studiecentrum Arbeid en Gezondheid (UVA) en de Sectie Experimentele en Arbeidspsychologie (RUG) veelvuldig worden gebruikt, en ook in deze onderzoeksgroepen zijn ontwikkeld. Het gaat hier om vermoeidheid/inspanningsverschijnselen/mentale belastingsgevoelens (SEB; Meijman, 1991) en slaapklachten (GSKS, Meijman et al., 1985). Besloten is om de SEB en GSKS niet op te nemen in de itempool.

Ten derde is "het werken met nieuwe technologie/werkwijzen en de introductie hiervan" een probleem: hiervoor zijn geen goede items gevonden in de bestaande literatuur. Hiervoor is een nieuwe reeks items geformuleerd. Hetzelfde geldt voor het aanverwante onderwerp "organisatieverandering".

Ten aanzien van verzuim en verloop tenslotte, is duidelijk dat deze parameters beter aan objectieve personeelsinformatiesystemen kunnen worden ontleend dan aan schattingen en herinneringen van werknemers. Wellicht is het begrip "verloopintentie" in dit verband nog het meest meetbaar met behulp van een vragenlijst. Hiermee wordt bedoeld de mate waarin de werknemer speelt met de gedachte c.q. probeert om de organisatie te verlaten. Het is de tegenhanger van "binding" aan de organisatie. Voor het meten van productiviteit en prestatie is het eveneens moeilijk geschikte items te vinden. Ook hier geldt dat er vaak betere metingen mogelijk zijn aan de hand van registratiesystemen op het bedrijf dan met behulp van vragenlijsten. Er is dan ook besloten de facetten met betrekking tot negatieve gevolgen voor de organisatie niet in te vullen binnen de itempool, met uitzondering van "verloopintentie".

5.4 Het bepalen van het type antwoordmogelijkheden

Wat betreft het type antwoordmogelijkheden bij vragenlijstonderzoek op het gebied van werkstress, adviseert DeFrank (1988) om per onderzocht facet gebruik te maken van drie typen antwoordmogelijkheden: frequentie, duur, en intensiteit. In een onderzoek naar de meerwaarde van deze drie typen antwoordmogelijkheden ten opzichte van de vaak gehanteerde ja/nee-keuze bij het meten van werkstressoren, concludeert Dewe (1991) het volgende: het hanteren van deze drie typen antwoordmogelijkheden naast elkaar, verklaart inderdaad een significante proportie variantie bovenop de verklaarde variantie door een ja/nee-keuze.

Het opnemen van drie vragen per onderwerp (frequentie, duur, intensiteit) is echter gezien de hoeveelheid facetten/onderwerpen die op dit moment voor de itempool in aanmerking komen onmogelijk. Voor het meten van de blokken "taak" en "context van de taak" is derhalve gekozen voor een globale frequentiebeoordeling in termen van nooit, soms, vaak, altijd.

Dit biedt in ieder geval vooruitgang ten opzichte van een simpele ja/nee-keuze. Het biedt de respondent enige ruimte tot nuancering. Door antwoordmogelijkheden te gebruiken die betrekking hebben op de frequentie waarmee zich iets voordoet hopen we daarnaast meer objectiveerbare antwoorden te krijgen dan wanneer we zouden kiezen voor constructies als "in hoeverre bent u het eens met onderstaande stelling?" of "hoe tevreden bent u met...?".

We volgen met de keuze voor frequentie-schattingen het advies van Zapf (1989). Hij stelt dat dit één van de manieren is om zoveel mogelijk te komen tot onafhankelijke metingen van stress-concepten met behulp van een vragenlijstmethode. Indien zou worden gekozen voor instemmingsschalen c.q. satisfactie-schalen, zou het veel waarschijnlijker worden dat de meting van afzonderlijke stress-begrippen zou worden vertekend door onderliggende psychologische disposities en antwoordtendensen. De keuze voor frequentieschattingen als antwoordmogelijkheid is een manier om confounding (Kasl, 1978) tegen te gaan.

De keuze voor antwoordmogelijkheden in termen van frequenties betekent, dat nogal wat schalen uit diverse bronnen een ander antwoordformaat krijgen dan in het originele instrument.

Voor het derde blok "mogelijke gevolgen" is gekozen voor afwijkende antwoordmogelijkheden. Dit vindt zijn legitimering hierin, dat het merendeel van de voor dit blok beoogde items afkomstig is uit eerder werk bij het Studiecentrum Arbeid en Gezondheid (UVA) en de Sectie Experimentele en Arbeidspsychologie (RUG). Besloten is om hier zo dicht mogelijk bij de originelen te blijven. Deze originelen zijn voor een deel vertalingen (Hogervorst, 1992), gebaseerd op schalen zoals ontwikkeld door Warr (1990b). Het betreft de schalen "affective well-being" (emotionele reacties), "competence" (competentie), "aspiration" (aspiratie), en "negative carry-over" (herstelbehoefte). Overigens zijn dit niet de enige vertalingen van deze schalen die in Nederland in omloop zijn. Ook Schalk et al. (1995) rapporteren onderzoek naar de schaalkwaliteiten van door hen vertaalde versies van de schalen van Warr (1990b).

5.5 Het samenstellen van de itempool

Op basis van het voorgaande is een itempool samengesteld. Bij het samenstellen van de itempool is nog een aantal aanvullende zaken naar voren gekomen:

In de eerste plaats werd duidelijk dat bij sommige facetten de concrete items erg veel op elkaar leken. Er is in deze gevallen voor gekozen om de facetten te combineren.

Er is voor gekozen om de aspecten "organisatieverandering" en "verloopintentie/binding aan de organisatie" in de vragenlijst op te nemen. Naast inhoudelijke overwegingen heeft hierbij de inzetbaarheid en acceptatie van de itempool bij een aantal beoogde pilotbedrijven mede een rol gespeeld.

Bij de samenstelling van de itempool heeft op details bijstelling plaatsgevonden naar aanleiding van een aantal proefafnames van de itempool.

Dit heeft vooral betrekking gehad op de begrijpelijkheid van de vragen (moeilijke woorden, dubbele ontkenningen), de formulering van de vragen (duidelijkheid, suggestiviteit), en de accepteerbaarheid van de vragen (niet te confronterend, niet te persoonlijk). Ook mede naar aanleiding van de proefafnames is besloten een tweetal vragen over ongewenste intimiteiten toe te voegen onder het kopje "emotionele belasting".

De volledige itempool (ruim 400 items) is te vinden in Van Veldhoven (1993a; bijlage 4). We beperken ons hier tot een opsomming van de itemverzamelingen in de itempool, die betrekking hadden op PSA en werkstress (zie tabel 5.2). Dit betekent dat bepaalde aspecten die bij de facetanalyse naar voren kwamen als belangrijk, bijvoorbeeld met betrekking tot fysieke belasting, fysisch-chemische arbeidsomstandigheden en leefgewoonten verder buiten beschouwing worden gelaten. De itempool werd onderverdeeld in de drie grote blokken die bij de facetanalyse naar voren kwamen: "taak", "context van de taak", en "mogelijke gevolgen".

Tabel 5.2: Overzicht van itemverzamelingen ten behoeve van de schaalconstructie

1. Taak

- werktempo en werkhoeveelheid
- variatie in het werk
- autonomie in het werk
- zinvolheid van het werk
- mentale belasting
- emotionele belasting
- verantwoordelijkheid
- problemen bij de uitvoering van het werk
- informatievoorziening in het werk
- verandering in het werk

2. Context van de taak

- rolproblemen
- arbeidsverhoudingen collega's
- arbeidsverhoudingen directe leiding
- sociaal isolement werkplek
- inspraak
- communicatie, samenwerking en cultuur
- organisatieverandering
- leer- en loopbaanmogelijkheden
- toekomstonzekerheid
- beloning
- overige arbeidsvoorwaarden

3. Mogelijke gevolgen van het werk

- opzien tegen het werk
- competentie
- aspiratie
- betrokkenheid bij de organisatie
- herstelbehoefte
- emotionele reacties

Hoofdstuk 6: Schaalconstructie-onderzoek

In het vorige hoofdstuk is een itempool ontwikkeld voor het meten van PSA en werkstress. De volgende stap is te onderzoeken in hoeverre de itempool, naar psychometrische criteria, goede schalen bevat. Het schaalconstructie-onderzoek wordt in dit hoofdstuk beschreven, te beginnen met een beschrijving van de gevolgde methode: Mokkenanalyse.

6.1 Mokkenanalyse: een korte inleiding

Sinds in de jaren zestig de klassieke testtheorie tot een min of meer definitieve vorm kwam, is het accent qua ontwikkeling in de testtheorie vooral komen te liggen op een klasse van modellen die tegenwoordig wordt aangeduid als item-response theorie. Bekende modellen op dit gebied zijn die van Rasch, Birnbaum, en Mokken (Drenth en Sijtsma, 1990).

In de item-respons theorie wordt ervan uitgegaan dat indicaties worden verkregen over de positie van een individu op een onderliggende dimensie oftewel latente trek. Ook van vragenlijstitems kan de positie worden aangegeven op hetzelfde continuüm van een latente trek. Op basis van deze aanname wordt een aantal assumpties geformuleerd waaraan item-responses moeten voldoen. De genoemde modellen verschillen in de assumpties die worden geformuleerd. Eén van de grote verschillen met de klassieke testtheorie is dat, op basis van de mate waarin schalen voldoen aan de in een bepaald model geformuleerde assumpties, kan worden afgeleid wat het meetniveau is dat met de schaal wordt bereikt. Voor de parametrische modellen van Rasch en Birnbaum is dat een intervalniveau, voor het non-parametrische Mokken-model een ordinaal meetniveau.

Het Mokken-model werd ontwikkeld door Mokken (1971). Het was in eerste instantie alleen van toepassing op dichotome items. Door Molenaar (1982) werd de toepassing van Mokkenanalyse uitgebreid tot polychotome items. Hieronder bespreken we eerst de assumpties van het Mokken-model, daarna de parameters waarin volgens het Mokken-model de kwaliteiten van een schaal worden uitgedrukt, en tenslotte het specifieke software-pakket waarmee in dit onderzoek de Mokkenanalyse wordt uitgevoerd.

In het Mokken-model worden de volgende assumpties geformuleerd (Mokken en Lewis, 1982; Debets en Brouwer, 1989). In de eerste plaats wordt verondersteld dat alle items dezelfde latente trek meten (assumptie van unidimensionaliteit).

In de tweede plaats wordt verondersteld dat voor elke willekeurige persoon (met een bepaalde positie op het continuüm van de latente trek) de antwoorden op een verzameling van items onafhankelijk van elkaar zijn (assumptie van locale onafhankelijkheid). Ten derde wordt aangenomen dat voor elk item de kans op een positief antwoord monotoon toeneemt met de positie op de latente trek (assumptie van monotone homogeniteit). De functie die de relatie beschrijft tussen deze kans op een positief antwoord en de latente trek wordt de "item characteristic curve (ICC)" genoemd. De vierde, en meest vergaande, assumptie van het Mokken-model is dat de ICC's van de afzonderlijke items elkaar niet mogen kruisen (assumptie van dubbele monotonie). De assumptie van de dubbele monotonie veronderstelt de assumptie van de monotone homogeniteit. Een verzameling van items die voldoet aan deze vier assumpties vormt een Mokkenschaal. Scores berekend op basis van een dergelijke verzameling van items hebben een ordinaal meetniveau.

De schaalkwaliteiten volgens het Mokken-model worden uitgedrukt in een aantal parameters. De eerste parameter is Loevinger's H, een maat voor de schaalbaarheid. Deze maat geeft een indicatie over de mate waarin de items een monotoon homogene set vormen. In dit onderzoek wordt naar deze parameter verwezen onder de afkorting H(t). Alhoewel er geen harde criteria zijn voor de H(t) wordt door Mokken en Lewis (1982) de volgende vuistregel gegeven:

- < .30 = geen schaal
- .30 - .40 = zwakke schaal
- .40 - .50 = middelmatige schaal
- > .50 = sterke schaal

Ook voor individuele items kan een H-waarde worden berekend. Deze wordt aangeduid in de literatuur met de afkorting H(i). Bij Mokkenanalyse kan een minimumwaarde worden geformuleerd voor opname van een item in een schaal aan de hand van de H(i). Onder deze zogenaamde "lowerbound" vallen items buiten een schaal. In dit onderzoek is telkens gebruik gemaakt van een lowerbound van .30. Voor de volledigheid moet hier worden opgemerkt dat bij de statistische evaluatie van de H(i) gebruik wordt gemaakt van de zogenaamde Delta-ster parameter, met een significantie-niveau van 0.05. Informatie per item (met name de H(i) en de Delta-ster) is te gedetailleerd om hier verder over te rapporteren. Voor informatie op itemniveau zie Van Veldhoven (1993b: bijlage 1).

Een andere parameter die bij Mokkenanalyse kan worden berekend is de Rho(t). Dit is een schatting van de betrouwbaarheid (interne consistentie) van een schaal, vergelijkbaar met de interne consistentie zoals berekend volgens Cronbach's alpha. Ook voor de betrouwbaarheid bestaan geen harde criteria. Het streven bij het schaalconstructie-onderzoek is om een Rho(t) van minimaal .80 te bereiken voor alle schalen.

De vierde, en meest vergaande, assumptie binnen het Mokken-model (dubbele monotonie) wordt onderzocht met behulp van de zogenaamde P-matrixen. Indien een verzameling van items voldoet aan de assumptie van dubbele monotonie, zouden deze P-matrixen van links naar rechts en van boven naar beneden uitsluitend oplopende reeksen van getallen moeten laten zien. Door de P-matrixen stap voor stap langs te lopen kan worden berekend hoeveel schendingen voorkomen ten opzichte van het totaal aantal stappen. Dit wordt uitgedrukt in een percentage. Als vuistregel geldt dat minder dan 5% schendingen acceptabel is. Bij de inspectie van de P-matrixen wordt alleen gekeken naar schendingen in de eerste decimaal. Deze zijn over het algemeen te beschouwen als vrij grove schendingen.

Een laatste aspect dat bij de interpretatie van de Mokkenanalyse zal worden betrokken is de spreiding van de itemgemiddelden. Deze spreiding houdt verband met de mate van dekking die de items bieden van het onderliggende continuüm van een latente trek. Er moet voor worden gewaakt dat de items slechts een bepaald gebied van het te meten begrip afdekken, bijvoorbeeld alleen lage of alleen hoge waarden.

Het zijn vooral de schaalbaarheid ($H(t)$), de betrouwbaarheid ($Rho(t)$), en de spreiding van de itemgemiddelden, die in eerste instantie bij het schaalconstructie-onderzoek zullen worden gehanteerd. Wanneer de schaalconstructie zover is gevorderd dat er sprake is van min of meer definitieve schalen, zal worden bekeken in hoeverre de geconstrueerde schalen voldoen aan de assumptie van dubbele monotonie. Het inspecteren van de P-matrixen om schendingen van de assumptie van dubbele monotonie vast te stellen is namelijk een tijdrovende bezigheid. Er is om praktische redenen niet voor gekozen dit bij elke stap in het analyseplan te doen, maar alleen aan het eind.

Voor de Mokkenanalyses is gebruik gemaakt van het softwarepakket MSP (Debets en Brouwer, 1989; Sijtsma, Debets en Molenaar, 1991). Dit programma is zowel in staat een stapsgewijze procedure uit te voeren om uit een reeks items een adequate schaal samen te stellen (searchprocedure), als de hierboven beschreven parameters te berekenen voor een vaststaande verzameling items (testprocedure).

6.2 Onderzoeksgroep

Het schaalconstructie-onderzoek heeft plaatsgevonden van december 1992 tot en met april 1993. In deze periode werd de itempool voorgelegd aan 1228 werknemers van drie bedrijven/instellingen die zijn aangesloten bij de Arbodienst West-Brabant. Bij alle drie de bedrijven/instellingen werd de hele populatie benaderd voor het onderzoek. In totaal 812 werknemers vulden de vragenlijst in. De respons bedraagt derhalve 66%.

Gezien de lengte van de lijst en het tijdsbeslag dat per werknemer hiermee was gemoeid, is dit te beschouwen als een bevredigend responspercentage. Bij alle drie de bedrijven is dezelfde itempool afgenomen, met één uitzondering: bij het eerste bedrijf ontbrak de schaal met vragen ten aanzien van "betrokkenheid bij de organisatie".

Bij géén van de drie bedrijven/instellingen is het onderzoek geïnitieerd door de arbodienst. Telkens is door de arbodienst de vragenlijst ingepast in een onderzoek naar aanleiding van een vraag van de kant van het bedrijf. Door te werken met deze onderzoeksgroepen, is bij dit schaalconstructie-onderzoek sprake van ontwikkeling in de praktijk. Ervaringen uit de adviestrajecten bij de drie genoemde bedrijven/instellingen werden in het schaalconstructie-onderzoek mede betrokken bij beslissingen over aanpassingen van de lijst.

Hieronder volgt een korte beschrijving van elk van de drie betrokken bedrijven/instellingen.

Bedrijf 1: gezondheidszorginstelling

De respons was hier 60%. De gemiddelde leeftijd was jong, namelijk 30,3 jaar (standaarddeviatie: 6,8 jaar). Slechts 14% van de groep bestond uit mannen. Het merendeel van de werknemers zit in de rechtstreekse zorgverlening.

Bedrijf 2: gezondheidszorginstelling

Hier was de respons 62%. Met een gemiddelde leeftijd van 34,1 jaar (standaarddeviatie: 7,8 jaar) en slechts 31% mannelijke werknemers wijkt deze instelling eveneens aanzienlijk af van de gemiddelde beroepsbevolking. Een groter deel van de werknemers binnen deze instelling dan bij bedrijf 1 zit in ondersteunende dan wel staffuncties.

Bedrijf 3: productiebedrijf

De respons bedroeg 87%. Van de respondenten is 84% mannelijk. De gemiddelde leeftijd is 36,1 jaar (standaarddeviatie: 9,9 jaar). Drie-kwart van de medewerkers in dit bedrijf is werkzaam in productie- en aanverwante functies.

Al met al is de onderzoeksgroep sterk afwijkend qua samenstelling ten opzichte van de gemiddelde beroepsbevolking in de regio West-Brabant. Relatief sterk vertegenwoordigd zijn: werknemers in de gezondheidszorg, vrouwen, en jonge werknemers. Dit is het gevolg van de keuze om te ontwikkelen in de praktijk.

6.3 Samenstellen ontwikkelgroep en testgroep

Besloten werd om de afwijkende samenstelling van de onderzoeksgroep te ondervangen door het maken van een "ontwikkelgroep" en een "testgroep". De

ontwikkelgroep werd eerst zodanig samengesteld dat alle leeftijden, sexen en functietypen gelijkelijk waren vertegenwoordigd. Dit met het oog op de generaliseerbaarheid van de resultaten. De testgroep werd gevormd door de resterende werknemers, Werknemers in de gezondheidszorg, vrouwen, en jonge werknemers zijn in de testgroep sterk oververtegenwoordigd.

Voor alle drie de bedrijven waren per werknemer gegevens beschikbaar over de leeftijd, het geslacht, en het type functie. Dit laatste is gencodeerd tot vier functieklassen te weten: kantoor/leidinggevend, productie, gezondheidszorg, en rest. Onder gezondheidszorg wordt in dit verband verstaan: werknemers met direct patiëntcontact in hun functie. Werknemers zonder direct patiëntcontact worden ingedeeld in de restcategorie.

Een bevredigende oplossing voor de ontwikkelgroep was alleen mogelijk door bij de vrouwen de werkneemsters in productie- en rest-functies samen te voegen, en door bij de leeftijd de grens voor splitsing tussen jong en oud te leggen bij 35 jaar. De kleinste groep die aldus werd gevonden, was de groep mannen in kantoor/leidinggevende functies onder de 35 jaar (N=27). Besloten werd om uit elk van de 12 cellen (jong/oud x man/vrouw x kantoor en leiding/productie/gezondheidszorg) een random steekproef te trekken van N=27 ten behoeve van de ontwikkelgroep. De overige werknemers werden geplaatst in de testgroep. Aldus werd een ontwikkelgroep samengesteld van 12 (cellen) x 27 (werknemers), oftewel 324 werknemers. De testgroep bestaat uit 488 werknemers (inclusief vier werknemers waarvan de benodigde persoonsgegevens ontbraken om ze in de hier beschreven indeling te kunnen plaatsen). Hieronder volgt in tabel 6.1 de verdeling van de ontwikkel- (O) en testgroep (T) over de kenmerken:

Tabel 6.1: Samenstelling ontwikkelgroep en testgroep Mokkenanalyse

	jonger dan 35				ouder dan 35			
	man		vrouw		man		vrouw	
	O	T	O	T	O	T	O	T
kantoor/leiding	27	0	27	17	27	29	27	1
productie (man)	27	14	-	-	27	2	-	-
productie/rest (vrouw)	-	-	27	44	-	-	27	17
gezondheidszorg	27	18	27	181	27	10	27	57
rest (man)	-	42	-	-	-	44	-	-

Vervolgens vond de schaalconstructie plaats aan de hand van een gedetailleerd analyseplan, dat vooraf werd opgesteld. Hieronder volgt een stapsgewijs verslag van de schaalconstructie op hoofdlijnen. Voor een gedetailleerd verslag verwijzen we naar Van Veldhoven (1993b).

6.4 Schaalconstructie

6.4.1 Stap 1: Mokken-searchprocedure op de itemverzamelingen

Alhoewel er geen hard criterium voor bestaat lijkt een verzameling van vier items het minimum te zijn, wil er nog gesproken kunnen worden van een schaal. Van de 28 onderzochte itemverzamelingen blijkt bij vier géén eerste cluster te vinden te zijn van vier of meer items. Deze itemverzamelingen zijn niet unidimensioneel. Het gaat om de itemverzamelingen: verantwoordelijkheid, problemen bij de taakuitvoering, organisatieverandering, en overige arbeidsvoorwaarden.

Bij de 24 overige itemverzamelingen blijken er wel minstens vier items te clusteren binnen de eerste schaal. Ook hier blijken echter bij meerdere itemverzamelingen "snipper-schalen" te ontstaan van twee à drie items. Dit duidt eveneens op niet-unidimensionaliteit van de itemverzamelingen. Soms lijkt het nodig om het label van de eerstgevonden schaal te veranderen ten gevolg van de selectie van items die blijkens de Mokken-searchprocedure binnen de eerste schaal vallen.

Er is één itemverzameling waarin twee schalen worden gevonden met vier of meer items. Dit is de schaal rolproblemen. De Mokken-searchprocedure formeert hieruit twee afzonderlijke schalen: rolconflict en rolonduidelijkheid. In het vervolg zullen we deze twee dan ook afzonderlijk meenemen in de schaalconstructie.

De resultaten van de Mokken-searchprocedure voor de 25 schalen met minimaal vier items zijn samengevat in tabel 6.2. In de tabel zijn opgenomen: label van de schaal, aantal items: $N(i)$, schaalbaarheid van de schaal: $H(t)$, betrouwbaarheid: $Rho(t)$, en de range van de itemgemiddelden: $Range(ig)$. De itemgemiddelden kunnen bij de meeste schalen lopen van minimaal 1 tot maximaal 4. Een uitzondering hierop vormen de schalen "opzien tegen het werk", "competentie", "aspiratie", "betrokkenheid bij de organisatie" en "herstelbehoefte". Hier loopt de range van de itemgemiddelden van minimaal 1 tot maximaal 2.

Tabel 6.2: Resultaten Mokken-searchprocedure ontwikkelgroep

Schaallabel	N(i)	H(t)	Rho(t)	Range(ig)
werktempo en -hoeveelheid	15	.46	.91	1.40 - 3.69
taakvariatie	7	.39	.83	1.57 - 2.60
taakautonomie	12	.44	.91	2.07 - 2.60
taakzinvolheid	5	.44	.80	1.52 - 2.00
mentale belasting	8	.46	.86	2.55 - 3.55
emotionele belasting	10	.38	.85	1.15 - 2.60
taakinformatie	9	.42	.86	1.98 - 2.94
taakverandering	14	.37	.85	1.38 - 2.34
rolconflict	7	.45	.80	1.40 - 1.87
rolonduidelijkheid	5	.45	.81	1.73 - 2.43
arbeidsverh. collega's	9	.52	.87	1.23 - 2.11
arbeidsverh. directe leiding	9	.58	.90	1.15 - 2.17
sociaal isolement werkplek	4	.46	.77	1.84 - 2.45
inspraak	10	.41	.88	1.85 - 2.90
comm., samenw. en cultuur	12	.40	.89	2.05 - 2.80
leer- en loopbaanmog.	12	.37	.89	2.03 - 3.44
toekomstonzekerheid	6	.71	.96	1.85 - 1.99
beloning	5	.45	.80	1.88 - 2.61
opzien tegen het werk	17	.46	.88	1.05 - 1.36
competentie	7	.38	.70	1.01 - 1.28
aspiratie	5	.48	.70	1.03 - 1.29
betrokkenheid bij org.	17	.46	.87	1.02 - 1.84
herstelbehoefte	17	.45	.90	1.06 - 1.29
emotionele reacties	12	.43	.89	1.41 - 2.17

Wanneer we deze schalen bekijken qua schaalbaarheid (criterium: H-waarde minimaal .40) en betrouwbaarheid (criterium: Rho-waarde minimaal .80) dan zien we het volgende. Eén schaal schiet ten aanzien van beide bovenstaande criteria tekort: dat is de schaal "competentie". Een viertal schalen voldoet niet qua schaalbaarheid. Dit zijn: "taakvariatie", "emotionele belasting", "taakverandering", en "leer- en loopbaanmogelijkheden". De schalen "sociaal isolement op de werkplek" en "aspiratie" tenslotte, voldoen niet qua betrouwbaarheid. Al met al resteren er in dit stadium 17 schalen die aan de gestelde criteria voldoen. Dit is een hoopgevende start.

6.4.2 Stap 2: Aanvullen gereduceerde itemverzamelingen vanuit de rest van de itempool

Van alle 25 schalen met vier items of meer werden schaalscores berekend door de itemscores op te tellen. Bij alle schaalscores is ervoor gezorgd dat de hoogste score is verbonden aan de veronderstelde negatieve pool van het gemeten begrip.

Vervolgens zijn alle afzonderlijke items gecorreleerd met de 25 schaalscores. Het doel hiervan was om items op te sporen die in een bepaalde itemverzameling zijn geplaatst, maar toch hoog correleren met een andere schaalscore. Als drempelwaarde voor mogelijke opname in een andere schaal is een correlatie van .70 gekozen. Dit omdat een dergelijke correlatie in de buurt komt van de ondergrens, die in psychometrisch onderzoek vaak wordt gesteld aan de interne consistentie van een schaal.

Geen enkel item bleek aan deze voorwaarde te voldoen. We kunnen over het resultaat van stap 2 dan ook kort zijn: deze stap leidt niet tot aanpassingen van de lijst.

Toch werden er aanzienlijke correlaties gevonden tussen de schaalscores en sommige items uit andere itemverzamelingen. In totaal werden zes correlaties gevonden van .60 of meer. Er werden 50 correlaties gevonden van .50 of meer. Tenslotte werden 243 correlaties gevonden van .40 of meer. Dit lijkt misschien veel, maar er werden meer dan 9000 correlaties berekend. Zo bezien is er eerder sprake van een redelijk goede classificatie van items in itemverzamelingen, met een acceptabele conceptuele overlap.

6.4.3 Stap 3: Verder inkorten en aanpassen itemverzamelingen en tussentijdse Mokken-testprocedure

Op basis van de informatie uit het schaalconstructie-onderzoek tot nu toe (met name de resultaten van de Mokken-searchprocedure), en de ervaringen bij de pilot-bedrijven, is het mogelijk om de itemverzamelingen verder in te korten en aan te passen. Een reeks van criteria speelt hierbij een rol.

In de eerste plaats zijn er psychometrische criteria voor itemreductie. Het streven is om, waar mogelijk, schalen te ontwikkelen van vier items of meer, met een schaalbaarheid (H-waarde) van .40 of hoger, en met een betrouwbaarheid (Rho-waarde) van .80 of hoger.

Verder is het wenselijk dat de schalen zo min mogelijk items bevatten met dezelfde moeilijkheidsgraad. Dit is vertaald in: zo min mogelijk items met gemiddelden binnen dezelfde .10 van de potentiële antwoordrange. In het geval

dat meerdere items ongeveer dezelfde moeilijkheidsgraad hebben, kan het item met de hoogste item H-waarde ($H(i)$) worden gekozen, mits ook inhoudelijk sprake is van dicht tegen elkaar aanliggende items. Vaak moet bij de inkorting van een schaal een afweging worden gemaakt van schaallengte ten opzichte van betrouwbaarheid (die zakt immers ten gevolge van schaalverkorting).

Ook is een zo adequaat mogelijke dekking gewenst van de potentiële antwoordrange door de items in de schaal. Met andere woorden: de itemgemiddelden horen redelijk gespreid te zijn qua moeilijkheidsgraad. Hierbij moet meestal een afweging worden gemaakt tussen dekking van de antwoordrange en schaalbaarheid van de schaal.

In de tweede plaats weegt een aantal inhoudelijke criteria mee bij deze fase. Denk bijvoorbeeld aan pragmatische bezwaren die bij het uittesten van de itempool zijn gerezen ten aanzien van bepaalde vragen. Het gaat hierbij zowel om aspecten die naar voren kwamen bij de dataverzameling als bij de advisering. Een ander aspect dat meeweegt is de plaats van het facet in de facetanalyse zoals gerapporteerd in hoofdstuk 5. Met andere woorden: er wordt rekening mee gehouden of een facet binnen de instrumenten-literatuur op het gebied van PSA en werkstress meer of minder vaak voorkomt. Een laatste aspect is inhoudelijke zuiverheid. Zo is er één vraag ("uitdaging") die in drie variaties in de itempool staat, binnen drie verschillende itemverzamelingen. Op dergelijke punten moet de lijst worden gezuiverd om item-overlap te vermijden.

De bovenstaande criteria kunnen niet blind op elke itemverzameling worden toegepast. Ze wegen ook niet bij elke itemverzameling even zwaar. Het inkorten van schalen is tot op zekere hoogte maatwerk. Op grond van deze criteria zijn per itemverzameling voorstellen gedaan om de schaalkwaliteiten te verbeteren en tegelijkertijd de schaal in te korten.

Bij een aantal itemverzamelingen is ervoor gekozen om verder geen schalen op basis hiervan te ontwikkelen. Dit is gebeurd bij: "taakzinvolheid", "verantwoordelijkheid", "problemen bij de taakuitvoering", en "organisatieverandering". Bij een tweetal itemverzamelingen is ervoor gekozen om deze op te splitsen in twee onderdelen. Zo werd "betrokkenheid bij de organisatie" gesplitst in "verloopintentie" en "betrokkenheid bij de organisatie". "Herstelbehoefte" werd gesplitst in "piekeren" en "herstelbehoefte". Daar waar "snipper-schalen" ontstonden, die niet tot acceptabele schalen leidden, zijn deze geplaatst onder de categorie restvragen. Hetzelfde gold voor enkele losse items, die onmisbaar leken in een vragenlijst naar PSA en werkstress.

Vervolgens zijn de aldus herziene schalen onderworpen aan een Mokken-testprocedure op de ontwikkelgroep. In tabel 6.3 worden de schaalkwaliteiten gerapporteerd voor deze schalen.

Tabel 6.3: Resultaten tussentijdse Mokken-testprocedure ontwikkelgroep

Schaallabel	N(i)	H(t)	Rho(t)	Range(ig)
werktempo en -hoeveelheid	11	.48	.89	1.40 - 3.69
taakvariatie	6	.42	.82	1.97 - 2.60
taakautonomie	11	.44	.90	2.07 - 2.60
mentale belasting	7	.51	.87	2.92 - 3.55
emotionele belasting	7	.45	.85	1.46 - 2.60
taakinformatie	7	.41	.83	1.98 - 2.94
taakverandering	5	.62	.86	1.38 - 2.05
rolconflict	6	.51	.80	1.40 - 1.87
rolonduidelijkheid	5	.45	.81	1.73 - 2.43
arbeidsverh. collega's	9	.52	.87	1.23 - 2.11
arbeidsverh. directe leiding	9	.58	.90	1.15 - 2.17
sociaal isolement werkplek	4	.46	.77	1.84 - 2.45
inspraak	8	.42	.85	1.85 - 2.90
communicatie	9	.45	.87	2.05 - 2.80
leer- en loopbaanmog.	9	.41	.87	2.03 - 3.44
toekomstonzekerheid	4	.75	.95	1.90 - 1.99
beloning	5	.45	.80	1.88 - 2.61
opzien tegen het werk	10	.49	.81	1.05 - 1.36
competentie	6	.38	.63	1.01 - 1.28
aspiratie	6	.36	.61	1.03 - 1.29
betrokkenheid bij org.	8	.54	.79	1.09 - 1.84
verloopintentie	6	.85	.92	1.08 - 1.46
herstelbehoefte	11	.48	.87	1.10 - 1.29
piekeren	6	.67	.86	1.06 - 1.25
emotionele reacties	12	.43	.89	1.41 - 2.17

Per kolom kunnen we een aantal opmerkingen plaatsen bij de resultaten.

Schaallengte

Twee schalen hebben het minimum van vier items. Voor "sociaal isolement" kan dit niet worden verbeterd: er zijn in de itempool niet meer van dergelijke items beschikbaar. Voor "toekomstonzekerheid" zijn niet meer items nodig, gezien de hoge H(t). Vijf schalen hebben tien items of meer, wat aan de lange kant is. Dit zijn echter meestal schalen die in de facetanalyse naar voren kwamen als centraal voor het gebied PSA en werkstress: het belang rechtvaardigt de lengte van deze schalen. De overige schalen hebben een schaallengte tussen de vijf en tien items.

Schaalbaarheid

Met uitzondering van de schalen "competentie" en "aspiratie" voldoen alle schalen aan het gestelde criterium van een $H(t)$ van minimaal .40. Van de overige schalen hebben er 15 een schaalbaarheid tussen .40 en .50. Volgens de normen van Mokken en Lewis (1982) geldt dit als een middelmatige schaal. Verder hebben negen schalen een schaalbaarheid boven de .50. Deze kunnen gelden als sterke schalen.

Schalen met een extreem hoge $H(t)$ kunnen waarschijnlijk verder worden ingekort. Voor "toekomstonzekerheid" heeft dit er reeds toe geleid dat het aantal items werd beperkt tot een minimum van vier. Voor "verloopintentie" en "piekeren" valt dit eveneens te overwegen.

De schaal "verloopintentie" blijkt inderdaad korter te kunnen. Er is voor gekozen om twee items te laten vervallen die betrekking hebben op de bedrijfstak. De gegevens in de voorgaande tabel worden na deze ingreep: $N(i)=4$, $H(t)=.90$, $Rho(t)=.90$, $Range=1.13-1.46$. De schaalbaarheid neemt door deze ingreep kennelijk nog verder toe.

Bij de schaal "piekeren" zijn de twee items met de meest ingewikkelde formulering geschrapt. De tabelgegevens worden dan: $N(i)=4$, $H(t)=.67$, $Rho(t)=.80$, $Range=1.07-1.25$. De schaalqualiteiten blijven bij deze schaal dus nagenoeg gelijk.

Betrouwbaarheid

De betrouwbaarheid is uitgedrukt in de parameter Rho . Over het algemeen wijkt deze maat niet erg veel af van de interne consistentie van een schaal, zoals berekend volgens de klassieke methode (Cronbach's α). Van drie schalen wisten we al dat de betrouwbaarheid onder het criterium van .80 zou blijven: "sociaal isolement", "competentie", en "aspiratie". Daar komt nu echter bij de schaal "betrokkenheid". Bekeken kan worden of de betrouwbaarheid hier nog kan worden verbeterd. Het toevoegen c.q. weglaten van bepaalde items blijkt niet te leiden tot een verbetering van de betrouwbaarheid van de "betrokkenheid"-schaal. Anderzijds ligt de betrouwbaarheid van .79 zo dicht bij het criterium, dat is besloten om de schaal in deze vorm te handhaven.

Spreiding van de itemgemiddelden over de potentiële antwoordrange

Dit punt is het moeilijkst uit te drukken in criteria. Bij aanpassing van de schalen is zoveel mogelijk geprobeerd de bestaande spreiding aan de extremen te bewaren, en in het middenstuk zo min mogelijk doublures te hebben qua populariteit van de items. Wanneer we de bovenstaande resultaten bezien, dan is er een aantal schalen te noemen waarvan de spreiding van de itemgemiddelden minder goed is.

Bij de volgende schalen is de afstand gering tussen het minimum en het maximum van de gemiddelde itemscores (minder dan 20% van de potentiële antwoordrange):

- taakautonomie
- fysieke inspanning
- rolconflict
- toekomstonzekerheid
- herstelbehoefte
- piekeren

Van deze schalen is vooral "toekomstonzekerheid" problematisch. Hier is de afstand tussen minimum en maximum minder dan 10% van de potentiële antwoordrange.

6.4.4 Stap 4: Redundantie-onderzoek voorgestelde schalen

De redundantie van de voorgestelde schalen ten opzichte van elkaar zal op een eenvoudige wijze worden bekeken door de schaalscores onderling te correleren. Voor alle schalen, zoals ze tot hier aan toe zijn geconstrueerd, zijn scores berekend die lopen van laag (gunstige score) naar hoog (ongunstige score). De somscores zijn gedeeld door het aantal items, om de antwoordrange van de schalen zoveel mogelijk te standaardiseren. Voor de correlaties worden vooraf de volgende criteria geformuleerd:

< .20	geldt als "gering"
tussen .20 en .40	geldt als "matig"
tussen .40 en .60	geldt als "redelijk (hoog)"
boven .60	geldt als "hoog"

De correlaties zijn berekend met behulp van pairwise deletion. Een compleet overzicht van de correlaties is te vinden in Van Veldhoven (1993b: paragraaf 3.4). We behandelen hier alleen de opvallende correlaties. Dit doen we blok voor blok.

We beginnen met de correlaties tussen taakaspecten onderling. De minimale N per correlatie is 290. We vinden dan dat de taakaspecten een redelijk grote onderlinge onafhankelijkheid vertonen. Er worden geen correlaties gevonden boven de .60, en slechts één tussen .40 en .60. "Emotionele belasting" correleert namelijk .41 met "werktempo/werkhoeveelheid".

Vervolgens gaan we naar de "context van de taak" onderling. De minimale N per correlatie, opnieuw met pairwise deletion, is 280. De situatie ligt hier duidelijk anders. In de eerste plaats zijn er twee correlaties aan te wijzen die als hoog mogen gelden. De beoordeling van de "arbeidsverhoudingen met collega's" versus de "directe leiding" correleert .60. Wellicht speelt hierin mee dat het in feite om

parallelversies gaat, waarin alleen het object van de vraag is gewijzigd. In de tweede plaats correleren de schalen "arbeidsverhoudingen met de leiding" en "communicatie" hoog met elkaar. Dit is niet zo vreemd, aangezien in de laatstgenoemde schaal een aantal expliciete vragen over de directe leiding zijn opgenomen. Verder kunnen we elf correlaties aanwijzen tussen .40 en .60. Dit geeft aan dat er nogal wat overlap is tussen aspecten op het gebied van de "context van de taak". Het zijn vooral de schalen "rolconflict", "rolonduidelijkheid", "arbeidsverhoudingen collega's/leiding", "inspraak", en "communicatie", die overlap vertonen met elkaar. Hier zal een beslissing moeten worden genomen over het al dan niet handhaven/combineren van schalen.

Vervolgens komen we bij de correlaties tussen de "mogelijke gevolgen" onderling. De minimale N per correlatie is 264. Dit is een gevolg van het feit dat de schaal "betrokkenheid" niet bij het eerste bedrijf is afgenomen. Er is een hoge correlatie (.65) tussen "opzien tegen het werk" en "verloopintentie". Verder zien we zes correlaties tussen .40 en .60. Verwachte correlaties zijn er bij "betrokkenheid" versus "verloopintentie" (.43) en bij "herstelbehoefte" versus "piekeren" (.60). Deze schalen komen immers voort uit dezelfde itemverzamelingen.

Hierna kijken we naar de correlaties tussen taakaspecten en aspecten van de context van de taak. De minimale N bedraagt hier 280. Er worden geen echt hoge correlaties gevonden, wel een aantal redelijk hoge. Zo correleert "taakverandering" .47 met "rolconflict". "Gebrek aan taakinformatie" correleert .52 met "rolonduidelijkheid". "Gebrek aan taakinformatie" correleert tevens .46 met een "arbeidsverhoudingen directe leiding". "Gebrek aan taakinformatie" correleert verder .53 met "gebrek aan inspraak". "Gebrek aan taakinformatie" correleert .58 met een minder goede "communicatie". Tenslotte vinden we een redelijk hoge correlatie bij de schaal "(gebrek aan) leer- en loopbaanmogelijkheden". Die correleert namelijk .58 met "gebrek aan taakvariatie".

We bespreken nu de correlaties tussen taakaspecten en mogelijke gevolgen. De minimale N bedraagt ditmaal 260. We vinden hier slechts twee redelijk hoge correlaties. De eerste zit tussen "gebrek aan taakvariatie" en "gebrek aan aspiratie" (.40). De tweede tussen "emotionele belasting" en "herstelbehoefte" (.46).

Tenslotte komen we bij de correlaties tussen context van de taak en mogelijke gevolgen. De minimale N bedraagt 259. Opnieuw vinden we geen echt hoge correlaties, wel redelijk hoge. "Gebrek aan leer- en loopbaanmogelijkheden" correleert .43 met "gebrek aan aspiratie". "Betrokkenheid" en "relatie met de leiding" correleren .42 met elkaar. "Rolconflict", tenslotte, heeft een correlatie van .43 met "verloopintentie".

De correlaties overziend, blijkt dat de tot hier aan toe geconstrueerde schalen redelijk onafhankelijk van elkaar zijn.

De enige uitzondering hierop vormen de voorgestelde schalen binnen het blok context van de taak, die betrekking hebben op sociaal-organisatorische aspecten. Voordat we hierop ingaan dient opgemerkt te worden dat géén van de correlaties zo hoog is, dat die in de buurt komt van de betrouwbaarheid (lees: interne consistentie) van de betrokken schalen. Dit betekent dat het onderscheiden van de schalen, zoals dat hier gebeurt, altijd meer informatie oplevert dan het weglaten of indikken van schalen. Dit pleit voor de gevolgde strategie van facetanalyse via Mokken-searchprocedure naar analyse van de redundantie van schalen.

Wat te doen met de redundantie in de sociaal-organisatorische schalen?

Wanneer we de correlatiematrix bekijken tussen de variabelen "rolconflict", "rolonduidelijkheid", "arbeidsverhoudingen collega's/leiding", "inspraak", en "communicatie" onderling, dan blijkt het volgende. "Rolconflict", "arbeidsverhoudingen collega's", en "inspraak" hebben elk slechts één correlatie van .50 of hoger met de overige variabelen. "Rolonduidelijkheid" heeft twee van dat soort correlaties. "Communicatie" heeft er drie, en "arbeidsverhoudingen directe leiding" vier. Met name de laatste twee blijken dus redundant te zijn.

De facetanalyse onderstreept het belang van arbeidsverhoudingen in de vragenlijst, met aparte schalen voor collega's en leiding. Qua praktische bruikbaarheid van de vragenlijst is een uitsplitsing naar collega's versus directe leiding wenselijk, en dit wordt ook veelvuldig teruggevonden in bestaande meetinstrumenten.

De beste kandidaat om te schrappen is de schaal "communicatie". Het is echter voorstelbaar dat de hoge correlaties die deze schaal genereert ontstaan doordat hierin vragen over de directe leiding en de directie zijn opgenomen. Laten we derhalve eens bekijken wat de correlaties zijn na weglating van deze items uit de schaal. We krijgen dan als correlaties met "communicatie":

rolconflict	: .33
rolonduidelijkheid	: .51
arbeidsverhoudingen collega's:	: .34
arbeidsverhoudingen directe leiding	: .51
inspraak	: .56

Nog steeds blijken er drie correlaties boven de .50 te bestaan. Wel is de hoogste correlatie (.63) tussen "communicatie" en "arbeidsverhoudingen directe leiding" teruggebracht tot .51. Deze correlaties zijn te accepteren mits de verkorte "communicatie"-schaal, die nu nog maar uit vier items bestaat, aan een aantal elementaire psychometrische voorwaarden voldoet. Uit een Mokken-testprocedure komen als resultaten: $N(i)=4$, $H(t)=.48$, $Rho=.79$, Range: 2.09-2.67. Gezien de schaallengte is dit acceptabel.

6.4.5 Stap 5: Mokken-testprocedure voorgestelde schalen op de ontwikkelgroep

In het bovenstaande zijn stapsgewijs schalen geconstrueerd op basis van de ontwikkelgroep. Vanwege de vele tussentijdse aanpassingen is het wenselijk de geconstrueerde schalen en hun gegevens hier nog eens op te nemen in een overzicht (zie tabel 6.4). We beperken ons dan tot de $N(i)$, $H(t)$, en $Rho(t)$. Tevens zullen we nu voor het eerst onderzoeken in hoeverre de assumptie van de dubbele monotonie wordt geschonden. Dit gebeurt door een inspectie van de P-matrixen. Er is sprake van een schending wanneer de getallen niet van links naar rechts en van boven naar beneden oplopen (in de eerste decimaal; dit betekent dat alleen wordt gekeken naar vrij grove schendingen). Als criterium geldt dat indien het percentage schendingen onder de 5% blijft, dit acceptabel is. Het percentage schendingen wordt in tabel 6.4 weergegeven in de kolom %S(t).

Samenvatting schaalconstructie

We kunnen nu samenvatten wat de resultaten zijn van het schaalconstructie-onderzoek bij de ontwikkelgroep.

De totale itempool omvatte ruim 400 vragen. Hieruit zijn 25 schalen geconstrueerd op het gebied van PSA en werkstress.

Qua schaalengte zitten vijf schalen op het minimum van vier items. Drie van deze schalen zijn echter juist zo kort gehouden vanwege een dusdanige schaalbaarheid dat meer items overbodig zijn. De andere twee waren op basis van de itempool niet langer te maken.

De schaalbaarheid voldoet slechts bij twee schalen ("competentie" en "aspiratie") niet aan het criterium van minimaal .40. Dit betekent dat de overige schalen variëren van middelmatige tot sterke schalen volgens de vuistregel van Mokken en Lewis (1982).

De betrouwbaarheid haalt bij vijf schalen niet het criterium van .80. Bij drie schalen zit de betrouwbaarheid er wel dicht tegen aan (tussen .70 en .80). De twee resterende schalen zijn opnieuw de schalen "competentie" en "aspiratie".

Wat betreft de assumptie van de dubbele monotonie is er een drietal schalen dat het criterium van 5% schendingen overschrijdt. Dit zijn: "beloning", "betrokkenheid", en "verloopintentie". Geen van deze schalen komt echter uit boven de 10% schendingen, zodat de situatie nergens echt kritiek wordt. Het merendeel van de geconstrueerde schalen voldoet aan de belangrijkste assumptie van het Mokken-model: die van de dubbele monotonie.

Tabel 6.4: Resultaten Mokken-testprocedure voorgestelde schalen ontwikkelgroep

Schaallabel	N(i)	H(t)	Rho(t)	%S(t)	Range(ig)
werktempo/hoeveelheid	11	.48	.89	0.6%	1.40 - 3.69
taakvariatie	6	.42	.82	1.3%	1.97 - 2.60
taakautonomie	11	.44	.90	1.6%	2.07 - 2.60
mentale belasting	7	.51	.87	0.7%	2.92 - 3.55
emotionele belasting	7	.45	.85	0.2%	1.46 - 2.60
taakinformatie	7	.41	.83	2.6%	1.98 - 2.94
taakverandering	5	.62	.86	0.0%	1.38 - 2.05
rolconflict	6	.51	.80	0.0%	1.40 - 1.87
rolonduidelijkheid	5	.45	.81	0.0%	1.73 - 2.43
arbeidsverh. collega's	9	.52	.87	0.7%	1.23 - 2.11
arbeidsverh. directe leiding	9	.58	.90	1.9%	1.15 - 2.17
sociaal isolement	4	.46	.77	0.0%	1.84 - 2.45
inspraak	8	.42	.85	3.4%	1.85 - 2.90
communicatie	4	.48	.79	0.0%	2.09 - 2.67
leer- en loopbaanmog.	9	.41	.87	2.8%	2.03 - 3.44
toekomstonzekerheid	4	.75	.95	0.0%	1.90 - 1.99
beloning	5	.45	.80	5.2%	1.88 - 2.61
opzien tegen het werk	9	.50	.79	0.0%	1.05 - 1.36
competentie	6	.38	.63	0.0%	1.01 - 1.28
aspiratie	6	.36	.61	0.0%	1.03 - 1.29
betrokkenheid bij org.	8	.54	.79	5.4%	1.09 - 1.84
verloopintentie	4	.90	.90	8.3%	1.13 - 1.46
herstelbehoefte	11	.48	.87	4.5%	1.10 - 1.29
piekeren	4	.67	.80	0.0%	1.07 - 1.25
emotionele reacties	12	.43	.89	1.9%	1.41 - 2.17

6.4.6 Stap 6: Mokken-testprocedure definitieve schalen op de testgroep

We kunnen nu voor de testgroep vergelijkbare berekeningen maken als bij stap 5 voor de ontwikkelgroep. Zo is een eerste controle op de waarde van de schalen mogelijk. De schaalqualiteiten in een ontwikkelgroep zijn waarschijnlijk geflatteerd omdat ze door het sleutelen aan de schalen kunnen worden opgeschroefd. We dienen wel te bedenken dat een testgroep uit andere bedrijven dan die van de ontwikkelgroep in dit opzicht een meer kritische controle zou hebben opgeleverd.

In de testgroep zitten 488 werknemers (zie paragraaf 6.3). De N per Mokkenanalyse fluctueert tussen minimaal 439 en maximaal 479. De "betrokkenheid"-schaal zit daar weer onder (N=357), omdat die alleen in bedrijf 2 en 3 werd afgenomen. Wanneer we de Mokken-testprocedure loslaten op de testgroep (N=488) dan krijgen we de resultaten van tabel 6.5. We vermelden hier tevens per schaal de range van de itemgemiddelden en het percentage schendingen van de assumptie van de dubbele monotonie.

Tabel 6.5: Resultaten Mokken-testprocedure definitieve schalen testgroep

Schaallabel	N(i)	H(t)	Rho(t)	%S(t)	Range(ig)
werktempo/hoeveelheid	11	.50	.89	0.5%	1.37 - 3.67
taakvariatie	6	.43	.82	1.6%	1.94 - 2.51
taakautonomie	11	.42	.89	1.0%	2.02 - 2.65
mentale belasting	7	.44	.83	2.6%	2.80 - 3.52
emotionele belasting	7	.41	.82	0.5%	1.46 - 2.58
taakinformatie	7	.42	.82	0.7%	2.06 - 2.97
taakverandering	5	.69	.87	1.0%	1.37 - 2.02
rolconflict	6	.43	.75	0.0%	1.37 - 1.84
rolonduidelijkheid	5	.38	.76	0.5%	1.77 - 2.64
arbeidsverh. collega's	9	.51	.87	1.4%	1.22 - 2.04
arbeidsverh. directe leiding	9	.56	.90	1.3%	1.12 - 2.29
sociaal isolement	4	.46	.76	1.5%	1.82 - 2.53
inspraak	8	.39	.83	4.0%	1.94 - 2.96
communicatie	4	.46	.78	0.0%	2.15 - 2.68
leer- en loopbaanmog.	9	.40	.86	2.8%	2.00 - 3.44
toekomstonzekerheid	4	.68	.93	0.8%	1.78 - 1.90
beloning	5	.48	.82	1.0%	2.02 - 2.68
opzien tegen het werk	9	.41	.72	0.0%	1.02 - 1.40
competentie	6	.41	.66	0.0%	1.02 - 1.31
aspiratie	6	.21	.45	0.0%	1.02 - 1.27
betrokkenheid bij org.	8	.47	.78	5.4%	1.08 - 1.83
verloopintentie	4	.89	.91	8.3%	1.13 - 1.48
herstelbehoefte	11	.50	.88	1.8%	1.10 - 1.28
piekeren	4	.66	.77	0.0%	1.04 - 1.24
emotionele reacties	12	.38	.87	2.3%	1.39 - 2.07

Wanneer we kijken naar de schaalbaarheid - $H(t)$ - dan valt één schaal door de mand, namelijk: "aspiratie" (.21). Drie andere schalen voldoen net niet aan het criterium van minimaal .40: "rolonduidelijkheid", "inspraak", en "emotionele reacties".

De $H(t)$ blijft hier echter steeds acceptabel (boven de .35). Er treedt in de testgroep een beperkte verslechtering op met betrekking tot de schaalbaarheid. Dit was op grond van de samenstelling van de testgroep (veel minder heterogeen dan de ontwikkelgroep) te verwachten.

Bij de betrouwbaarheid - $Rho(t)$ - valt opnieuw de schaal "aspiratie" door de mand (.45). De schaal "competentie" zit eveneens onder de .70. Ook qua betrouwbaarheid is er een aantal schalen dat net niet het criterium van minimaal .80 haalt. Deze blijven echter allemaal tussen de .70 en de .80, wat acceptabel is. Deze schalen zijn: "rolconflict", "rolonduidelijkheid", "sociaal isolement", "communicatie", "opzien tegen het werk", "betrokkenheid", en "piekeren". Het zijn vooral de wat kortere schalen die hier problemen mee hebben. Ook ten aanzien van de betrouwbaarheid is het resultaat bij de testgroep dus iets minder goed dan bij de ontwikkelgroep, wat te verwachten was.

Voor wat betreft de schendingen van de assumptie van dubbele monotonie is het resultaat bij de testgroep iets beter dan bij de ontwikkelgroep. Twee schalen halen, net als bij de ontwikkelgroep, niet het criterium van minder dan 5% schendingen. Dit zijn "betrokkenheid" en "verloopintentie". Deze schalen blijven echter ook deze keer onder de 10% schendingen. De schaal "beloning" blijft bij de testgroep, in tegenstelling tot bij de ontwikkelgroep, onder de 5% schendingen.

De resultaten van de Mokken-testprocedure bij de testgroep vallen zeker niet tegen. Het is wel wat minder fraai dan bij de ontwikkelgroep, maar een groot deel van de schalen blijft goede schaalkwaliteiten houden. Deze conclusie mag zeker worden getrokken als we bedenken dat de testgroep ten opzichte van de ontwikkelgroep uitermate onevenwichtig is samengesteld, omdat er veel meer jonge, vrouwelijke werknemers in de gezondheidszorg in zitten. Wel kan op grond van deze test extra worden getwijfeld aan de schaalkwaliteiten van de schaal "aspiratie".

Dit is een zoveelste argument tegen de schaalkwaliteiten van de twee Warr-vertalingen "competentie" en "aspiratie". Warr (1990b) rapporteert zelf voor de schaal "werkgerelateerde competentie" een Cronbach's Alpha van .68 en voor "werkgerelateerde aspiratie" een Cronbach's Alpha van .62. Dit is vergelijkbaar met de $Rho(t)$ die in dit onderzoek wordt gevonden bij de ontwikkelgroep, en voldoet niet aan de vooraf door ons geformuleerde criteria. We dienen hierbij te bedenken dat Warr een vijf-polig antwoordformaat gebruikte, terwijl wij hier met een ja/nee-formaat hebben gewerkt. Bij de operationalisatie van PSA en werkstress met behulp van vragenlijstschalen in dit proefschrift nemen we, gezien het voorgaande, deze twee begrippen verder niet mee.

Nu we toch stilstaan bij de vertalingen van de schalen van Warr (1990b), willen we ook nog even ingaan op de andere twee schalen die hier zijn onderzocht: "herstelbehoefte" en "emotionele reacties". De "herstelbehoefte"-schaal van Warr

(1990b) heeft vier items en een vijf-polig antwoordformaat. Warr rapporteert voor deze schaal een Cronbach's Alpha van .78. In dit onderzoek vinden we een Rho(t) van .87, bij een schaal bestaande uit elf ja/nee-items. De hier geconstrueerde schaal voldoet additioneel aan de eisen van het Mokken-model.

Voor de schaal "emotionele reacties" ligt een vergelijking met Warr (1990b) wat gecompliceerder. Bij een Mokkenanalyse op de twaalf items van deze schaal vonden we dat de schaal als unidimensioneel beschouwd kan worden. De H(t) was .43 in de ontwikkelgroep, en de Rho(t) .89. Warr (1990b) splitst deze schaal echter in twee subschalen van zes items, te weten: "gespannen-rustig" (Cronbach's Alpha: .76) en "depressief-enthousiast" (Cronbach's Alpha: .80). Studies van Sevastos et al. (1992) en Schalk et al. (1995) met behulp van confirmatorische factoranalyse, komen eveneens tot een twee-factor structuur op basis van dezelfde twaalf items. Het vervelende is echter dat de gevonden twee factoren niet in elk van deze drie oplossingen dezelfde items bevatten. Schalk et al. (1995) concluderen dan ook dat meerdere scoringsmogelijkheden zinvol kunnen zijn, afhankelijk van de onderzoeksdoelstelling. Omdat de genoemde confirmatorische factoranalyses op de "emotionele reacties"-schaal niet tot een eenduidige twee-factor oplossing hebben geleid, en omdat de Mokkenanalyse wijst in de richting van een unidimensionele schaal, handhaven we in dit proefschrift verder de scoringswijze waarin slechts één totaalscore wordt berekend op basis van de twaalf items in de schaal "emotionele reacties".

6.4.7 Stap 7: Cronbach's Alpha

Tot nu toe zijn de kwaliteiten van de schalen weergegeven in parameters behorend bij het Mokken-model. Aangezien deze techniek nog niet zo is ingeburgerd, geven we in tabel 6.6 de Cronbach's alpha (een maat voor interne consistentie volgens de klassieke testtheorie) weer voor de ontwikkelgroep en de testgroep.

Ook bij de Cronbach's alpha zien we dat de testgroep iets minder hoge waarden kent dan de ontwikkelgroep. Bij één schaal vallen de alpha's tegen, namelijk bij "opzien tegen het werk". Dit is ook de enige schaal waarvoor de alpha bij de ontwikkelgroep onder de .80, en bij de testgroep onder de .70 belandt. Bij de overige schalen is er een aantal dat noch in de testgroep, noch in de ontwikkelgroep het criterium van een alpha boven de .80 haalt, namelijk: "rolconflict", "rolonduidelijkheid", "sociaal isolement werkplek", "communicatie", "betrokkenheid bij de organisatie", en "piekeren". Dit zijn kennelijk iets minder intern consistente schalen. Het zijn veelal ook wat kortere schalen, wat te verwachten was gezien het effect van schaalverkorting op de interne consistentie. Met waarden die boven de .70 blijven, zijn de alpha's van deze schalen echter ook niet echt slecht te noemen.

Tabel 6.6: Cronbach's alpha ontwikkelgroep en testgroep

	Ontwikkelgroep		Testgroep	
Schaal	N	Alpha	N	Alpha
werktempo/hoeveelheid	303	.88	457	.88
taakvariatie	310	.82	472	.80
taakautonomie	305	.90	470	.88
mentale belasting	310	.87	471	.83
emotionele belasting	306	.83	470	.81
taakinformatie	310	.81	471	.81
taakverandering	308	.84	470	.84
rolconflict	304	.76	469	.71
rolonduidelijkheid	311	.79	472	.74
relatie collega's	302	.86	460	.87
relatie directe leiding	305	.88	465	.88
sociaal isolement	314	.76	482	.74
inspraak	307	.85	472	.83
communicatie	308	.76	475	.75
leer-en loopbaanmog.	303	.87	465	.85
toekomstonzekerheid	314	.94	474	.93
beloning	307	.80	470	.82
opzien tegen het werk	309	.77	471	.67
betrokkenheid bij org.	275	.76	374	.74
verloopintentie	293	.82	395	.81
herstelbehoefte	307	.87	464	.87
piekeren	313	.74	469	.70
emotionele reacties	305	.89	475	.87

Verder is uit de bovenstaande tabel af te lezen dat er 17 schalen zijn geconstrueerd met een betrouwbaarheid van .80 of meer in zowel de testgroep als de ontwikkelgroep. Dit is, ook naar de maatstaven van de klassieke testtheorie, een goed resultaat.

6.5 Externe test van de schaalkwaliteiten

Door Weel en Broersen (1995) is onderzoek gedaan naar de schaalkwaliteiten van een deel van de geconstrueerde schalen voor het meten van PSA en werkstress bij

een landelijke onderzoekspopulatie. De keuze van schalen werd gemaakt door een groep van deskundigen werkzaam bij arbodiensten, die bij het project betrokken was. De keuzes werden gemaakt op basis van een aantal criteria: het belang van het onderwerp, de schaal kwaliteiten zoals gerapporteerd in Van Veldhoven en Meijman (1994), en de toepasbaarheid van de schaal bij in principe alle typen organisaties en functies. Uiteindelijk werden 15 schalen geselecteerd.

De selectie van schalen werd beantwoord door 601 werknemers afkomstig uit elf verschillende organisaties, verspreid over het hele land. In totaal werkten negen Arbodiensten mee aan de verzameling van deze gegevens.

De betrokken organisaties liepen qua type erg uiteen: van dienstverlenend tot productie, van gezondheidszorg tot kantoor. Van de totale onderzoeksgroep bestond 38% uit vrouwen en 34% was 40 jaar of ouder. De groep bestond voor 57% uit arbeiders en lagere employés. 49% van de groep had een opleiding van maximaal voortgezet onderwijs, lagere trap. Tenslotte: 49% van de groep werkte in een ander dienstrooster dan dagdienst. De samenstelling van deze onderzoeksgroep is met andere woorden uitermate heterogeen, en dus ideaal geschikt voor een test van de schaal kwaliteiten. De N per schaal fluctueert tussen 561 en 593.

In tabel 6.7 worden de $H(t)$ en $Rho(t)$ gepresenteerd voor de 15 geselecteerde schalen. Ter vergelijking zijn dezelfde gegevens van de ontwikkelgroep opgenomen in de tabel.

Wanneer we de in dit hoofdstuk gehanteerde criteria toepassen op het door Weel en Broersen (1995) gevonden resultaat, dan zien we dat de $H(t)$ bij alle schalen voldoet aan de ondergrens van minimaal .40. De meeste schalen zitten hier zelfs wat boven. De $Rho(t)$ komt - op een uitzondering na - op .80 of hoger. Het verschil tussen de resultaten van Weel en Broersen (1995) en die van de ontwikkelgroep is over het algemeen klein.

De uitzondering op de regel is de schaal "betrokkenheid bij de organisatie", die met een $Rho(t)$ van .72 altijd nog redelijk acceptabel is. Een mogelijke verklaring waarom juist deze schaal het minder goed doet is de volgende. De items in deze schaal hebben betrekking op het organisatie-niveau. Peeters en Meijer (1995) maken een onderscheid naar betrokkenheid bij het werk, bij de afdeling, en bij de organisatie. Zij laten zien dat werknemers zich doorgaans meer betrokken voelen bij hun eigen werk en afdeling dan bij de organisatie. Wellicht dat de relatief lage betrouwbaarheid van de betrokkenheid-schaal ontstaat doordat in de hier geselecteerde onderzoekspopulaties bedrijven en instellingen zitten waarbij het organisatie-niveau op wisselende afstand van de werknemers staat.

De conclusie uit deze gegevens is dat de ontwikkelde schalen voor het meten van PSA en werkstress een goede schaalbaarheid en betrouwbaarheid hebben in een onafhankelijke test van de schaal kwaliteiten. Dit is een goed resultaat, zeker wanneer we bedenken dat de gegevens van Weel en Broersen (1995) zijn verzameld in totaal andere regio's van het land, en bij een onderzoeksgroep die qua samenstelling heterogener is dan de ontwikkelgroep.

Tabel 6.7: Resultaten externe test schaalkwaliteiten

	Weel en Broersen (1995)		Ontwikkelgroep (1993)	
Schaal	H(t)	Rho(t)	H(t)	Rho(t)
werktempo/hoeveelheid	.46	.88	.48	.89
taakvariatie	.40	.80	.42	.82
taakautonomie	.41	.88	.44	.90
taakinformatie	.48	.87	.41	.83
taakverandering	.51	.81	.62	.86
rolonduidelijkheid	.52	.82	.45	.81
relatie collega's	.45	.83	.52	.87
relatie directe leiding	.57	.90	.58	.90
inspraak	.47	.86	.42	.85
toekomstonzekerheid	.73	.95	.75	.95
beloning	.46	.80	.45	.80
opzien tegen het werk	.50	.80	.50	.79
betrokkenheid bij organisatie	.42	.72	.54	.79
herstelbehoefte	.46	.86	.48	.87
piekeren	.76	.82	.67	.80

Bovendien was een belangrijk verschil dat in het schaalconstructie-onderzoek is gewerkt met complete bedrijfspopulaties, terwijl bij het onderzoek van Weel en Broersen (1995) is gewerkt met probleemafdelingen. In dit onderzoek werden de deelnemende Arbodiensten namelijk verzocht een proefgroep van circa 50 medewerkers aan te leveren. Een dergelijke opzet leidt tot selectiviteit wat betreft de aangeleverde onderzoeksgroepen. Dit wordt bevestigd doordat de gemiddelde schaalscores in het onderzoek van Weel en Broersen (1995) boven het niveau van de ontwikkelgroep liggen. Dat desondanks de schaalkwaliteiten vergelijkbaar zijn, is een goed, additioneel argument voor de kwaliteit van de geconstrueerde schalen.

6.6 Exploratieve factoranalyse op de schaalscores

Tot nu toe zijn de schalen telkens ingedeeld in blokken, die zijn ontleend aan de facetanalyse zoals beschreven in hoofdstuk 5. De indeling van de schalen in deze

blokken is tot nu toe slechts onderbouwd met kwalitatieve gegevens.

We zullen in deze paragraaf een exploratieve factoranalyse uitvoeren op de schaalscores. Dit kan ons op ideeën brengen over welke dimensies we kunnen onderscheiden in de verzameling van geconstrueerde schalen. Een meer uitvoerige analyse over de dimensionalisering van de schalen voor het meten van PSA en werkstress, met behulp van confirmatorische factoranalyse, volgt in hoofdstuk 7.

6.6.1 Additionele gegevens en aangepaste scoring

Alvorens te beginnen met de exploratieve factoranalyse zijn de gegevens van een vierde bedrijf toegevoegd aan het gegevensbestand. Het totaal aantal werknemers waarvan vragenlijsten beschikbaar zijn bedraagt hierdoor 1004. Hieronder volgen enkele kenmerken van bedrijf 4, zoals we die ook hebben beschreven van de drie andere bedrijven.

Bedrijf 4: gezondheidszorginstelling

Het onderzoek speelde in op een onderzoeksvraag van de kant van de instelling. De respons bedroeg 57%. Alle personen van 45 jaar en ouder werden uitgenodigd voor het onderzoek, evenals alle medewerkers van een zestal afdelingen. De gemiddelde leeftijd van de onderzoeksgroep is, door deze opzet, voor een gezondheidszorginstelling aan de hoge kant: gemiddeld 41,0 jaar (standaarddeviatie: 11,2 jaar). Het overgrote deel van de onderzochte populatie is vrouwelijk: 74%. Bij deze organisatie werd de 'emotionele reacties'-schaal vervangen door de trait-versie van de zelfbeoordelvragenlijst (Van der Ploeg, Defares, en Spielberger, 1980). Verder was de vragenlijst identiek aan de itempool die in de eerdere onderzoeken werd afgenomen.

Voor de exploratieve factoranalyse is in de scoring van de "gevolgen"-schalen een aantal wijzigingen aangebracht, zodat alle "gevolgen"-schalen een vergelijkbare score-range hebben.

Verder zijn in de exploratieve factoranalyse een tweetal scores toegevoegd die betrekking hebben op actuele gezondheidsklachten. Deze gegevens waren beschikbaar in de itempool. In de eerste plaats is er een score toegevoegd voor psychische klachten op basis van acht dichotome items (Cronbach's $\alpha = .66$). In de tweede plaats een vergelijkbare score voor klachten van het houding- en bewegingsapparaat (HBA-klachten; Cronbach's $\alpha = .59$). Alhoewel deze twee schalen niet voldoen aan de betrouwbaarheidseisen zoals oorspronkelijk gesteld, nemen we ze toch mee in de analyse.

Voor de exploratieve factoranalyse zijn de schalen in twee groepen gesplitst. De eerste groep bevat alle schalen uit de blokken "taak" en "context van de taak".

Van nu af aan zullen we deze schalen collectief benoemen als schalen voor het meten van PSA. Het tweede deel bevat alle "gevolgen"-schalen. Hiernaar zullen we in het vervolg collectief verwijzen als schalen voor het meten van werkstress.

De redenen om dit te splitsen waren tweeledig. In de eerste plaats verschillen deze twee groepen schalen wat betreft antwoordmogelijkheden (altijd/vaak/soms/nooit versus ja/nee). In de tweede plaats bleken de correlaties tussen PSA-schalen en werkstress-schalen in paragraaf 6.4.4 over het algemeen van een dusdanig laag niveau te zijn, dat een grote mate van redundantie tussen deze twee groepen schalen niet valt te verwachten.

6.6.2 Exploratieve factoranalyse PSA

De 17 schaalscores met betrekking tot PSA zijn beschikbaar voor 866 werknemers. Hierop is een principale componentenanalyse uitgevoerd met varimax-rotatie.

Op grond van de criteria "eigenwaarde groter dan 1" en "bijdrage aan de verklaarde variantie" zijn drie factoren te onderscheiden, die gezamenlijk 46,8% van de variantie verklaren. De eerste factor verklaart 24,1%, de tweede 15,6%, en de derde 7,1%. De factorladingen (na rotatie) van de schaalscores op deze drie factoren worden weergegeven in tabel 6.8. De factor-ladingen worden alleen weergegeven als ze $>.35$ of $<-.35$ zijn.

Uit de tabel blijkt dat de eerste factor wordt bepaald door (in volgorde van factorlading): "gebrek steun leiding", "rolonduidelijkheid", "gebrek steun collega's", "rolconflict", "gebrek communicatie", "gebrek taakinformatie", "taakverandering", en "gebrek aan inspraak". Deze factor lijkt vooral te maken te hebben met sociaal-organisatorische aspecten.

Op de tweede factor laden de volgende schalen (in volgorde van factorlading): "mentale belasting", "emotionele belasting", "gebrek taakvariatie", "werkdruk", en "gebrek leer- en loopbaanmogelijkheden". Deze schalen hebben allen te maken met psychologische taakeisen.

Tenslotte de derde factor. Hierop laden hoog: "gebrek taakautonomie", "gebrek inspraak", en "sociaal isolement". Deze dimensie kunnen we aanduiden met de term regelmogelijkheden.

De factoranalyse bij de PSA-schalen laat een structuur zien die goed aansluit bij het herziene demand-control model van Karasek en Theorell (1990). Er kan echter wel een aantal kanttekeningen worden geplaatst bij dit resultaat.

Tabel 6.8: Resultaten exploratieve factoranalyse PSA

Schaal	Factor 1	Factor 2	Factor 3
werktempo/hoeveelheid		.61	
taakvariatie		-.67	
gebrek taakautonomie			.75
mentale belasting		.74	
emotionele belasting		.67	
gebrek taakinformatie	.54		
taakverandering	.53		
rolconflict	.70		
rolonduidelijkheid	.75		
gebrek steun collega's	.71		
gebrek steun leiding	.76		
sociaal isolement			.58
gebrek inspraak	.38		.68
gebrek communicatie	.58		
gebrek leermogelijkheden		-.52	
toekomstonzekerheid			
beloning			

In de eerste plaats valt een tweetal schalen buiten de drie gevonden factoren. Het gaat hier om de schalen "toekomstonzekerheid" en "beloning". Zij vormen samen ook geen duidelijke vierde factor meer. Dit betekent dat aspecten op het gebied van "arbeidsvoorwaarden" niet clusteren met aspecten op het gebied van "arbeidsinhoud" en "arbeidsorganisatie", en dat ze niet clusteren met elkaar. Dit is enerzijds een resultaat dat de validiteit van de hier gevonden drie dimensies ondersteunt, anderzijds levert dit de volgende vraag op voor het derde deel van dit proefschrift: nemen we "arbeidsvoorwaarden" hierin mee of niet?

Opvallend is ook, dat de schaal leer- en loopbaanmogelijkheden clustert onder de factor "psychologische taakeisen". Deze schaal werd verondersteld een aspect van de "arbeidsvoorwaarden" te meten. Ook voor deze schaal moet worden bekeken hoe we hier in het derde deel van het proefschrift mee omgaan.

Verder is er een schaal die $>.35$ of $<-.35$ laadt op meer dan één factor, namelijk: "gebrek aan inspraak". Deze schaal laadt .38 op de factor "sociaal-organisatorische aspecten" en .68 op de factor "regelmogelijkheden". Dit verschil in factorlading is echter ruim voldoende aanleiding om deze schaal te plaatsen onder "regelmogelijkheden".

Tenslotte een opmerking over de schaal "sociaal isolement". Dat deze schaal clustert onder "regelmogelijkheden" werkt wellicht bevreemdend. Door het label te veranderen in "gebrek aan contactmogelijkheden" wordt de plaatsing onder "regelmogelijkheden" qua naamgeving logischer. Het hebben van contactmogelijkheden biedt immers regelmogelijkheden, zoals het kunnen verlaten van de werkplek en het afleiding kunnen zoeken in een gesprek.

6.6.3 Exploratieve factoranalyse werkstress

Op eenzelfde manier als voor de PSA-schalen is ook een principale componenten-analyse gedaan op de werkstress-schalen. Hierbij is de beschikbare N aanzienlijk kleiner omdat bij bedrijf 1 de schalen "betrokkenheid" en "verloopintentie" niet werden afgenomen, en bij bedrijf 4 de schaal "emotionele reacties" niet werd afgenomen.

In totaal zijn van 641 personen alle werkstress-schalen beschikbaar. Op deze acht schaalscores blijken twee factoren te worden gevonden die beide voldoen aan het criterium "eigenwaarde groter dan 1". Gezamenlijk verklaren ze 56,2% van de variantie, waarvan 37% door factor 1 en 19,2% door factor 2. De factorladingen (na rotatie) voor de werkstress-schalen met een waarde van $>.35$ of $<-.35$ zijn opgenomen in tabel 6.9.

Tabel 6.9: Resultaten exploratieve factoranalyse werkstress

Schaal	Factor 1	Factor 2
opzien tegen het werk		.78
herstelbehoefte	.77	
piekeren	.74	
gebrek betrokkenheid		.81
verloopintentie		.81
emotionele reacties	.68	
HBA-klachten	.43	
psychische klachten	.78	

De schaalscores blijken zich netjes te verdelen over de twee factoren. De eerste factor wordt gevormd door: "herstelbehoefte", "piekeren", "emotionele reacties", "HBA-klachten", en "psychische klachten". Als we de gezondheidsklachten-schalen hieruit weglaten, dan houden we een dimensie over, die we willen aanduiden met de term arbeidsgerelateerde spanningsverschijnselen.

De tweede factor bevat de schalen "opzien tegen het werk", "gebrek aan betrokkenheid", en "verloopintentie". Deze dimensie kan worden aangeduid met meer traditionele termen als "arbeidssatisfactie", "arbeidsmotivatie", of "welbevinden in het werk". Naar analogie met de term arbeidsgerelateerde spanningsverschijnselen willen we deze factor arbeidsgerelateerd welbevinden noemen.

Bij de resultaten van de exploratieve factoranalyse op de werkstress-schalen valt eigenlijk weinig op te merken, behalve dan dat de factorlading van de HBA-klachtenscore op de eerste factor wat aan de lage kant is.

6.7 Slot deel II

We hebben nu een zelfbeoordelingsinstrument in handen waarmee we PSA en werkstress goed kunnen operationaliseren, althans volgens gangbare psychometrische criteria. Op het gebied van PSA hebben we 17 schalen ontwikkeld: vijf over "psychologische taakeisen", zeven over "sociaal-organisatorische aspecten", drie over "regelmogelijkheden", en tenslotte twee over "arbeidsvoorwaarden". Ten aanzien van werkstress beschikken we over zes schalen: drie over "arbeidsgerelateerde spanningsverschijnselen", en drie over "arbeidsgerelateerd welbevinden". De ontwikkelde schalen zijn redelijk onafhankelijk van elkaar, met uitzondering van de schalen over "sociaal-organisatorische aspecten". Deze vertonen hoge onderlinge correlaties.

In het kader van dit proefschrift betekent dit, dat we nu de aandacht kunnen richten op de vraagstellingen, die we in deel I hebben geformuleerd. Aan het eind van hoofdstuk 4 zijn deze vraagstellingen tegen de achtergrond van de literatuur beschreven. We roepen ze hier kort in herinnering:

- wat is de structuur van PSA en werkstress?
- hoe is de relatie tussen PSA en werkstress?
- wat is de impact van PSA en werkstress op ziekteverzuim?

Tussen de regels door, keert bij de beantwoording van deze vragen in deel III telkens de vraag terug naar de (meer-)waarde van een vragenlijstmethode voor onderzoek naar PSA en werkstress.

Deel III: Structuur en betekenis van psychosociale arbeidsbelasting en werkstress

Hoofdstuk 7: Dimensies in PSA en werkstress

In het vorige hoofdstuk is een exploratieve factoranalyse uitgevoerd om te onderzoeken welke dimensies zijn aan te geven in de geconstrueerde schalen. We vonden een structuur bestaande uit vijf dimensies: drie voor PSA en twee voor werkstress.

Omdat we nu een verzameling ter beschikking hebben van psychometrisch goede schalen, en omdat we een idee hebben van de achterliggende dimensies, kunnen we meer fundamentele vragen gaan beantwoorden over de structuur van PSA en werkstress, zoals geoperationaliseerd met een zelfbeoordelingsmethode. In psychometrisch jargon: dit hoofdstuk gaat over de convergente en divergente validiteit van zelfbeoordelingsschalen voor het meten van PSA en werkstress. Hoe abstract dit ook moge klinken, het belang hiervan is heel concreet. De analyses in dit hoofdstuk geven antwoorden op vragen als:

1. Slaan de afzonderlijke zelfbeoordelingsschalen voor het meten van PSA en werkstress niet allemaal op hetzelfde, bijvoorbeeld op een algemene onvrede bij de werknemers?
2. Hoeveel zelfbeoordelingsschalen hebben we nodig om PSA en werkstress te meten? Wanneer wordt een vragenlijst op dit gebied onnodig lang?
3. Zijn de zelfbeoordelingsschalen voor PSA en werkstress te gebruiken voor toetsing van theoretische modellen? Zo ja: welke schalen dekken dan welke theoretische begrippen? Wat dit laatste punt betreft zijn we vooral op zoek naar een goede operationalisatie van het "decision latitude"-begrip uit de theorie van Karasek (1979).

De eerste vraag is een standaard dooddoener bij menige nabespreking van een werkbelevingsonderzoek. Als differentiatie in schalen alleen maar dient als algemene graadmeter voor onvrede, waarom dan al de moeite? Achter deze, ogenschijnlijk methodologische, kritiek schuilt in de praktijk vaak een onvermogen c.q. onwil om werkstressproblematiek aan te pakken. Het maakt immers verschil of we het hebben over een vage algemene onvrede, of dat we het hebben over een klinkklaar signaal van absurde werkdruk, uit de hand lopende conflicten, of totale afwezigheid van autonomie. Zolang iets vaag is, hoeft er niets te gebeuren.

De tweede vraag is vooral van praktisch belang, omdat werknemers tegenwoordig veelvuldig geplaagd worden met vragenlijsten en enquêtes. Iedere vraag teveel werkt vragenlijstmoeheid en slechte respons in de hand.

De derde vraag is van belang voor zowel wetenschap als praktijk. Een antwoord hierop geeft enerzijds aan of een zelfbeoordelingsmethode van wetenschappelijke waarde kan zijn. Anderzijds kunnen we uit het antwoord op deze vraag afleiden welke keuze van schalen hoort bij welk type vraagstelling. Alleen op deze manier kunnen we komen tot enige vorm van nader onderzoek naar specifieke vormen van PSA en werkstress.

Een conceptueel heldere operationalisatie van "decision latitude" (zie 2.5) tenslotte, is voor theorie en praktijk van belang. Veel onderling tegenstrijdige resultaten, die zijn verkregen met het demand-control model kunnen wellicht worden verklaard als deze kwestie is opgelost.

7.1 Voorbereiding van de gegevens

Voor de analyses in dit hoofdstuk zijn de gegevens van twee additionele bedrijven beschikbaar. Het gaat om twee productiebedrijven. Opnieuw werd door de Arbodienst ingespeeld op onderzoeksvragen van de kant van de bedrijven. In het ene bedrijf werd de totale populatie uitgenodigd, waarbij de respons 96% was. In het tweede bedrijf werd alleen het management uitgenodigd tot deelname. Hier was de respons 82%. De gemiddelde leeftijd van de deelnemersgroepen was respectievelijk 38,9 jaar (standaarddeviatie: 10,7 jaar) en 45,7 jaar (standaarddeviatie: 7,7 jaar). Bij beide groepen gaat het hoofdzakelijk om mannen: respectievelijk 74% en 99%. Waar de medewerkers van de voorgaande vier bedrijven de uitgebreide itempool invulden, kregen de medewerkers van deze twee productiebedrijven de lijst voorgelegd in de verkorte vorm. Door toevoeging van deze twee bedrijven komt het totaal aantal werknemers op 1306, waarbij de groep een stuk minder onevenwichtig is geworden in samenstelling qua leeftijd, geslacht, en type bedrijf. De enige groep die nu nog relatief ondervertegenwoordigd is in het cumulatieve gegevensbestand, is die van de traditionele productie-arbeid. Als volgende stap werden alleen die werknemers geselecteerd waarvan bij alle schalen op het gebied van PSA en werkstress een score beschikbaar was. Doordat enkele schalen systematisch ontbreken bij twee bedrijven, vallen de respondenten van deze bedrijven af voor de hier beoogde analyse. Deze selectie leverde een bestand op van 816 werknemers, afkomstig uit vier bedrijven (een grote steekproef uit een gezondheidszorginstelling en drie kleinere steekproeven uit drie productiebedrijven). Alle variabelen werden vervolgens gestandaardiseerd, waarna met de analyses werd begonnen.

Op grond van de exploratieve factoranalyse (zie 6.6) werd besloten om de schalen met betrekking tot "arbeidsvoorwaarden" buiten de hiernavolgende analyse te houden. Het vervolg van dit proefschrift beperkt zich verder dan ook tot PSA, voor zover dit betrekking heeft op de inhoud en organisatie van de arbeid.

De keuze om "arbeidsvoorwaarden" buiten de analyse te houden brengt voor nog een schaal problemen met zich mee, namelijk "leer- en loopbaanmogelijkheden". "Loopbaanmogelijkheden" gelden vooral als een arbeidsvoorwaardelijk aspect, "leermogelijkheden" vooral als een kenmerk dat al dan niet aanwezig is in een bepaalde arbeidstaak. Uiteraard zijn er banen te noemen waarin het ontbreken van het één gepaard gaat met het ontbreken van het ander. Dit is echter niet altijd zo: een baan kan duidelijke leermogelijkheden bevatten terwijl de loopbaanmogelijkheden (doorgroei-mogelijkheden, salaris-ontwikkeling) beperkt zijn. De koppeling van leer- en loopbaanmogelijkheden in één schaal, is achteraf gezien niet zo gelukkig gekozen. Dit ondanks dat de Mokkenanalyse heeft aangetoond dat een samenvoeging van leer- en loopbaan-items wel degelijk een goede, unidimensionele schaal oplevert.

Het is vooral het aspect "leermogelijkheden" dat we nodig hebben voor het vervolg. We weten uit het vorige hoofdstuk dat de schaal "leer- en loopbaanmogelijkheden" .58 correleert met "taakvariatie". Een interpretatie van schalen als "leer- en loopbaanmogelijkheden" en "taakvariatie" in de richting van theoretische begrippen als "skill discretion" en "skill utilization" (Karasek en Theorell, 1990) ligt dan ook voor de hand. Het zijn juist dergelijke, meer taakinhoudelijke items binnen de schaal "leer- en loopbaanmogelijkheden" die we in het vervolg mee willen nemen.

De items van de schalen "taakvariatie" en "leer- en loopbaanmogelijkheden" zijn gezamenlijk in een Mokken-searchprocedure ingevoerd. Dit leidt bij de ontwikkelgroep tot een eerste schaal waarin alle "taakvariatie"-items vallen plus vier items uit de schaal "leer- en loopbaanmogelijkheden". Dit is het deel "leermogelijkheden". Vier andere items uit de schaal "leer- en loopbaanmogelijkheden" vormen samen een tweede schaal: "loopbaanmogelijkheden". Eén item is niet plaatsbaar in één van beide schalen.

Een afzonderlijke Mokken-testprocedure voor de leermogelijkheden-items bij de ontwikkelgroep levert een $H(t)$ op van .57, en een $Rho(t)$ van .84. Voor de "loopbaanmogelijkheden"-items wordt een $H(t)$ gevonden van .46 en een $Rho(t)$ van .77. Aldus blijken twee goede schalen gemaakt te kunnen worden. De schaal "leermogelijkheden" nemen we mee als taakinhoudelijk aspect in het vervolg van dit proefschrift.

Aldus houden we 21 schalen over voor de analyses in dit hoofdstuk. Voor zover we dat op voorhand kunnen overzien, heeft het "decision latitude"-begrip raakvlakken met een aantal van deze schalen. In de eerste plaats met "taakvariatie" en "leermogelijkheden" ("skill discretion"), en in de tweede plaats met "taakautonomie", "inspraak" en "contactmogelijkheden" ("task authority").

7.2 De interne consistentie van dimensies in PSA en werkstress

Onze eerstvolgende stap is niet erg gebruikelijk, maar wel voor de hand liggend. We voeren de schaalscores, die uit de exploratieve factoranalyse naar voren kwamen als behorend bij een bepaalde factor, samen in, in een reliability-procedure (tweede-orde analyse). We bekijken dan of dit een hoge interne consistentie van de desbetreffende factor oplevert.

De nieuwe schaal "leermogelijkheden" wordt ingedeeld bij de factor "psychologische taakeisen", net als de voorganger hiervan: "leer- en loopbaanmogelijkheden". Voor deze analyse zijn alle schalen binnen een bepaalde dimensie in één richting gescoord. De schalen "taakvariatie" en "leermogelijkheden" moesten hiervoor worden omgeklapt: een hogere score geeft in dit hoofdstuk meer "taakvariatie" en meer "leermogelijkheden" aan!

In tabel 7.1 staan drie soorten gegevens. In de onderste rij staat de Cronbach's Alpha voor de vijf dimensies. Vetgedrukt staan de item-restcorrelaties van de schalen met de andere schalen binnen hun dimensie. In gewone druk staan de Pearson-correlaties van de afzonderlijke schalen met de somscores van de schalen in de vijf dimensies (dimensie-scores: somscores van alle schalen die in een bepaalde dimensie zijn ingedeeld).

De Cronbach's Alpha varieert van .67 tot .81. De onderscheiden dimensies hebben dus redelijke tot goede interne consistenties. Dit kunnen we zo interpreteren, dat de schalen binnen de veronderstelde dimensies inderdaad in sterke mate onderling verband houden.

Belangrijker is het om per kolom en per rij de cijfers in de tabel te bekijken. Als criterium voor een redelijk correlatie gebruiken we hierbij .40. We bespreken de gegevens hier eerst per dimensie (kolom).

Bij "psychologische taakeisen" correleert "werktempo en werkhoeveelheid" slechts .34 met de overige schalen binnen de dimensie. Dit is laag voor zo'n belangrijk aspect. Er is verder een schaal van een andere dimensie, die hoger correleert met "psychologische taakeisen" dan de schaal "werktempo en -hoeveelheid". "Inspraak" correleert namelijk -.46 met de dimensiescore van "psychologische taakeisen". Bij de dimensie "regelmogelijkheden" treffen we eveneens een schaal aan die een betrekkelijk lage item-rest correlatie heeft: "contactmogelijkheden" (.38). Maar liefst vier schalen uit andere dimensies correleren sterker dan .40 met "regelmogelijkheden". Dit zijn de schalen "taakvariatie" (-.46), "leermogelijkheden" (-.45), "taakinformatie" (.46), en "communicatie" (.45).

Tabel 7.1: Resultaten tweede-orde analyse interne consistentie

Schaal	Dimensie				
	psy. eisen	regelmog.	soc.-org. aspecten	spanning	welbev.
werktempo/hoeveelheid	.34	.02	.22	.42	.04
mentale belasting	.50	-.24	-.08	.14	-.15
emotionele belasting	.54	-.18	.21	.37	.12
taakvariatie	.63	-.46	-.24	.06	-.33
leermogelijkheden	.49	-.45	-.34	-.04	-.34
taakautonomie	-.32	.52	.28	.08	.18
inspraak	-.46	.54	.50	.05	.35
contactmogelijkheden	-.09	.38	.30	.13	.15
rolconflict	.05	.25	.59	.29	.39
rolonduidelijkheden	-.05	.36	.60	.21	.32
relatie collega's	-.03	.30	.54	.31	.34
relatie leiding	-.07	.39	.68	.27	.39
informatie	-.28	.46	.52	.13	.30
communicatie	-.12	.45	.56	.21	.38
verandering	.20	.03	.30	.13	.15
herstelbehoefte	.30	.10	.25	.52	.25
piekeren	.27	.00	.22	.52	.11
emotionele reacties	.07	.16	.30	.47	.31
opzien	-.20	.25	.37	.34	.59
betrokkenheid	-.21	.29	.40	.14	.55
verlooptentatie	-.05	.17	.33	.22	.58
Cronbach's Alpha	.74	.67	.81	.69	.75

Ook bij de "sociaal-organisatorische" dimensie treffen we opnieuw één schaal aan die een lage item-rest correlatie heeft. Deze keer is dat "taakverandering" (.30). Twee schalen van buiten de "sociaal-organisatorische" dimensie correleren hoog, namelijk: "inspraak" (.50) en "betrokkenheid" (.40).

De dimensie "spanning" heeft intern geen schalen met een lage item-rest correlatie,

wel een externe schaal met een correlatie van meer dan .40, namelijk "werktempo en werkhoeveelheid". Bij de "welbevinden"-dimensie zien we noch intern, noch extern opvallende correlaties.

We kunnen ook per schaal (rij) de tabel langslopen, op zoek naar opmerkelijke correlaties. We vragen ons hierbij af: zijn er schalen die een hogere correlatie hebben met een andere dimensie, dan met de dimensie waarbij ze zijn ingedeeld? Er is inderdaad zo'n schaal te vinden, namelijk: "werktempo en werkhoeveelheid". De item-rest correlatie met "psychologische taakeisen" is .34, terwijl de correlatie met de dimensie-score van "spanning" .40 is. Verder valt op dat de schalen uit de "sociaal-organisatorische" dimensie vaak redelijk hoog correleren met de dimensie-scores van "regelmogelijkheden" en "welbevinden".

Uit deze tweede-orde analyse kunnen we een aantal conclusies trekken: Een vijf-factor structuur van PSA en werkstress is redelijk adequaat, maar niet zonder problemen. De samenvoeging van de schaal "werktempo en -hoeveelheid" met een aantal andere schalen onder de dimensie "psychologische taakeisen", is discutabel. De schaal "werktempo en -hoeveelheid" hangt te weinig samen met de overige schalen in de dimensie "psychologische taakeisen", en hangt zelfs meer samen met "spanning" dan met "psychologische taakeisen". Het is onduidelijk of dit een gevolg is van de schaal "werktempo- en hoeveelheid" zelf, danwel van de samenstelling van de rest van de "psychologische taakeisen"-dimensie. De lage item-rest correlatie van de schaal "werktempo en -hoeveelheid" (.34) is vooralsnog geen reden om dit aspect verder buiten beschouwing te laten: daarvoor is het conceptueel en praktisch te belangrijk.

De schaal "inspraak" hangt samen met alle PSA-dimensies. Deze schaal draagt mogelijk bij aan een onheldere dimensie-structuur van PSA. Ook dit aspect kan om inhoudelijke en praktische redenen echter niet worden gemist.

De schalen "contactmogelijkheden" en "taakverandering" hebben een lage samenhang met de andere schalen in de dimensie waarbij ze op grond van de exploratieve factoranalyse zijn ingedeeld. Dit is voldoende reden om deze schalen in het onderstaande verder buiten beschouwing te laten. Het gaat om begrippen die inhoudelijk en pragmatisch zeker relevant zijn, maar die toch minder centraal zijn dan onderwerpen als "werktempo/werkhoeveelheid" en "inspraak".

In hoeverre geeft de tweede-orde analyse van de interne consistentie informatie over de operationalisatie van het begrip "decision latitude"? Het blijkt dat binnen de dimensie "regelmogelijkheden" ("task authority") de schaal "contactmogelijkheden" relatief slecht past. De schalen "taakvariatie" en "leermogelijkheden" ("skill discretion") vallen binnen de problematische "psychologische taakeisen"-dimensie. De tweede-orde analyse van de interne consistentie geeft ons dus geen duidelijk antwoord op de vraag hoe we "decision latitude" kunnen operationaliseren.

7.3 Confirmatorische factoranalyse

In de vorige paragraaf hebben we aan de hand van een tweede-orde analyse van de interne consistentie gezien dat de vijf-factor structuur, zoals die resulteerde uit de exploratieve factoranalyse, een redelijke dimensie-structuur biedt voor PSA en werkstress. Met name de psychologische taakeisen-dimensie leverde echter problemen op.

De analyse in de vorige paragraaf heeft een aantal beperkingen. Uit de analyse kunnen in de eerste plaats weinig zinvolle suggesties voor een betere structuur worden gehaald. Ten tweede is er geen manier om te toetsen hoe goed de totale vijf-factor structuur past bij de gegevens. Tenslotte kunnen we niet handig vergelijken of deze vijf-factor structuur beter is dan een simpeler structuur, bijvoorbeeld een structuur die alles terugbrengt tot één gezamenlijke factor (algemene onvrede).

Samengevat bieden de exploratieve factoranalyse en de tweede orde-analyse van de interne consistentie een eerste indruk van de structuur van PSA en werkstress waarnaar we op zoek zijn in dit hoofdstuk, zonder alle vragen te beantwoorden.

We hebben nu een methode nodig die meer de gelegenheid biedt te toetsen hoe goed onze veronderstelde structuur past bij de onderzoeksgegevens, en die tevens mogelijkheden biedt de structuur op onderdelen aan te scherpen. Dit laatste geldt met name voor het aspect "decision latitude". Een techniek die ons deze mogelijkheden biedt is confirmatorische factoranalyse.

Om een dergelijke analyse te kunnen uitvoeren zijn de gegevens van de 816 werknemers, zoals gebruikt voor de analyse in de vorige paragraaf, geconverteerd naar PRELIS/LISREL (Jöreskog en Sörbom, 1993). Hoewel de meeste schalen enigszins niet-normaal verdeeld zijn (vooral de werkstress-schalen zijn nogal scheef) en hoewel de antwoordrange van sommige schalen beperkt is tot minder dan 15 categorieën, zijn alle variabelen in PRELIS behandeld als continue variabelen. Op basis van de schalen is een gewone correlatie-matrix aangemaakt voor analyse in LISREL. Dit heeft als voordeel dat de resultaten van de tweede orde-analyse van de interne consistentie en de confirmatorische factoranalyse vergelijkbaar uitgangsmateriaal hebben.

In plaats van één matrix te maken voor de totale groep van 816 werknemers, is een LISREL-ontwikkelgroep en een LISREL-testgroep gemaakt door random steekproeven uit de totale groep te trekken. De LISREL-ontwikkelgroep bevat 409 werknemers, de LISREL-testgroep 407. Het onderscheiden van deze twee groepen biedt de mogelijkheid een structuur voor PSA en werkstress te ontwerpen bij de LISREL-ontwikkelgroep, en deze structuur vervolgens onafhankelijk daarvan te kruisvalideren in de LISREL-testgroep.

Voor de duidelijkheid geven we hier nog een keer aan dat deze confirmatorische factoranalyse zich beperkt tot 19 schalen die betrekking hebben op PSA (arbeidsinhoud en arbeidsorganisatie, geen arbeidsvoorwaarden) en werkstress. De schalen "contactmogelijkheden" en "taakverandering" zijn naar aanleiding van de vorige paragraaf in het vervolg weggelaten.

Het basismodel voor de LISREL-analyse is afgeleid van de exploratieve factoranalyse in het vorige hoofdstuk. We verwachten in eerste instantie dezelfde problemen tegen te komen als in de vorige paragraaf bij de tweede-orde analyse van de interne consistentie.

LISREL geeft in de vorm van passingsmaten aan, of de indeling van schalen over de veronderstelde dimensies (hierna te noemen: het model) juist is. Tevens worden de ladingen van de schalen op hun onderliggende dimensies geschat, en worden (in de vorm van modificatie-indexen) hints gegeven voor verbetering van het model.

Van een goed passend model is sprake wanneer de overschrijdingskans van de Chi-kwadraatverdeelde passingsmaat groter is dan .05. Aangezien deze grootheid erg gevoelig is voor de steekproefgrootte, en daardoor gauw tot verwerping van het model leidt, bestaan er alternatieve passingsmaten waarmee de passing van een model beoordeeld kan worden. Een alternatieve passingsmaat is de Adjusted Goodness of Fit Index (AGFI), die een waarde tussen 0 (slechte passing) en 1 (perfecte passing) kan aannemen. Een andere maat is de Root Mean Square Residual (RMSR). Ervaring leert dat sprake is van een goed passend model bij een AGFI groter dan .90 en, na standaardisatie, een RMSR kleiner dan 0.05 (Jöreskog en Sörbom, 1993).

Tevens zal gebruik worden gemaakt van een maat voor de relatieve passing van een model ten opzichte van een ander (genest) model: de Tucker-Lewis Index (TLI) (Tucker en Lewis, 1973). Dit is een "type-2 incremental fit"-index, die onafhankelijk is van de steekproefgrootte, en die wordt aanbevolen door Marsh et al. (1988). Bentler en Bonett (1980) stellen dat een waarde van een "incremental fit"-index onder de .90 aangeeft dat het model waarschijnlijk nog verbeterd kan worden.

In de onderstaande analyses zal de verbetering van de passing van een model telkens worden beoordeeld ten opzichte van een nulmodel, dat ervan uitgaat dat alle schalen zijn terug te brengen tot één onderliggende dimensie (algemene onvrede). De TLI geeft aan in welke mate een complex model beter past bij de gegevens dan een dergelijk nulmodel.

We dienen te bedenken dat de keuze voor dit type nulmodel anders is dan gebruikelijk. Vaak wordt als nulmodel de afwezigheid van structuur genomen. Dit zou in dit geval corresponderen met een 19-factor model. Wij nemen een één-factor model als nulmodel. De TLI geeft in onze analyse hierdoor niet aan, of een model beter is dan géén model, maar of het veronderstelde model beter is dan een zo simpel mogelijk, maar wel betekenisvol model. Dit is, vanuit de praktijk geredeneerd, een zinvollere vraag om te beantwoorden. We geven hiermee rechtstreeks antwoord aan diegenen die sceptisch zijn ten aanzien van werkbelevingsonderzoek. Als het hier te ontwikkelen model een betere passing heeft dan een "algemene onvrede"-model, dan mogen we de resultaten van een werkbelevingsonderzoek niet meer opvatten als algemeen klaaggedrag. Hogere scores op specifieke PSA-factoren enerzijds en specifieke werkstress-symptomen anderzijds, moeten in dit geval geïnterpreteerd worden als specifieke aanduidingen van het soort probleem. Het kunnen onderscheiden van specifieke oorzaken en gevolgen is een voorwaarde om te kunnen komen tot een gerichte aanpak. Zolang dit niet is aangetoond, is elke aanpak in principe geschikt, als de werknemers maar ophouden met klagen.

Bij de confirmatorische factoranalyse wordt aangenomen dat de latente begrippen vrij met elkaar zijn gecorreleerd. Deze assumptie past bij de empirie. Tussen de schalen die de latente begrippen afdekken zijn immers op voorhand de nodige correlaties bekend (zie paragraaf 6.4.4).

Als eerste model dat we gaan onderzoeken, nemen we een model dat rechtstreeks is afgeleid van de exploratieve factoranalyse in hoofdstuk 6. De 19 schalen zijn hierbij ingedeeld in dezelfde vijf dimensies, als in paragraaf 7.2 (model 1). In de tabel 7.2 presenteren we de gegevens van het nulmodel (alle 19 schalen zijn terug te brengen tot één dimensie) en model 1, voor de LISREL-ontwikkelgroep en de LISREL-testgroep afzonderlijk. De specificaties van het nulmodel en model 1 zijn te vinden in bijlage 1.1 en 1.2.

Tabel 7.2 Confirmatorische factoranalyse: passingsmaten nulmodel/model 1

Model	N	Chi	Df	Sign.	AGFI	RMSR	TLI
nulmodel/ontwikkelgroep	409	1482.56	152	.00	.55	.14	NVT
nulmodel/testgroep	407	1522.96	152	.00	.51	.15	NVT
model 1/ontwikkelgroep	409	659.76	142	.00	.80	.11	.58
model 1/testgroep	407	730.19	142	.00	.77	.12	.54

Uit tabel 7.2 worden drie dingen duidelijk. Ten eerste valt op dat het nulmodel niet zo heel erg slecht is. Het model dat alle schalen herleidt tot één gezamenlijke dimensie leidt tot een AGFI die halverwege de nul en de één ligt. Dit heeft zijn consequenties voor de hierna te berekenen TLI: hoe beter het nulmodel, hoe sterker het alternatieve model moet zijn om een hoge TLI te behalen.

In de tweede plaats valt op dat model 1, ondanks dat het een flinke verbetering inhoudt ten opzichte van het nulmodel, niet voldoet aan de criteria voor een goed model (Significantie $>.05$, AGFI $>.90$, RMSR $<.05$).

Ten derde valt op dat de twee modellen ongeveer dezelfde waarden opleveren bij de LISREL-ontwikkelgroep en de LISREL-testgroep, wat erop duidt dat de modellen redelijk stabiel zijn.

7.4 Verbetering van het model op de LISREL-ontwikkelgroep: "zonder weglating van schalen"

In eerste instantie gaan we proberen zonder verdere weglating van schalen tot een beter passend model te komen. Hiertoe gaan we drie criteria gebruiken: onderscheidbaarheid, specificiteit, en informatiewaarde (Oort-Marburger et al., 1993). We gebruiken de LISREL-ontwikkelgroep om verbeteringen in het model te ontdekken. Het resulterende model toetsen we straks in de LISREL-testgroep.

7.4.1 Onderscheidbaarheid

Dit criterium heeft betrekking op de volgende vraag: hangt een schaal die is ingedeeld bij dimensie X sterker samen met dimensie Y dan met dimensie X? We kunnen dit onderzoeken door de modificatie-indexen te bekijken in de LISREL-output voor andere paden dan die welke in het model zijn gespecificeerd. Een modificatie-index gedraagt zich als een Chi-kwadraat verdeelde parameter met één vrijheidsgraad. LISREL signaleert default alleen modificatie-indexen met een waarde groter dan 7.882, wat correspondeert met een éézijdig significantie-niveau van .005 (Jöreskog en Sörbom, 1993). Aangezien de hier gevonden modificatie-indexen veel groter zijn dan deze drempel, gaan we hier niet verder op de significantie van de modificatie-indexen in. We kijken alleen naar de absolute hoogte van de modificatie-indexen.

Wanneer we deze modificatie-indexen (voor de LISREL-ontwikkelgroep) rangschikken van hoog naar laag, dan zijn dit de belangrijkste (tabel 7.3):

Tabel 7.3: *Confirmatorische factoranalyse: modificatie-indexen model 1*

pad van	pad naar	verwachte afname in Chi-kwadraat
emotionele belasting	welbevinden	85.0
emotionele belasting	spanning	83.4
emotionele belasting	sociaal-organisatorisch	77.8
werktempo/hoeveelheid	spanning	66.0
leermogelijkheden	sociaal-organisatorisch	50.1
leermogelijkheden	spanning	41.6

Uit de tweede-orde analyse van de interne consistentie wisten we al dat de dimensie "psychologische taakeisen" problematisch was (paragraaf 7.2). Dit had daar vooral betrekking op de schaal "werktempo en werkhoeveelheid" binnen deze dimensie. Nu blijken echter ook bij "emotionele belasting" en "leermogelijkheden" problemen te bestaan. De conceptuele diversiteit binnen de dimensie "psychologische taakeisen" komt bij de LISREL-analyse dus nog sterker naar voren dan bij de tweede-orde analyse van de interne consistentie.

De meest voor de hand liggende verandering in het model, op grond van de modificatie-indexen, is het splitsen van de dimensie "psychologische taakeisen" in twee onderdelen, namelijk: in een dimensie "psychologische taakeisen" ("werktempo/werkhoeveelheid", "mentale belasting", "emotionele belasting") en in een dimensie die we willen benoemen als "veelzijdigheid van het werk" ("taakvariatie" en "leermogelijkheden").

Dit is een cruciale stap, wat betreft het onderwerp "decision latitude". Het komt er op neer dat we de "skill discretion" component als aparte dimensie in het model opnemen. De "task authority" component zat al apart in het model. Aldus ontstaat model 2. "Decision latitude" is hierin dus opgesplitst in twee delen, en wordt strikt gescheiden van "psychologische taakeisen" (in model 1 viel de "skill discretion" component hier nog onder). De specificatie van model 2 is te vinden in bijlage 1.3.

De term "veelzijdigheid" is hier gekozen als label voor een dimensie, die wordt vertegenwoordigd door de schalen "taakvariatie" en "leermogelijkheden". Er was een aantal alternatieven. De term "taakvariatie" ("skill variety") was te smal: het leeraspect zit hier niet in. De term "volledigheid" roept teveel associaties op met het gelijknamige begrip in de handelingstheorie (paragraaf 2.2) en in de WEBA-methode (Pot et al., 1989). Het volledigheidsbegrip is daar breder gedefinieerd dan we hier aan schalen beschikbaar hebben. "Skill utilization" en "skill discretion" waren als termen geschikt geweest, maar zijn niet gemakkelijk letterlijk te vertalen in het Nederlands. "Veelzijdigheid" komt als vertaling nog het meest in de buurt.

De gegevens van model 2 bij de LISREL-ontwikkelgroep zijn als volgt: Chi-kwadraat= 476.18 bij 137 vrijheidsgraden ($p=.00$), AGFI= .85, en RMSR= .08. De TLI ten opzichte van het nulmodel is .72. Model 2 is een duidelijke verbetering ten opzichte van model 1. We gaan in het onderstaande verder op basis van model 2.

7.4.2 Specificiteit

De volgende vraag die we ons kunnen stellen is: meten we de onderscheiden dimensies zo efficiënt mogelijk? Dit onderzoeken we door de modificatie-indexen te bekijken (bij de LISREL-ontwikkelgroep) voor het toevoegen van error-covarianties tussen bepaalde schalen. Ook deze modificatie-indexen kunnen we weer rangschikken van hoog naar laag, met tabel 7.4 als resultaat.

De gevonden modificatie-indexen hebben vooral betrekking op de dimensie "sociaal-organisatorische" aspecten. Bij deze dimensie bestaat relatief veel overlap tussen de schalen en is de spaarzaamheid van het model minder dan bij de andere dimensies. Dit bevestigt wat tijdens de schaalconstructie al bleek uit de hoge onderlinge correlaties van de afzonderlijke schalen (paragraaf 6.5.4), en wat uit de tweede-orde analyse van de interne consistentie bleek in de vorm van een Cronbach's Alpha van .81 voor deze dimensie (paragraaf 7.2): de schalen in de "sociaal-organisatorische" dimensie zijn redundant. Op dit gebied lopen we het risico om werknemers te plagen met een lange, gedifferentieerde vragenlijst, terwijl de gemeten begrippen onderling voor een groot deel naar hetzelfde verwijzen.

Tabel 7.4: *Confirmatorische factoranalyse: modificatie-indexen model 2*

error covariantie tussen	en	verwachte afname in Chi-kwadraat
relatie met collega's	rolconflict	23.9
relatie met collega's	inspraak	16.6
relatie met collega's	informatie	15.2
betrokkenheid	piekeren	14.3
relatie met leiding	autonomie	13.1
communicatie	relatie met collega's	13.0
rolconflict	inspraak	12.3
relatie met collega's	relatie met leiding	12.3

Aangezien we in eerste instantie het model willen verbeteren zonder schalen te schrappen, leidt deze informatie nu niet meteen tot voorstellen voor het laten vallen van schalen in het LISREL-model. In eerste instantie gaan we bekijken of de situatie ook hier verbeterd kan worden door de dimensie te splitsen. We splitsen in: "relaties op het werk" (met "collega's" en "directe leiding") en "werkorganisatie" ("informatie" en "communicatie", "rolconflict" en "rolonduidelijkheid").

Op deze manier ontstaat model 3. Dit model bekijken we voor de LISREL-ontwikkelgroep. Dit levert de volgende resultaten op: De Chi-kwadraat is 453.85 bij 131 vrijheidsgraden ($p=.00$). De AGFI is .85, en de RMSR is 0.08. De TLI van model 3 ten opzichte van het nulmodel is .72. Model 3 is niet of nauwelijks beter dan model 2. We blijven dan ook voorlopig bij model 2: een dimensie minder draagt immers bij aan de spaarzaamheid van het model.

7.4.3 Informatiewaarde

Een volgende stap is het bekijken van de informatiewaarde van de schalen. De informatiewaarde heeft betrekking op de bijdrage van een individuele schaal aan het totale model. De informatiewaarde kan worden berekend door de gemeenschappelijke (verklaarde) variantie van een schaal te delen door het residu (onverklaarde variantie). Als vuistregel kan een informatiewaarde gelden van .50. Onder deze waarde draagt een individuele schaal onvoldoende bij aan het totale model.

In tabel 7.5 presenteren we voor model 2 per schaal de informatiewaarde. Bij vier schalen treffen we een informatiewaarde aan onder de .50. De schaal die het meest ongunstig scoort in dit opzicht is "mentale belasting" (informatiewaarde = .23). Deze schaal draagt relatief weinig bij aan het totale model.

Kan het zijn dat deze schaal verkeerd is geplaatst in het model? We vinden voor "mentale belasting" een redelijk grote modificatie-index (add path), en wel in de richting van "veelzijdigheid". De verwachte afname van de Chi-kwadraat bedraagt 29.3. Model 4 is gelijk aan model 2, maar dan met "mentale belasting" geplaatst onder de dimensie "veelzijdigheid", in plaats van onder "psychologische taakeisen".

Model 4 levert de volgende passingsmaten op: een Chi-kwadraat van 384.98 bij 120 vrijheidsgraden ($p=.00$), een AGFI van .86, en een RMSR van .06. De TLI ten opzichte van het nulmodel is .75. Model 4 houdt een lichte verbetering in ten opzichte van model 2. We vinden de verbetering echter niet voldoende groot om ook inhoudelijk te rechtvaardigen dat "mentale belasting" wordt herplaatst onder de dimensie "veelzijdigheid". We blijven derhalve bij model 2.

Tabel 7.5: Confirmatorische factoranalyse: informatiewaarde model 2

schaal	dimensie	verklaard	onverklaard	informatie- waarde
werktempo/hoev.	psychologische taakeisen	.30	.70	.44
mentale belasting	psychologische taakeisen	.19	.81	.23
emotionele belasting	psychologische taakeisen	.55	.45	1.24
taakvariatie	veelzijdigheid	.69	.31	2.25
leermogelijkheden	veelzijdigheid	.59	.41	1.44
taakautonomie	regelmogelijkheden	.33	.67	.50
inspraak	regelmogelijkheden	.85	.15	5.90
rolconflict	sociaal-organisatorisch	.31	.69	.45
rolonduidelijkheid	sociaal-organisatorisch	.43	.57	.77
relatie collega's	sociaal-organisatorisch	.33	.67	.49
relatie leiding	sociaal-organisatorisch	.60	.40	1.52
informatie	sociaal-organisatorisch	.45	.55	.80
communicatie	sociaal-organisatorisch	.46	.54	.86
herstelbehoefte	spanning	.51	.49	1.05
piekeren	spanning	.40	.60	.66
emot. reacties	spanning	.39	.61	.64
opzien	welbevinden	.60	.40	1.52
betrokkenheid	welbevinden	.52	.48	1.10
verlooptentatie	welbevinden	.52	.48	1.10

De overige schalen met een informatiewaarde onder de .50 ("werktempo- en hoeveelheid", "rolconflict", en "relatie met collega's"), wijken onvoldoende van deze norm af, om tot veranderingen in het model aanleiding te geven.

7.4.4 Tussentijdse conclusie "zonder weglating van schalen"

Model 2 blijkt inhoudelijk en na toepassing van de criteria van onderscheidbaarheid, specificiteit, en informatiewaarde het beste op de data te passen. Op grond van de Chi-kwadraat moet Model 2 worden verworpen als een goed model.

Volgens de AGFI (.85) en RMSR (.08) is model 2 echter een redelijk goed model te noemen. De TLI (.72) geeft bovendien aan dat model 2 duidelijk beter is dan een nulmodel dat alle schalen herleidt tot één gezamenlijke factor.

De belangrijkste bezwaren tegen model 2 zijn de lage informatiewaarde van de schaal "mentale belasting" en de onderlinge redundantie van de "sociaal-organisatorische" schalen. Tot nu toe hebben we geprobeerd tot verbeteringen te komen zonder schalen te schrappen. Model 2 noemen we dan ook het uitgebreide model.

Het uitgebreide model is inhoudelijk valide. De gevonden dimensies bij de PSA-schalen sluiten goed aan bij het model van Karasek en Theorell uit 1990. De enige afwijking is, dat de "decision latitude"-component is gesplitst in twee afzonderlijke onderdelen: "veelzijdigheid" en "regelmogelijkheden". Uit alle uitgevoerde structuur-analyses tot nu toe (exploratieve factoranalyse, tweede-orde analyse van de interne consistentie, confirmatorische factoranalyse) komt dat "veelzijdigheid" en "regelmogelijkheden" niet samen in één dimensie clusteren. Als er al ergens een vereenvoudiging mogelijk is, dan is de meest voor de hand liggende combinatie die van "psychologische taakeisen" met "veelzijdigheid". Deze oplossing blijkt echter inferieur te zijn aan een opsplitsing in afzonderlijke factoren (model 2 versus model 1).

De gevonden dimensies bij de werkstress-schalen sluiten aan bij soortgelijke bevindingen van Schaufeli en Van Dierendonck (1993). Zij vinden dat een twee-factor oplossing het best passende model oplevert bij het verklaren van de achterliggende structuur van acht schalen op het gebied van burnout en psychische klachten. De eerste factor labelen zij als een affectieve component (uitputting), de tweede als een attitude-component (negatieve houding). Deze resultaten passen goed bij de hier gevonden dimensies van "spanning" en "welbevinden".

7.5 Verbetering van het model op de LISREL-ontwikkelgroep: "met weglating van schalen"

Kan het gevonden model (bestaande uit zes dimensies) worden afgedekt met minder schalen, terwijl de passing van het model op de data gelijk blijft, of er zelfs op vooruit gaat?

Het zal duidelijk zijn dat het weglaten van schalen altijd tot een verlies van informatie leidt. In hoofdstuk 5 is middels een facetanalyse in beeld gebracht welke concepten een vragenlijst naar PSA en werkstress zou moeten afdekken. Weglating van schalen leidt altijd tot een verminderde dekking van de begrippen die uit de facetanalyse als belangrijk naar voren zijn gekomen. Een kortere vragenlijst heeft echter voor de praktijk dusdanige voordelen in termen van acceptatie, gebruiksgemak, en kosten/baten-verhouding, dat we er hier gepaste aandacht aan schenken.

Welke schalen in model 2 kunnen gemist worden? Hiertoe zijn de criteria onderscheidbaarheid en informatiewaarde tegelijk bekeken voor model 2. Het criterium van de specificiteit betrekken we in eerste instantie niet in de overweging. Dit speelde zich immers uitsluitend binnen de "sociaal-organisatorische" dimensie af. We bekijken eerst of de andere twee criteria leiden tot voorstellen om schalen te schrappen.

Qua onderscheidbaarheid vallen drie schalen in ongunstige zin op (verwachte afname van de Chi-kwadraat van 20 of meer): "mentale belasting", "informatie", en "emotionele reacties". Bij vier schalen wordt een informatiewaarde gevonden van minder dan .50, namelijk bij: "mentale belasting", "werktempo/werkhoeveelheid", "relatie met collega's", en "rolconflict".

Voor de schalen "werktempo/werkhoeveelheid" en "relatie met collega's" geldt, dat het weglaten van deze schalen tot een inacceptabel verlies van inhoudelijke dekking zou leiden van het model: beide begrippen lijken centraal voor een goede operationalisatie van PSA en werkstress. Aldus blijven over om te schrappen: "mentale belasting", "informatie", "rolconflict", en "emotionele reacties". Dit brengt ons bij model 5, waaruit deze vier schalen zijn weggelaten.

De passingsmaten voor model 5 zijn als volgt: de Chi-kwadraat is 246.96 bij 75 vrijheidsgraden ($p=.00$). De AGFI is .88 en de RMSR is .05. Voor de berekening van de TLI is een nieuw nulmodel berekend (model 5 is immers niet genest binnen het oorspronkelijk model van 19 schalen!). Ten opzichte van dit nulmodel is model 5 een verbetering van .80, uitgedrukt in de TLI. Model 5 is dus een duidelijke verbetering ten opzichte van model 2. De specificatie van model 5 is opgenomen in bijlage 1.4.

Als we model 5 bekijken, dan valt op dat alleen de "sociaal-organisatorische" dimensie en de "welbevinden"-dimensie nog worden afgedekt door meer dan twee schalen. De overige vier dimensies ("psychologische taakeisen", "veelzijdigheid", "regelmogelijkheden", en "spanning") bestaan uit twee schalen. Welk model ontstaat, wanneer we ook de dimensies "sociaal-organisatorische aspecten" en "welbevinden" terugbrengen tot een minimum van twee schalen?

Bij "welbevinden" is hiertoe de schaal "verloopintentie" buiten beschouwing gelaten. Bij de "sociaal-organisatorische" dimensie is de keuze gemaakt om dit te beperken tot de "relatie met collega's en leiding". Hierdoor versmallen we deze dimensie tot "sociale ondersteuning". Anderzijds weten we dat juist binnen deze dimensie een grote overlap bestond tussen de schalen (lage specificiteit), zodat de te verwachten overlap tussen "sociale ondersteuning" en "sociaal-organisatorische aspecten" aanzienlijk is.

Model 6 bevat dus slechts twaalf schalen: twee per dimensie. De Chi-kwadraat van dit model is 143.90 bij 39 vrijheidsgraden ($p=.00$). De bijbehorende AGFI is .89, de RMSR is .05. Ook voor model 6 is een nieuw nulmodel berekend (model 6 is immers niet genest binnen de eerder berekende modellen). Ten opzichte van dit eigen nulmodel heeft model 6 een TLI van .81. Model 6 is, net als model 5, een verbetering ten opzichte van model 2. Van model 6 is de specificatie te vinden in bijlage 1.5.

De dimensiestructuur van model 2 blijft, zowel bij model 5 als bij model 6, behouden. Vooral model 6 lijkt een bruikbaar beknopt model te zijn voor onderzoek naar PSA en werkstress.

7.6 Samenvatting en conclusie confirmatorische factoranalyse

In tabel 7.6 nemen we nog een keer de LISREL-gegevens op van het uitgebreide model (model 2 en het bijbehorende nulmodel) voor de LISREL-ontwikkelgroep, maar ditmaal presenteren we tevens de gegevens van de LISREL-testgroep.

Tabel 7.6: Confirmatorische factoranalyse: passingsmaten nulmodel/model 2

Model	N	Chi	Df	Sign	AGFI	RMSR	TLI
nulmodel/ontwikkelgroep	409	1482.56	152	.00	.55	.14	NVT
nulmodel/testgroep	407	1522.96	152	.00	.51	.15	NVT
model 2/ontwikkelgroep	409	476.18	137	.00	.85	.08	.72
model 2/testgroep	407	519.11	137	.00	.83	.09	.69

Vervolgens presenteren we dezelfde gegevens voor het beknopte model (model 6 en het bijbehorende nulmodel) in tabel 7.7:

Tabel 7.7: Confirmatorische factoranalyse: passingsmaten nulmodel/model 6

Model	N	Chi	Df	Sign	AGFI	RMSR	TLI
nulmodel/ontwikkelgroep	409	823.02	54	.00	.54	.16	NVT
nulmodel/testgroep	407	794.91	54	.00	.54	.16	NVT
model 6/ontwikkelgroep	409	143.90	39	.00	.89	.05	.81
model 6/testgroep	407	121.06	39	.00	.91	.06	.85

Voor alle modellen geldt dat ze opvallend consistent zijn qua gegevens van de LISREL-ontwikkelgroep en de LISREL-testgroep. De berekende modellen kunnen als stabiel worden beschouwd, dat wil zeggen als generaliseerbaar naar andere heterogene werknemerspopulaties.

Wanneer we alle modellen tot nu toe bekijken, moeten ze op grond van de Chi-kwadraat allemaal worden verworpen als goed passende modellen. Volgens de AGFI en RMSR zijn model 2 en vooral model 6 echter wel degelijk goed te noemen. Welke antwoorden kunnen we geven op de vragen die we ons aan het begin van dit hoofdstuk hebben gesteld?

In de eerste plaats de kritiek dat alles herleidbaar is tot één onderliggende dimensie (algemene onvrede): Zowel model 2 als model 6 zijn duidelijk beter dan de bijbehorende nulmodellen, ondanks dat de TLI niet het criterium haalt van .90. Dit is een gevolg van het feit dat het nulmodel zelf niet echt slecht is. Een nulmodel dat de afwezigheid van elke structuur veronderstelt (evenveel dimensies als schalen) had geleid tot een hogere TLI dan een dergelijk één-factor model. Met name model 6, met een TLI boven de .80 in zowel de LISREL-ontwikkelgroep als de LISREL-testgroep komt behoorlijk in de goede richting. De belangrijkste conclusie is dan ook de volgende: de schalen zijn niet goed terug te brengen tot één algemene, achterliggende dimensie. Het onderscheiden van aparte dimensies heeft wel degelijk zin.

In de tweede plaats de lengte van een vragenlijst voor PSA en werkstress: Model 6 doet het iets beter dan model 2. De onderzoeker die een zo kort mogelijke vragenlijst wil én een zo breed mogelijke dekking van "arbeidsinhoud", "arbeidsorganisatie", en "werkstress", kan derhalve het beste de twaalf schalen van model 6 afnemen. In dit geval is er nagenoeg geen redundantie meer tussen de schalen, met andere woorden: de vragenlijst is geen schaal te lang.

Ten derde de kwestie van de structuur van PSA en werkstress in relatie tot de theorie, met specieke aandacht voor het begrip "decision latitude". Het blijkt dat een model met zes dimensies het beste de structuur van PSA en werkstress beschrijft. Deze zes dimensies zijn redelijk stabiel, ondanks wisselingen in de dekking van deze dimensies qua aantallen schalen. Deze zes dimensies sluiten bovendien goed aan bij de begrippen uit theoretische kaders die veel worden gebruikt, met name de theorie van Karasek en Theorell (1990). Het uitgebreide model is te beschouwen als een luxe operationalisatie van het begrippenkader van deze theorie met behulp van vragenlijsten, het beknopte model als een beperkte (maar wel even heldere en dekkende) operationalisatie.

Model 2 en model 6 lenen zich goed voor het (multivariaat) toetsen van hypothesen gebaseerd op theorieën, zoals die van Karasek en Theorell (1990). Dit zullen we in hoofdstuk 8 dan ook gaan doen.

Inhoudelijk het belangrijkste resultaat van de confirmatorische factoranalyse is, dat de dimensies "veelzijdigheid" en "regelmogelijkheden" als afzonderlijk dienen te worden beschouwd. Met name in het Karasek-model (1979) is het gangbaar deze twee dimensies te combineren tot het begrip "decision latitude". De hier gevonden resultaten pleiten hier tegen. Als er al een combinatie mogelijk was geweest binnen de drie factoren "psychological demands", "skill discretion" en "task authority", dan was dat de combinatie van "psychological demands" en "skill discretion" geweest, en niet die van "skill discretion" met "task authority".

Een laatste vraag in dit hoofdstuk was, welke schalen welke theoretische begrippen afdekken. We beëindigen de confirmatorische factoranalyse derhalve met een overzicht van de schalen in model 2 en model 6 (zie tabel 7.8), de dimensie waarbij ze zijn ingedeeld, en de t-waarde behorend bij de ladingen op deze dimensie (afzonderlijk voor de LISREL-ontwikkelgroep (O) en de LISREL-testgroep (T)). Bij groepen van deze omvang is een t-waarde significant op het niveau van .01 (tweezijdig), wanneer deze boven de 2.59 komt. Alle t-waarden liggen hier ruim boven. De laagste t-waarde in de tabel bedraagt 8.24, de hoogste 19.57.

Over het algemeen zijn de t-waarden heel stabiel voor de LISREL-ontwikkelgroep en de LISREL-testgroep: een volgend argument voor de gevonden structuur van PSA en werkstress. Relatief grote verschillen tussen LISREL-ontwikkelgroep en LISREL-testgroep vinden we bij "rolconflict", "betrokkenheid bij de organisatie", en "verloopintentie".

Wanneer we model 2 en 6 onderling vergelijken, is er maar één schaal die opvalt door een relatief groot verschil bij zowel de LISREL-ontwikkelgroep als de LISREL-testgroep. Dit is de schaal "opzien tegen het werk". Verder blijven de t-waarden in model 2 versus model 6 redelijk stabiel: opnieuw een argument voor de gevonden structuur van PSA en werkstress.

Tabel 7.8: Confirmatorische factoranalyse: t-waarden model 2 en model 6

Schaal	Dimensie	t-waarden model 2		t-waarden model 6	
		O	T	O	T
werktempo/hoeveelheid	psychologische taakeisen	10.86	10.69	10.15	10.92
emotionele belasting	psychologische taakeisen	14.93	16.67	13.19	14.28
mentale belasting	psychologische taakeisen	8.24	10.48		
taakvariatie	veelzijdigheid	18.29	19.57	18.39	19.50
leermogelijkheden	veelzijdigheid	16.67	16.20	16.46	15.92
taakautonomie	regelmogelijkheden	11.90	11.11	11.62	10.96
inspraak	regelmogelijkheden	18.76	18.57	17.23	17.72
relatie collega's	sociaal-organisatorisch	11.85	12.56	12.36	12.28
relatie leiding	sociaal-organisatorisch	17.61	16.96	15.92	16.76
informatie	sociaal-organisatorisch	14.33	12.65		
communicatie	sociaal-organisatorisch	14.69	15.04		
rolconflict	sociaal-organisatorisch	11.40	14.51		
rolonduidelijkheid	sociaal-organisatorisch	14.10	13.44		
herstelbehoefte	spanning	13.94	14.16	13.14	13.51
piekeren	spanning	12.13	12.37	11.20	10.97
emotionele reacties	spanning	12.01	10.37		
opzien tegen het werk	welbevinden	16.56	16.07	14.37	13.58
betrokkenheid bij org.	welbevinden	15.21	11.58	13.30	10.66
verloopintentie	welbevinden	15.21	11.26		

7.7 Generaliseerbaarheid

In dit hoofdstuk is de dimensionalisering van PSA en werkstress onderzocht bij twee heterogene werknemersgroepen. We kunnen op basis van de resultaten tot nu toe weinig uitspraken doen over de vraag, of de gevonden dimensie-structuur generaliseerbaar is naar specifieke werknemersgroepen. Om hier toch ten dele een antwoord op te geven, zijn de passingsmaten van model 6 en het bijbehorend nulmodel berekend voor medewerkers in productiebedrijven en gezondheidszorginstellingen afzonderlijk. De resultaten zijn weergegeven in tabel 7.9.

Tabel 7.9: Passingsmaten nulmodel/model 6 voor productie en gezondheidszorg afzonderlijk

Model	N	Chi	Df	Sign	AGFI	RMSR	TLI
nulmodel/productie	398	702.89	54	.00	.60	.14	NVT
nulmodel/gez. zorg	418	1026.22	54	.00	.46	.19	NVT
model 6/productie	398	139.74	39	.00	.89	.06	.79
model 6/gez. zorg	418	160.08	39	.00	.88	.06	.83

Zoals uit bovenstaande tabel blijkt, is de dimensie-structuur qua passing ongeveer gelijk bij productiebedrijven en gezondheidszorginstellingen. De AGFI bedraagt .89 in de ene, en .88 in de andere groep. De RMSR is .06 in beide groepen. Omdat het nulmodel het in de productiebedrijven relatief goed doet, is de TLI voor model 6 hier iets kleiner.

De hier gevonden waarden van de passingsmaten wijken nauwelijks af van die bij de LISREL-ontwikkelgroep en de LISREL-testgroep. Alhoewel niet afdoende, is hiermee toch enig bewijs geleverd voor de generaliseerbaarheid van de dimensie-structuur naar specifieke werknemersgroepen.

7.8 Een model voor de beleving en beoordeling van PSA en werkstress

Op basis van de literatuur (Meijman, 1989; Karasek en Theorell, 1990; Van Dijk et al., 1990; Warr, 1994) kunnen verbanden worden verondersteld tussen de vier PSA-dimensies en de twee werkstress-dimensies. Nu de structuur van PSA en werkstress is verhelderd kan de aandacht worden gevestigd op deze mogelijke verbanden. Dit levert vragen op als: Wat is het effect op de passingsmaten wanneer we dergelijke verbanden toevoegen aan het model? Welke PSA-dimensies hangen samen met welke werkstress-dimensies?

Het beantwoorden van deze vragen geeft ons een eerste inzicht in de relatie tussen PSA en werkstress, een volgende kernvraag van dit proefschrift. We gaan op deze vragen in aan de hand van het beknopte model (model 6).

In dit hoofdstuk is tot nu toe gewerkt met latente begrippen. De hier gestelde vragen zijn beter te beantwoorden door een LISREL-analyse uit te voeren op zogenaamde "dimensiescores".

Deze scores worden verkregen door de schaalscores behorend bij een bepaalde dimensie bij elkaar op te tellen. Zodoende beperkt de LISREL-analyse in deze paragraaf zich uitsluitend tot geobserveerde scores. Er zitten geen latente begrippen meer in.

Heel expliciet is op dit moment in de analyse gekeken in hoeverre sprake is van item-overlap tussen de schalen die de PSA-dimensies meten en schalen die de werkstress-dimensies meten. Dit in verband met de "triviality trap" (zie 4.1.3). Dergelijke item-overlap bleek nagenoeg nihil te zijn. De hier gerapporteerde relaties tussen PSA en werkstress zijn dus niet of nauwelijks aan item-overlap toe te schrijven.

In deze paragraaf vindt de analyse niet meer plaats op de ontwikkelgroep en de testgroep afzonderlijk, maar op de totale groep beschikbare respondenten (N=816). Het gaat immers niet meer om kruisvalidatie, maar om modeltoetsing. Alvorens een LISREL-model te berekenen op de dimensiescores worden in tabel 7.10 eerst de correlaties gepresenteerd tussen de zes dimensiescores onderling.

Tabel 7.10: Correlaties tussen de dimensiescores

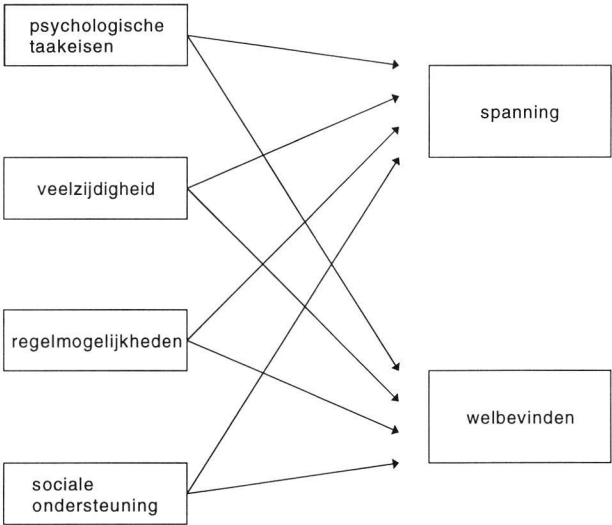
	psy. eisen	gebrek veelzijd.	gebrek regelmog.	gebrek soc. steun	spanning	onwelbev.
psychologische eisen	1.00					
gebrek veelzijdigheid	-.37	1.00				
gebrek regelmogelijkheden	-.19	.54	1.00			
gebrek sociale steun	.22	.28	.36	1.00		
spanning	.49	-.10	.02	.28	1.00	
onwelbevinden	.05	.41	.34	.40	.20	1.00

Hoge correlaties (>.60) komen tussen de dimensie-scores niet voor. Wanneer .40 als criterium wordt genomen, dan valt een aantal correlaties op. Zo is er een correlatie tussen "psychologische taakeisen" en "spanning" (.49). Ook tussen "gebrek veelzijdigheid" en "gebrek regelmogelijkheden" bestaat een correlatie (.54). De correlatie tussen "gebrek veelzijdigheid" en "onwelbevinden" bedraagt .41. De correlatie tussen "gebrek sociale steun" en "onwelbevinden" is .40.

Over de relatie tussen "veelzijdigheid" en "regelmogelijkheden" is in het voorgaande reeds veel gezegd. Uit de gepresenteerde correlaties blijkt opnieuw dat er een relatie tussen deze aspecten bestaat. We handhaven beide begrippen verder afzonderlijk, op grond van de resultaten uit de confirmatorische factoranalyse.

Bij de confirmatorische factoranalyse zijn de correlaties tussen de latente onafhankelijke variabelen vrijgelaten. Gezien de bovenstaande correlaties tussen de geobserveerde dimensiescores, zullen ook bij de analyse op de geobserveerde dimensiescores de correlaties tussen de onafhankelijke variabelen worden vrijgelaten.

We onderzoeken nu de passingsmaten voor een LISREL-model, waarin elk van de vier PSA-dimensies een relatie heeft met elk van de twee werkstress-dimensies (zie figuur 7.1).



Figuur 7.1: Een model voor de beleving en beoordeling van PSA en werkstress

In tabel 7.11 staan de passingsmaten voor dit model. Het model voldoet net niet aan het AGFI-criterium (.90). Het voldoet niet aan het Chi-kwadraat criterium ($p < .05$). Het model voldoet wel aan het RMSR-criterium ($< .05$). Gegeven deze resultaten gaat het om een redelijk goed passend model. De hoogte van de chi-kwadraat geeft aan dat de error-termen voor de twee werkstress-dimensies een significante correlatie hebben. De lage RMSR geeft echter aan dat dit effect voor het totale model niet zeer belangrijk is.

Tabel 7.11: Passingsmaten model met correlaties tussen PSA en werkstress

N	Chi	Df	Sign	AGFI	RMSR
816	14.23	1	.00	.88	.02

Tenslotte zijn we geïnteresseerd in de vraag welke PSA-dimensies samenhangen met welke werkstress-dimensies in dit model. Dit bekijken we aan de hand van de t-waarden, die zijn opgenomen in tabel 7.12. Met behulp van sterretjes wordt de significantie aangegeven. De kritieke grenzen liggen bij <.10 (*), <.05 (**), <.01 (***). Tevens zijn de R-kwadraten van de vergelijking ter voorspelling van "spanning" en "welbevinden" in de tabel opgenomen.

Tabel 7.12: T-waarden model met correlaties tussen PSA en werkstress

PSA-schaal	t-waarde spanning		t-waarde welbevinden	
psy. taakeisen	13.38	***	3.94	***
veelzijdigheid	-0.12		-8.94	***
regelmogelijkheden	1.22		2.38	**
sociale ondersteuning	4.72	***	7.08	***
R-kwadraat	.27		.28	

Uit deze tabel blijkt dat "psychologische taakeisen" een significante samenhang heeft met zowel "spanning" als "welbevinden". De relatie met "spanning" is het meest duidelijk. "Veelzijdigheid" heeft alleen een significante relatie met "welbevinden". "Regelmogelijkheden" hangen eveneens alleen samen met "welbevinden", zij het in minder sterke mate. "Sociale ondersteuning" tenslotte, hangt zowel samen met "spanning" als met "welbevinden". De relatie met "welbevinden" is het meest duidelijk.

Bij "spanning" is de t-waarde van "psychologische taakeisen" opvallend groot, in verhouding tot de andere PSA-dimensies. Alhoewel alle vier de PSA-dimensies samenhangen met "welbevinden", zijn "veelzijdigheid" en "sociale ondersteuning" de beste voorspellers. De R-kwadraten komen zowel bij "spanning" als bij "welbevinden" in de buurt van de 30%.

Deze resultaten geven een eerste indruk van de aard van de verbanden tussen PSA en werkstress. Op de relatie tussen PSA en werkstress wordt in het volgende hoofdstuk dieper ingegaan.

Hoofdstuk 8: De relatie tussen PSA en werkstress

In hoofdstuk 7 is gekeken naar de dimensies die in PSA en werkstress zijn te onderscheiden. Hiermee is het onderwerp van de begripsvaliditeit (convergent en divergent) aan de orde geweest.

Andere vormen van validiteit zijn die van de concurrente validiteit en predictieve validiteit. Beide zijn vormen van criterium-validiteit: de validiteit van een schaal wordt vastgesteld aan de hand van een bepaald criterium. Dit gebeurt tegelijkertijd met de vragenlijstafname (concurrent) of op een later tijdstip (predictief). Beide aspecten van validiteit worden in het vervolg van dit proefschrift op een multivariate manier onderzocht. De concurrente validiteit komt aan bod in hoofdstuk 8. De predictieve validiteit ten aanzien van ziekteverzuim wordt in hoofdstuk 9 behandeld.

Bij de literatuur over de relatie tussen PSA en werkstress zijn in deel I een aantal thema's besproken die hier kort worden herhaald. Hieraan worden een aantal opmerkingen toegevoegd die specifiek betrekking hebben op dit hoofdstuk.

Confounding

Een van de voornaamste kritiekpunten op cross-sectioneel vragenlijstonderzoek naar PSA en werkstress was, dat dit confounding in de hand werkt. Er ontstaan schijnverbanden tussen PSA en werkstress, doordat beide aspecten worden vertekend door algemene psychologische factoren die de beantwoording van vragenlijsten beïnvloeden (Kasl, 1978; 1986; Leventhal en Tomarken, 1987).

Een variabele die de afgelopen tijd in dit kader meerdere malen is onderzocht is negatieve affectiviteit (NA) (Watson en Pennebaker, 1989). In dit hoofdstuk zullen we onderzoeken in hoeverre NA inderdaad optreedt als confounder bij vragenlijst-onderzoek naar relaties tussen PSA en werkstress. Wat ons hierbij in het bijzonder interesseert, nu de structuur in PSA en werkstress helder is geworden, is of het vertekende effect van NA bij spanning net zo groot is als bij welbevinden. Onderzoek op dit gebied doet namelijk vermoeden dat NA een grotere rol speelt als confounder bij spanning dan bij welbevinden (Brief et al., 1988; Parkes, 1990; Chen en Spector, 1991; Parkes et al., 1994).

Ervaren onbillijkheid (EO) is een begrip uit de sociale vergelijkingstheorie dat de laatste tijd veelvuldig is onderzocht in relatie tot zowel PSA als werkstress (Buunk en Schaufeli, 1993; Geurts, 1994; Van Yperen, 1995). EO blijkt te correleren met een negatieve beoordeling van zowel PSA-factoren (hoge werkdruk, lastige patiënten) als werkstress-symptomen (burnout, dissatisfactie).

De vraag kan dan ook gesteld worden of EO niet, net als NA, werkt als confounder bij cross-sectioneel onderzoek naar PSA en werkstress met vragenlijsten. Eventuele samenhangen tussen PSA- en werkstress-factoren zouden kunnen ontstaan doordat EO over de uitwisselingsrelatie tussen werknemer en organisatie een negatieve beoordeling van zowel PSA- als werkstressfactoren in de hand werkt. In dit hoofdstuk onderzoeken we deze mogelijkheid.

De limiet van 10% à 15% verklaarde variantie

Semmer, Zapf en Greif (1995) beargumenteren dat er een limiet bestaat van 10% à 15% verklaarbare variantie door PSA-factoren in individueel arbeidsgerelateerd welbevinden. Zij weten aan te tonen dat de standaardverklaringen hiervoor (slechte meetinstrumenten, slechte onderzoeksdesigns) niet juist zijn. Het lijkt erop dat de multicausale aard van werkstress een betere verklaring biedt.

Zij noemen daarom 10% à 15% het "theoretisch maximum" voor de verklaarbare variantie in individuele uitkomstmaten. De enige manier om het niveau van verklaarde variantie volgens hen te verhogen, is het verklaren van variantie op een geaggregeerd niveau. Zij onderzoeken dit op het gebied van welbevinden in het werk, en introduceren hiertoe een nieuw begrip: "shared well-being".

Onderzoek op dit terrein is van belang, omdat een verklaringspercentage van 10% à 15% in de praktijk soms weinig houvast en argumenten biedt voor verandering en probleemaanpak. We gaan in dit hoofdstuk een relatief nieuwe statistische techniek, multi-level analyse (Bosker en Snijders, 1990; Bryk en Raudenbush, 1992; Longford, 1993), gebruiken om de door Semmer, Zapf en Greif voorgestelde oplossingsrichting nader te onderzoeken. Wellicht dat we door het simultaan onderzoeken van verklaringen voor werkstress op zowel individueel als geaggregeerd niveau een grotere mate van verklaarde variantie kunnen bereiken.

Een tweede vraagstuk op dit gebied is, of spanning en welbevinden verschillend zijn wat betreft de verdeling van de variantie over individuele versus geaggregeerde verklaringsniveaus. Inzicht hierin is voor de praktijk relevant, omdat dit aanwijzingen geeft over het niveau van beïnvloeding binnen een organisatie, waarvan het meeste resultaat te verwachten is. We zullen in dit hoofdstuk multi-level analyse gebruiken om ook deze vraag te beantwoorden.

Eén of meer werkstress-mechanismen?

Het demand-control model van Karasek (1979) is op dit moment het meest populaire onderzoeksmodel van PSA en werkstress. Wat betreft de relatie tussen PSA en werkstress is het model weinig genuanceerd. In het boek van Karasek en Theorell (1990) staat een enorme hoeveelheid onderzoek naar het interactie-effect van "psychological job demands" en "decision latitude" op uiteenlopende parameters als: depressie, angst, medicijngebruik, uitputting, hart- en vaatziekten, ziekenhuisopname, bloeddruk, productiviteit, verzuim, arbeidssatisfactie et cetera.

In hoofdlijnen werkt het demand-control model op al deze parameters hetzelfde in: hoe meer eisen en hoe minder regelmogelijkheden, des te meer risico op werkstress-symptomen. Het uitgebreide demand-control model (Karasek en Theorell, 1990) verandert niets aan deze opzet, maar voegt een as toe aan het model: hoe minder sociale ondersteuning, hoe meer risico.

De theorie van Karasek (1979) wordt hier verder gebruikt als prototypisch voor werkstress-theorieën die uitgaan van één algemeen mechanisme voor het ontstaan van werkstress. Eerlijkheidshalve dient wel opgemerkt te worden dat Karasek en Theorell (1990) in hun boek wel degelijk twee mechanismen onderscheiden die binnen het demand-control model werkzaam zijn. In de eerste plaats bespreken zij het "strain"-mechanisme, dat verantwoordelijk is voor het ontstaan van werkstress-symptomen. In de tweede plaats noemen ze het "active learning"-mechanisme, dat verantwoordelijk is voor het al dan niet optreden van groei en ontwikkeling in het werk. Over de relatie tussen deze twee mechanismen doen Karasek en Theorell (1990) bovendien een interessante uitspraak: "*Stress inhibits learning, learning inhibits stress*" (p. 89). Deze stelling vormt de basis voor hun werkstress-aanpak. Het accent in hun boek ligt echter duidelijk bij het "strain"-mechanisme, en dan vooral qua effecten op gezondheidkundige uitkomstmaten.

Warr (1987; 1994) maakt onderscheid tussen de effecten van verschillende typen PSA op verschillende soorten werkstress. Hij onderscheidt enerzijds PSA-factoren die nooit teveel aanwezig kunnen zijn: de CE-factoren. Anderzijds onderscheidt hij PSA-factoren die bij een bepaalde dosis een specifiek "toxisch" effect hebben: de AD-factoren. De theorie van Warr pleit voor differentiatie in een aantal werkstress-mechanismen, die parallel aan elkaar werkzaam zijn.

Dit hoofdstuk gaat in op de vraag, of de vier PSA-factoren ("psychologische taakeisen", "veelzijdigheid", "regelmogelijkheden" en "sociale steun") een algemeen effect hebben op "spanning" en "welbevinden" (Karasek), of dat deze PSA-factoren gedifferentieerd samenhangen met "spanning" en "welbevinden" (Warr). We zagen in paragraaf 7.8 al enkele resultaten, die wijzen in de richting van differentiatie.

Dit deel van het proefschrift bouwt verder op de resultaten van het beknopte model in paragraaf 7.8. Alvorens nader in te gaan op de techniek van multi-level analyse, wordt eerst behandeld op welke wijze NA en EO in dit onderzoek zijn geoperationaliseerd.

8.1 Operationalisatie negatieve affectiviteit

Watson en Pennebaker (1989) spreken over NA als een algemene dimensie van subjectieve spanning, die moet worden gezien als een persoonlijkheidskenmerk.

Over de meting van NA zeggen zij het volgende: "*The trait NA construct is measured by a large number of commonly used scales and has been variously called neuroticism, trait anxiety and general adjustment by other investigators*" (p. 235).

In artikelen over NA komen we bijvoorbeeld operationalisaties tegen als de Taylor Manifest Anxiety Scale (Brief et al., 1988), de Neuroticisme-schaal uit de Eysenck Personality Scale (Parkes, 1990; Parkes et al., 1994), en de trait-anxiety schaal uit de State-Trait Anxiety Inventory (Chen en Spector, 1991). Watson, Clark en Tellegen (1988) ontwikkelden een tweetal nieuwe schalen voor het meten van negatieve en positieve affectiviteit: de PANAS (positive and negative affectivity schedule). De diverse operationalisaties van NA zijn zeer verschillend, vooral wat betreft de hoeveelheid items die betrekking hebben op somatische klachten en neurotisch gedrag.

Deze maten hebben gemeenschappelijk dat ze NA meten als een persoonlijkheidskenmerk en dat de items raken aan het onderwerp emotionaliteit. Een dergelijke maat is echter binnen de context van dit onderzoek niet beschikbaar. De enige maat die in de buurt komt is de emotionele reacties-schaal. Deze heeft echter betrekking op de afgelopen week in het werk. Deze schaal is meer te beschouwen als een toestandsmaat, als tegengesteld aan een dispositie-maat (Spielberger, 1966).

De item-overlap tussen de emotionele reacties-schaal en bekende maten van NA is groot. Alle twaalf items hebben bijvoorbeeld letterlijk en/of inhoudelijk overlap met de ZBV (Van Der Ploeg, Defares, en Spielberger, 1980). Dit is de Nederlandse vertaling van de eerder genoemde State-Trait Anxiety Inventory (Spielberger et al., 1970).

De emotionele reacties-schaal is qua opzet vergelijkbaar met de PANAS, de maat die door Watson en Tellegen (1988) is geconstrueerd voor het meten van NA. Er is namelijk gestreefd naar zo min mogelijk item-overlap met somatische klachten-schalen en met schalen die psychologische begrippen meten zoals competentie en motivatie. Dit is niet zo vreemd, want de emotionele reacties-schaal is een vertaling van een schaal van Warr (1990b). Warr (1990b) baseerde zich bij zijn schaalconstructie en de achterliggende theorie (Warr, 1987) op het eerdere werk van Watson en Tellegen (1985) en Watson, Clark en Tellegen (1988).

Met de PANAS deelt de emotionele reacties-schaal ook de overeenkomst dat er exact evenveel positief als negatief geformuleerde items zijn in de schaal. Een belangrijk verschil met de PANAS is echter, dat deze twee groepen items daarin als theoretisch verschillend worden beschouwd, en twee afzonderlijke begrippen meten: negatieve en positieve affectiviteit. De schaal emotionele reacties meet slechts één concept.

De emotionele reacties-schaal lijkt een goede, eenduidige operationalisatie van NA te zijn, met betrekkelijk weinig item-overlap met somatische klachtenlijsten. Onverklaard is wel waarom de twee-dimensionele structuur bij deze vertaling van Warr (1990b) niet werd teruggevonden tijdens de schaalconstructie (hoofdstuk 6). Warr (1990b) vindt namelijk wel zo'n twee-dimensionele structuur, en deze structuur sluit aan bij zijn theorie over welbevinden in het werk. Sevastos et al. (1992) en Schalk et al. (1995) toonden echter aan dat deze twee-dimensionele structuur niet zonder problemen is, zodat het werken met één totaalscore voor deze schaal voorlopig een goede optie is.

Het belangrijkste probleem om de emotionele reacties-schaal te gebruiken als maat voor NA zit derhalve hierin, dat het meer een toestands-maat is dan een dispositie-maat. Voor de claim dat NA een confounder is in onderzoek naar de relatie tussen PSA en werkstress, lijkt dit punt echter niet zo'n bezwaar. Juist wanneer wordt gewerkt met een cross-sectionele onderzoeksopzet, en een werknemer hoog scoort op NA in het werk gedurende de afgelopen week, mag immers een vertekening worden verwacht in een gelijktijdige beoordeling van PSA en werkstress. Bovendien is de correlatie tussen state- en trait-maten van NA hoog. De correlaties tussen de ZBV-state en de ZBV-trait liggen bijvoorbeeld in de orde van grootte tussen .52 en .84 bij uiteenlopende onderzoekspopulaties (Van der Ploeg, Defares en Spielberger, 1980).

8.2 Operationalisatie ervaren onbillijkheid

Op grond van de billijkheidstheorie van Adams (1965) kan voorspeld worden dat werknemers die een ongelijkheid ervaren tussen de kosten en opbrengsten van het werk de neiging zullen hebben deze ongelijkheid te compenseren. Zo is het mogelijk om ziekteverzuim te gaan zien als compensatiegedrag (Geurts, 1994). Ook burnout-symptomen kunnen worden gezien als het resultaat van EO (Buunk en Schaufeli, 1993).

Alhoewel de billijkheidstheorie (en de sociale vergelijkingstheorie in het algemeen) kan worden toegepast op allerlei afzonderlijke facetten van arbeid en organisatie (onbillijkheid over werkdruk, onbillijkheid over taakhoud, onbillijkheid over inspraakmogelijkheden) is het gebruikelijk om EO op te vatten als een globale beoordeling van de uitwisselingsrelatie tussen de werknemer en de organisatie.

Het is wel gebruikelijk om hierbij te kijken naar de aard van de vergelijkingsspersoon (collega's, vrienden, familie) en de aard van de vergelijkingsdimensie (beloning, waardering, status).

In de vragenlijstschalen, die in deel II zijn geconstrueerd, is gezocht naar items die in aanmerking komen als operationalisatie van EO. Deze werden gevonden in 3 items uit de schaal "beloning". Deze items zijn verder inhoudelijk bekeken.

Het item "Vindt u dat u voldoende betaald krijgt voor het werk dat u levert?" is een echt onbillijkheidsitem, waarbij de "input" vaag is gehouden en de "opbrengst" concreet is geformuleerd in termen van financiële beloning. Bij dit item is niet gespecificeerd wie de vergelijkingspersoon of vergelijkingsgroep is.

Het volgende item is: "Vindt u dat u rechtvaardig betaald wordt in vergelijking met anderen op uw afdeling". Dit item vraagt eigenlijk naar hetzelfde als het vorige item (de "input" is niet gespecificeerd, de "opbrengst" is gespecificeerd als financiële beloning), maar geeft een vergelijkingsgroep aan: anderen op de afdeling. Dit is, zoals blijkt uit onderzoek van Geurts (1994), ook de meest relevante vergelijkingsgroep.

Het derde item is: "Denkt u dat het loon in uw bedrijf lager ligt dan in vergelijkbare bedrijven". Dit item vergelijkt het ene bedrijf met het andere. Onduidelijk is wat een antwoord op deze vraag zegt over het individu dat deze vraag invult. Een werknemer van een bedrijf kan dit item bevestigend beantwoorden voor zijn bedrijf, en toch absoluut vrij zijn van onbillijkheidsgevoelens naar zijn eigen inkomen toe. Hierdoor is het niet zo'n geschikt item voor het meten van EO in de hier beoogde context.

Aldus komen we voor dit onderzoek tot een operationalisatie van EO met behulp van twee items die betrekking hebben op "beloning". Dit is geen brede operationalisatie van EO, maar wel een operationalisatie op het gebied van de meest tastbare opbrengst van werk: geld. Op dit terrein zijn in ieder geval duidelijke onbillijkheidsgevoelens te verwachten, wanneer iemand vindt dat er sprake is van over- of onderbetaling (Thierry, 1990; Hermkens, 1995). De Cronbach's Alpha van deze twee items is .72.

8.3 Multi-level analyse

In werkbelevingsonderzoek wordt het effect onderzocht van het uitoefenen van een bepaald beroep c.q. het werken in een bepaalde organisatie op het functioneren van individuele werknemers. Tot nu toe hadden onderzoekers twee mogelijkheden om dit te doen.

De eerste en meest gebruikte methode is, om gegevens van werkbeleving en relevante uitkomsten (verzuim, verloop, intrede WAO) beide te verzamelen op het niveau van de individuele werknemer en hierop multivariate analyses te doen. Door in de eerste stap van de analyse een aantal individuele kenmerken in te voeren en

vervolgens te kijken wat de werk- en organisatievariabelen hier bovenop aan variantie verklaren, concludeert men wat de bijdrage is van werk en organisatie aan de uitkomst-variabelen.

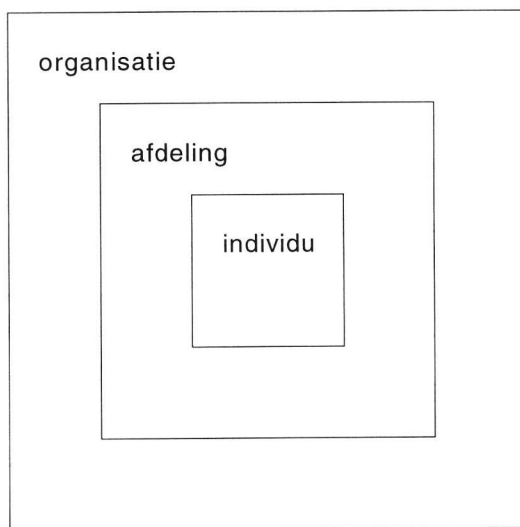
De hierbij gebruikte standaardtechnieken laten echter niet toe, om er rekening mee te houden dat de individuen systematisch verdeeld zijn over de groepen waar we conclusies over willen trekken. Anders gezegd: deze technieken gaan ervan uit dat de individuen waarvan afhankelijke en onafhankelijke variabelen worden verzameld, afkomstig zijn uit een enkelvoudige, random steekproef. Bij werkbelevingsonderzoek is dit doorgaans niet het geval: de individuen komen juist uit systematische groepen (functies, afdelingen, bedrijven), en hebben tot op zekere hoogte iets gemeenschappelijks. De meeste van dit soort analyses schenden op dit punt dan ook de assumpties die aan technieken als (multi-pele) regressie en (multi-pele) variantie-analyse ten grondslag liggen.

Het alternatief is, om de gegevens van de individuele werknemers eerst te aggregeren tot scores van functiegroepen en afdelingen, alvorens dan hierop een vergelijkbare multivariate analyse te doen. De problemen hiervan zijn echter legio: in de eerste plaats komen de afdelingen en functiegroepen meestal systematisch uit bepaalde bedrijven, zodat ook hier geen sprake is van een echte random steekproef. Het probleem herhaalt zich, maar dan een analyseniveau hoger. In de tweede plaats neemt door het aggregeren de steekproefomvang af, waardoor de multivariate analyse minder elementen in de vergelijking toestaat, en zeer grote onderzoeksgroepen nodig zijn om een goede analyse uit te kunnen voeren. Verder zijn bij een dergelijke procedure de eenheden van analyse niet de individuele werknemers, zoals de bedoeling is, maar functiegroepen of afdelingen.

Semmer, Zapf en Greif (1995) lossen dit probleem op met behulp van LISREL. Ze gebruiken gegevens van individuele werknemers en externe beoordelaars om een latent begrip op afdelingsniveau te voorspellen, dat ze "shared well-being" noemen. Zij gaan er vervolgens van uit dat "shared well-being" beter te verklaren is op basis van PSA, dan het welbevinden van individuele werknemers.

Er bestaat echter nog een alternatieve manier om dit methodologische probleem op te lossen: multi-level analyse (Bosker en Snijders, 1990; Bryk en Raudenbush, 1992; Longford, 1993). Deze techniek is speciaal geschikt voor multivariate analyses op zogenaamde hiërarchisch geneste datastructuren, zoals individuen binnen functiegroepen, die weer vallen binnen afdelingen, die weer vallen binnen bedrijven, die weer vallen binnen sectoren enzovoort.

Een manier om een dergelijke datastructuur voor te stellen, is als een serie in elkaar vallende vierkanten (zie figuur 8.1). Het laagste niveau (individuele werknemers) is volledig ingebed in het tweede niveau (afdelingen of functies). Dit tweede niveau is weer volledig ingebed in het derde niveau (organisaties of organisatie-onderdelen).



Figuur 8.1: Illustratie van een organisatie als een hiërarchisch geneste datastructuur

Multi-level analyse gaat uit van een geneste datastructuur, die perfect aansluit bij de manier waarop gegevens van bedrijven en instellingen meestal worden verzameld bij werkbelevingsonderzoek met vragenlijsten.

Wanneer we een afhankelijke variabele hebben op het niveau van de individuele werknemer, dan is het met deze techniek mogelijk om onafhankelijke variabelen die op verschillende niveaus zijn verzameld tegelijkertijd op hun effecten te onderzoeken.

We gebruiken in deze paragraaf multi-level analyse om de relatie te onderzoeken tussen PSA (op individu-niveau, op afdelingsniveau, en op organisatie-niveau) en werkstress (op individu-niveau), waarbij we corrigeren voor persoonskenmerken en confounders (NA en EO) op individueel niveau. De multi-level analyse werd uitgevoerd met behulp van het computerprogramma VARCL (Longford, 1993).

In de presentatie van de resultaten zullen we zoveel mogelijk aansluiten bij de conventies voor het presenteren van gegevens uit een multiële regressie-analyse, om de leesbaarheid en interpretatie van de resultaten uit deze nieuwe techniek te vergemakkelijken.

8.4 Gegevensbestand en voorbereiding van de analyse

De gegevens voor de multi-level analyse zijn afkomstig van een gezondheidszorg-instelling. De keuze voor deze organisatie is gemaakt vanwege de beschikbaarheid van goede informatie over de nesting van individuen binnen organisatie-onderdelen, en vanwege de continuïteit met het volgende hoofdstuk, dat betrekking heeft op verzuimcijfers.

Er zal worden gewerkt met een multi-level analyse op drie niveaus. Het eerste niveau is dat van individuele werknemers ($N=411$). Het tweede niveau is dat van afdelingen ($N=44$). Het derde niveau is dat van sectoren ($N=8$). Voor de voorbereiding van de gegevens ten behoeve de multi-level analyse houdt dit in, dat er drie gegevensbestanden worden aangemaakt: één per niveau. Deze drie bestanden worden gelijktijdig verwerkt in de analyse.

De drie bestanden moeten zorgvuldig zijn opgebouwd. Alle individuen in het eerste bestand ($N=411$) zijn geordend in afdelingsvolgorde, en de afdelingen in dit bestand staan in sectorvolgorde. Het tweede bestand bevat de afdelingen ($N=44$), geordend naar sector. Het derde bestand ($N=8$) bevat de sectoren in de juiste volgorde.

Voor een goed begrip van de resultaten van een multi-level analyse is het van belang om goed in het oog te houden dat er twee toevoegingen zijn ten opzichte van een standaard multivariate analyse. Een standaard-analyse bevat afhankelijke en onafhankelijke variabelen, beide gemeten op het niveau van de individuele werknemer. De eerste toevoeging is dat er onafhankelijke variabelen in de multi-level analyse zijn opgenomen, die zijn gemeten op hogere aggregatieniveaus (in dit geval: afdelingen en sectoren). De tweede, en minstens even belangrijke, toevoeging is de nestingstructuur. Deze stelt ons in staat om de variantie in de afhankelijke variabele op te splitsen in componenten naar verklaringsniveau (in dit geval: individuele werknemers, afdelingen en sectoren).

Voor de indeling in afdelingen en sectoren is de lijnorganisatie binnen deze instelling gevolgd. De sectoren zijn relatief autonome units binnen de organisatie, en hangen samen met bepaalde vormen van dienstverlening dan wel interne ondersteuning. De sectoren hebben elk hun eigen managementteam, met aan het hoofd een eigen sectormanager. De andere leden van het management zijn de afdelingshoofden. Elke sector heeft dus een eigen middenkader.

De nestingstructuur is eenduidig, dat wil zeggen: elk individu is slechts onder één afdeling geplaatst, en elke afdeling is slechts onder één sector geplaatst.

De afhankelijke variabelen zullen afzonderlijk (niet multivariaat) worden geanalyseerd. De afhankelijke variabelen zijn enigszins scheef verdeeld, maar zullen desalniettemin worden geanalyseerd met behulp van een normale foutenverdeling.

Er zullen geen interactietermen tussen onafhankelijke variabelen worden onderzocht, ook niet tussen "psychologische taakeisen" en "regelmogelijkheden" (Karasek, 1979; Karasek en Theorell, 1990). Recent onderzoek spreekt het bestaan van dergelijke interacties tegen (Spector, 1987; Parkes, 1990; Warr, 1990a; Fletcher en Jones, 1993). Voorlopig gaan we ervan uit dat in onderzoek, waarin interactietermen zijn onderzocht met behulp van regressie-technieken, geen aanwijzingen zijn gevonden dat een "demand-control" interactie bestaat. Aangezien multi-level analyse een aan regressie-analyse verwante techniek is, is de verwachting dat een "demand-control" interactie ook hiermee niet zal kunnen worden aangetoond. Het toevoegen van interactietermen aan de multi-level analyse (op elk van de drie niveaus) zou de analyse bovendien verder compliceren. Om inhoudelijke redenen en om de analyse relatief eenvoudig te houden wordt onderzoek naar interactietermen hier verder achterwege gelaten.

Bij multi-level analyse moet een keuze gemaakt worden tussen het werken met een fixed- dan wel random-effects model voor alle onafhankelijke variabelen, ofwel met beide. Er zal hier alleen worden gewerkt met fixed-effects van de onafhankelijke variabelen. Deze keuze is gemaakt omdat onvoldoende onderbouwing is te geven aan de veronderstelling dat bepaalde onafhankelijke variabelen in groep A een ander effect zullen hebben dan in groep B. Een random-effects model zou bijvoorbeeld pas zin hebben als redelijkerwijs te verwachten is, dat een hoge score op "werktempo en werkhoeveelheid" in afdeling X een ander effect heeft dan in de afdeling Y. Voorlopig is het aannemelijker dat in beide afdelingen werknemers hiervan gelijksoortige effecten ervaren. Ook voor het toevoegen van random-effects aan het model geldt, dat het de analyse verder zou compliceren. Er wordt opnieuw voor gekozen om de analyse betrekkelijk eenvoudig te houden, en de specificatie van random effects achterwege te laten.

Op alle drie de niveaus is het mogelijk te werken met kenmerken van dat niveau als onafhankelijke variabelen. Op het niveau van de individuele werknemer zal worden gewerkt met de kenmerken: geslacht, dienstjaren, leeftijd, uren werk per week, en functieniveau (laag/hoog). Bij deze laatste variabele ligt de grens tussen laag en hoog bij HBO-niveau. Ook op het niveau van de afdelingen is een kenmerk opgenomen in de analyse: het al dan niet werken met direct patiëntencontact. Op sectorniveau tenslotte is eveneens een kenmerk opgenomen in de analyse, namelijk het soort sector (ondersteunend versus dienstverlenend).

PSA zal op alle drie de niveaus worden uitgedrukt in dezelfde parameters. Op individu-niveau worden hiertoe de individuele scores gebruikt. Op afdelings- en sector-niveau worden de desbetreffende groepsgemiddelden gebruikt. De keuze en operationalisatie van deze variabelen is gebaseerd op het beknopte model uit hoofdstuk 7. Dit beknopte model bevat de volgende PSA-variabelen: "psychologische taakeisen" (gebaseerd op de schalen "werktempo/werkhoeveelheid" en "emotionele belasting"), "veelzijdigheid" (gebaseerd op de schalen "gebrek aan

taakvariatie" en "gebrek aan leermogelijkheden"), "regelmogelijkheden" (gebaseerd op de schalen "gebrek aan taakautonomie" en "gebrek aan inspraak"), en "sociale ondersteuning" (gebaseerd op de schalen "problemen in de relatie met collega's en leiding").

Als confounders zullen worden onderzocht: NA en EO (zie 8.1 en 8.2). Deze variabelen worden uitsluitend op het niveau van de individuele werknemer in de vergelijking opgenomen.

Voor de keuze van de afhankelijke variabelen is eveneens teruggegrepen op het beknopte model uit hoofdstuk 7. De twee dimensies "spanning" (gebaseerd op de schalen "herstelbehoefte" en "piekeren") en "welbevinden" (gebaseerd op de schalen "opzien" en "betrokkenheid") zijn meegenomen in de multi-level analyse.

Alle labels van de vragenlijstschalen zijn zo kort en eenvoudig mogelijk gehouden. Dit maakt het nodig om voorafgaand aan de multi-level analyse duidelijk aan te geven in welke richting hoge en lage scores dienen te worden geïnterpreteerd. In tabel 8.1 wordt voor elk van de zes PSA- en werkstress-dimensies aangegeven wat het label is, de betekenis van een lage dimensiescore, en de betekenis van een hoge dimensiescore.

Tabel 8.1: Interpretatie lage/hoge waarden van de dimensiescores

dimensiescore	lage score	hoge score
psychologische taakeisen	weinig	veel
veelzijdigheid	veel	weinig
regelmogelijkheden	veel	weinig
sociale ondersteuning	veel	weinig
spanning	weinig	veel
welbevinden	veel	weinig

De multi-level analyse is qua uitvoering gemodelleerd op een hiërarchische multiële regressie-techniek. In de analyse wordt een aantal stappen onderscheiden. We bespreken deze stappen hier kort, omdat het om een nieuwe en nog betrekkelijk onbekende statistische techniek gaat.

- stap 0 : leeg model
Er zijn nog geen onafhankelijke variabelen ingevoerd in de vergelijking. De bruto verdeling van de variantie over de drie niveaus (individu, afdeling, sector) wordt verkregen.
- stap 1 : kenmerken op individuniveau
Van hier af aan worden delen van de variantie op elk van de drie niveaus verklaard. Als eerste worden biografische kenmerken van de individuele werknemers ingevoerd in de vergelijking.
- stap 2 : confounders op individuniveau
Ten tweede worden NA en EO ingevoerd in de vergelijking. Binnen stap 2 worden telkens drie afzonderlijke mogelijkheden onderzocht:
- stap 2a : alleen NA
- stap 2b : alleen EO
- stap 2c : beide confounders
Door 2a versus 2c en 2b versus 2c te vergelijken, kunnen we zien hoe groot het confounding-effect is van NA en EO, relatief ten opzichte van elkaar. Na stap 2 zijn we toe aan de verklaring van variantie (op elk van de drie niveaus gelijktijdig) door PSA.
- stap 3 : psychosociale arbeidsbelasting op individuniveau
De dimensiescores voor de vier PSA-dimensies van de individuele werknemers (zoals we die eerder hebben gebruikt in paragraaf 7.8) worden hier ingevoerd in de vergelijking. Tot hier aan toe is de analyse vergelijkbaar met andere multivariate technieken, met uitzondering van het feit dat de verklaarde variantie wordt uitgesplitst per niveau.
- stap 4 : kenmerken op afdelingsniveau
Nu worden kenmerken op afdelingsniveau ingevoerd in de vergelijking. In dit geval betreft dit één dichotome variabele, die aangeeft of de afdeling al dan niet werkt met patiëntencontact.
- stap 5 : psychosociale arbeidsbelasting op afdelingsniveau
Deze stap lijkt op stap 3, maar nu worden de per afdeling de geaggregeerde dimensiescores (gemiddelden) ingevoerd in de vergelijking.
- stap 6 : kenmerken op sectorniveau
Deze stap komt overeen met stap 4, maar nu betreft het een sector-kenmerk. Dit is opnieuw een dichotome variabele, die aangeeft of de sector een dienstverlenende danwel een intern ondersteunende taak heeft.
- stap 7 : psychosociale arbeidsbelasting op sectorniveau
Deze stap is vergelijkbaar met de stappen 3 en 5, maar nu zijn het de geaggregeerde dimensiescores (gemiddelden) op sectorniveau, die in de vergelijking worden ingevoerd.

Alvorens te starten met de analyse zijn alle ordinale en continue variabelen gestandaardiseerd.

Voor de volledigheid volgt in tabel 8.2 een overzicht van alle onafhankelijke en afhankelijke variabelen in de multi-level analyse. Telkens wordt vermeld wat het meetniveau van de variabele is, en wat de variabele betekent cq. waaruit de variabele is berekend. In het maximale model zitten 21 onafhankelijke variabelen. Er zijn twee afhankelijke variabelen onderzocht. De afhankelijke variabelen kunnen bij multi-level analyse alleen op het laagste niveau worden onderzocht. In dit geval is dat het niveau van de individuele werknemer. De ordinale variabelen worden in de analyse behandeld als intervalvariabelen. Omdat het steeds gaat om samengestelde schalen en/of dimensies, is dit gerechtvaardigd.

Tabel 8.2: Overzicht van variabelen in de multi-level analyse

afhankelijke variabelen (1 niveau: 411 individuen)		
variabele	meetniveau	betekenis/samenstelling
spanning	ordinaal	somscore herstelbehoefte en piekeren
welbevinden	ordinaal	somscore opzien en betrokkenheid

onafhankelijke variabelen (3 niveaus: 411 individuen, 44 afdelingen, 8 sectoren)		
1. Biografische kenmerken individueel niveau (N=411): geslacht dienstjaren leeftijd uren werk per week func- tieniveau	dichotoom/nominaal continu continu continu dichotoom/nominaal	man/vrouw aantal jaren aantal jaren aantal uren laag (tot HBO) en hoog (HBO of hoger)
2. Confounders individueel niveau (N=411): NA EO	ordinaal ordinaal	schaal emotionele reacties twee items uit schaal beloning

Tabel 8.2: Overzicht van variabelen in de multi-level analyse (vervolg)

variabele	meetniveau	betekenis/samenstelling
3. PSA individueel niveau (N=411): psychologische taakeisen	ordinaal	somscore werktempo/werkhoeveelheid en emotionele belasting
veelzijdigheid	ordinaal	somscore gebrek aan taakvariatie en gebrek aan leermogelijkheden
regelmogelijkheden	ordinaal	somscore gebrek aan taakautonomie en gebrek aan inspraak
sociale ondersteuning	ordinaal	somscore problemen in relatie tot collega's en leiding
4. Beschrijvende kenmerken afdelingsniveau (N=44): type afdeling	dichotoom/nominaal	al dan niet direct patiëntcontact
5. PSA afdelingsniveau (N=44): psychologische taakeisen	ordinaal	geaggregeerde somscore werkdruk en emotionele belasting
veelzijdigheid	ordinaal	geaggregeerde somscore gebrek aan taakvariatie en gebrek leermogelijkheden
regelmogelijkheden	ordinaal	geaggregeerde somscore gebrek aan taakautonomie en gebrek aan inspraak
sociale ondersteuning	ordinaal	geaggregeerde somscore problemen in relatie collega's en leiding
6. Beschrijvende kenmerken sectorniveau (N=8): type sector	dichotoom/nominaal	ondersteunend dan wel dienstverlenend
7. PSA sectorniveau (N=8): psychologische taakeisen	ordinaal	geaggregeerde somscore werkdruk en emotionele belasting
veelzijdigheid	ordinaal	geaggregeerde somscore gebrek aan taakvariatie en gebrek leermogelijkheden
regelmogelijkheden	ordinaal	geaggregeerde somscore gebrek aan taakautonomie en gebrek aan inspraak
sociale ondersteuning	ordinaal	geaggregeerde somscore problemen in relatie collega's en leiding

8.5 Resultaten multi-level analyse

De resultaten vallen uiteen in drie delen. In de eerste plaats bekijken we de verdeling van de totale variantie over de drie onderscheiden analyseniveaus in het lege model. Dit geeft een indruk van het belang van de drie onderscheiden niveaus ten aanzien van de afhankelijke variabele.

Vervolgens bekijken we de verklaarde varianties per stap in de analyse. Dit geeft aan wat het verklarend vermogen is van de diverse stappen in de analyse. Nu doet zich bij multi-level analyse de vervelende complicatie voor, dat het toevoegen van onafhankelijke variabelen aan een model kan leiden tot het toenemen van de onverklaarde variantie op één van de niveaus. Dit komt omdat de door de diverse niveaus verklaarde varianties elkaar beïnvloeden.

Snijders en Bosker (1994) stellen een alternatieve berekeningswijze voor van de proportie verklaarde variantie (R-kwadraat), om negatieve R-kwadraten te vermijden. Deze alternatieve methode wordt in dit hoofdstuk toegepast. Om de methode van Snijders en Bosker toe te passen in het geval van een ongebalanceerd design (ongelijke aantallen werknemers in de afdelingen en ongelijke aantallen afdelingen in de sectoren, waarvan hier inderdaad sprake is), is het nodig een representatief getal te kiezen voor het aantal werknemers per afdeling, en voor het aantal afdelingen per sector. Hierbij moet de populatie, en niet de steekproef, als uitgangspunt worden genomen.

Omdat bij de instelling waar de multi-level analyse wordt toegepast exacte gegevens bekend zijn over de uitgenodigde onderzoekspopulatie en de respondende steekproef, zijn de benodigde representatieve getallen bepaald door gemiddelden te berekenen op basis van de uitgenodigde populatie. Het gemiddeld aantal werknemers per afdeling bedraagt dan 16.5. Het gemiddeld aantal afdelingen per sector is 6.4.

Na de verdeling van de ruwe variantie en de verdeling van de verklaarde variantie, worden tenslotte de t-waarden gepresenteerd van de afzonderlijke variabelen in het maximale model. Dit zegt ons wat de significantie is van de afzonderlijke onafhankelijke variabelen voor een bepaalde afhankelijke variabele, wanneer wordt gecontroleerd voor alle andere onafhankelijke variabelen in het model.

8.5.1 Verdeling van de variantie in het lege model

In tabel 8.3 zien we per afhankelijke variabele de verdeling van de variantie over de drie onderscheiden niveaus in het lege model.

Tabel 8.3: Verdeling van de variantie in het lege model

variabele	individu	afdeling	sector
spanning	91.0%	2.3%	6.8%
welbevinden	81.4%	15.0%	3.6%

Uit deze gegevens wordt duidelijk dat het individuele niveau het analyseniveau is waar de meeste variantie valt te verklaren. Dit is niet verbazingwekkend. Uit onderzoek bij scholen en leerlingen is als vuistregel af te leiden dat bij multi-level analyse doorgaans ongeveer 10% van de variantie op een hoger verklaringsniveau ligt dan dat van het individu (Bosker en Scheerens, 1989). Een vergelijkbare vuistregel voor bedrijven en werknemers bestaat nog niet. Tegen deze achtergrond is er een opvallend verschil in het resultaat bij "spanning" versus "welbevinden".

De variantie in "arbeidsgerelateerde spanning" ligt voor 9% op een boven-individueel niveau. Dit is zo'n beetje wat we van een multi-level analyse verwachten (Bosker en Scheerens, 1989). Opvallend is dat bij "welbevinden" 19% van de te verklaren variantie ligt op boven-individueel niveau. Dit is een relatief groot deel van de te verklaren variantie. Bij "welbevinden" is bovendien sprake van een afname van de te verklaren variantie met de toename van het analyseniveau. Dit is conform de verwachting. Waarom dit bij "spanning" niet zo is, is onduidelijk: hier ligt iets meer te verklaren variantie op sectorniveau dan op afdelingsniveau. Wellicht kan dit voor een deel worden verklaard door het gegeven, dat op sectorniveau meer balans bestaat tussen het aantal primair dienstverlenende en het aantal primair facilitaire organisatie-onderdelen, dan op afdelingsniveau. Op afdelingsniveau zijn de primair dienstverlenende afdelingen duidelijk groter in aantal. Verder is, vanwege het kleine aantal sectoren, de schatting van de variantie op sectorniveau waarschijnlijk niet erg nauwkeurig.

Een meer directe manier om te verwoorden wat hierboven staat, is dat over "arbeidsgerelateerd welbevinden" meer consensus bestaat binnen groepen werknemers dan over "arbeidsgerelateerde spanningsverschijnselen". Dit is een belangrijk resultaat, los van wat verklaard kan worden aan variantie in deze afhankelijke variabelen op de diverse niveaus.

De multi-level analyse levert al iets op, voordat nog maar één onafhankelijke variabele in de vergelijking is ingevoerd. Dit resultaat is volledig toe te schrijven aan de specificatie van de nestingsstructuur.

8.5.2 Verklaarde variantie per stap

We bekijken dit eerst met "spanning" als afhankelijke variabele. De resultaten zijn weergegeven in de tabellen 8.4.A t/m C.

In tabel 8.4.A is per stap de unieke bijdrage van die stap weergegeven aan de variantie op elk van de drie niveaus. Het gaat telkens om de verandering in de R-kwadraat ten opzichte van de vorige stap. Hierbij is telkens uitgegaan van de totale te verklaren variantie als basis (100%). Let wel: bij deze multi-level analyse valt op elk van de drie niveaus 100% van de variantie te verklaren! De verandering in R-kwadraat moet dan ook uitsluitend worden geïnterpreteerd binnen het desbetreffende verklaringsniveau (individueel, afdeling of sector).

Stap 2a, 2b, en 2c worden alle drie bepaald ten opzichte van stap 1. Stap 3 wordt bepaald ten opzichte van stap 2c. Tabel 8.4.A presenteert per stap, per niveau de verklaarde variantie, en is dus niet cumulatief. In tabel 8.4.B zijn dezelfde gegevens cumulatief weergegeven.

Om te toetsen in hoeverre het introduceren van een bepaalde onafhankelijke variabele bijdraagt aan het verklaren van de variantie over alle drie de niveaus heen, wordt bij multi-level analyse gebruik gemaakt van de zogenaamde deviance-maat. Het verschil in deviance tussen twee stappen in de analyse gedraagt zich als een parameter met een Chi-kwadraatverdeling, met evenveel vrijheidsgraden als er variabelen zijn geïntroduceerd bij de nieuwe analysestap.

In tabel 8.4.C staan per stap: de deviance (Dev.), het verschil in deviance met de vorige stap (Chi), het aantal vrijheidsgraden, en het resultaat van een significantietoets op het verschil in deviance met de vorige stap.

In tabel 8.4.A valt op dat, ondanks de door Snijders en Bosker (1994) voorgestelde berekeningswijze, toch enkele negatieve R-kwadraten voorkomen. Snijders en Bosker (1994) geven zelf aan, dat dit nog steeds kan voorkomen. In tabel 8.4.A is goed te zien, dat in een multi-level analyse bij elke stap de verklaarde variantie verandert op elk van de drie niveaus tegelijkertijd. Het invoeren van NA in de vergelijking, levert bijvoorbeeld een verbetering op in de R-kwadraat van 24.5% op individueel niveau, 43.5% op afdelingsniveau, en 40.1% op sectorniveau. Na deze inleidende opmerkingen is het nu tijd voor de resultaten zelf.

In paragraaf 8.5.1. werd gevonden dat 91% van de ruwe variantie zich bevindt op het niveau van de individuele medewerker. Hier staan we dan ook het eerst bij stil. Op dit individuele verklaringsniveau kan uiteindelijk 43% van de variantie worden verklaard in "spanning" (zie het getal bij stap 7 in tabel 8.4.B). De stappen met de grootste verandering in R-kwadraat (zie tabel 8.4.A) zijn: NA (24.5%) en PSA (individuele scores: 11.4%). De individuele en geaggregeerde PSA-gegevens verklaren gezamenlijk 12.4% van de variantie op het individuele verklaringsniveau (11.4% + 0.4% + 0.6%).

*Tabel 8.4.A: Stapsgewijs verklaarde variantie (R-kwadraat)
per niveau - niet cumulatief*

Y: Spanning			
Niveau:	Individu	Afdeling	Sector
stap 1 : biografische kenmerken	4.9%	-7.9%	-2.1%
stap 2a : NA	24.5%	43.5%	40.1%
stap 2b : EO	0.8%	-2.2%	-1.5%
stap 2c : confounders samen	25.5%	46.1%	42.9%
stap 3 : PSA individu	11.4%	37.5%	52.3%
stap 4 : kenmerken afdeling	0.1%	0.0%	0.0%
stap 5 : PSA afdeling	0.4%	0.2%	0.0%
stap 6 : kenmerken sector	0.1%	0.1%	0.0%
stap 7 : PSA sector	0.6%	0.3%	0.1%

*Tabel 8.4.B: Stapsgewijs verklaarde variantie (R-kwadraat)
per niveau - cumulatief*

Y: Spanning			
Niveau:	Individu	Afdeling	Sector
stap 1 : biografische kenmerken	4.9%	-7.9%	-2.1%
stap 2a : NA	29.3%	35.6%	37.9%
stap 2b : EO	5.7%	-10.1%	-3.6%
stap 2c : confounders samen	30.3%	38.2%	40.8%
stap 3 : PSA individu	41.8%	75.8%	93.1%
stap 4 : kenmerken afdeling	41.8%	75.8%	93.1%
stap 5 : PSA afdeling	42.2%	76.0%	93.1%
stap 6 : kenmerken sector	42.4%	76.0%	93.1%
stap 7 : PSA sector	43.0%	76.3%	93.2%

Tabel 8.4.C: Significantie-toets per stap (over alle drie de niveaus gezamenlijk)

Y: Spanning

	Dev.	Chi	Df	Sign.
stap 0 : leeg model	1149.46	nvt	nvt	nvt
stap 1 : biografische kenmerken	1124.26	25.20	5	<.001
stap 2a : NA	1010.67	113.59	1	<.001
stap 2b : EO	1119.16	5.10	1	.024
stap 2c : confounders samen	1005.82	118.44	2	<.001
stap 3 : PSA individu	947.56	58.26	4	<.001
stap 4 : kenmerken afdeling	947.10	0.46	1	.498
stap 5 : PSA afdeling	944.25	2.85	4	.583
stap 6 : kenmerken sector	943.37	0.88	1	.348
stap 7 : PSA sector	939.00	4.37	4	.358

Op afdelingsniveau stijgt de verklaarde variantie op basis van het maximale model tot 76.3% (zie tabel 8.4.B). Opnieuw zijn NA (43.5%) en PSA (individuele scores: 37.5%) de stappen die de grootste verandering in R-kwadraat laten zien. Op sectorniveau is de verklaarde variantie uiteindelijk maar liefst 93.2%. Ook hier zijn PSA (individuele scores: 52.3%) en NA (40.1%) de stappen met grootste verandering in R-kwadraat.

Tabel 8.4.A en 8.4.B leren ons dat het merendeel van de variantie op bovenindividueel niveau verklaard kan worden met dit model. Omdat de hoeveelheid "ruwe variantie" op bovenindividueel niveau relatief klein bleek te zijn, leidt dit automatisch tot de conclusie, dat de onverklaarde variantie zich nagenoeg uitsluitend bevindt op individueel niveau.

In tabel 8.4.C worden de verklaarde varianties voor alle drie de niveaus gezamenlijk getoetst per stap. Hier is overigens te zien dat de deviance netjes trapsgewijs afloopt: elke stap levert een lagere deviance op dan de vorige.

Over de drie niveaus gezamenlijk treedt er dus geen negatieve verklaarde variantie op. Uit deze tabel blijkt dat vooral de stappen NA en PSA (individuele scores) significant bijdragen aan de verklaarde variantie, zoals op grond van het voorgaande te verwachten was. Het effect van NA is duidelijk het grootst.

Ook de individuele biografische kenmerken blijken gezamenlijk een significante proportie van de variantie te verklaren. De bijdrage van EO levert eveneens een lichte significantie op. De overige stappen in de analyse dragen nauwelijks meer bij aan de verklaarde variantie. De geaggregeerde kenmerken en PSA-gegevens voegen dus nauwelijks meer iets toe. Dezelfde gegevens volgen nu voor "arbeids-gerelateerd welbevinden" in de tabellen 8.5.A t/m C.

Ook in tabel 8.5.A zien we negatieve R-kwadraten, ondanks de toegepaste berekeningswijze. Er wordt zelfs een waarde van -23% gevonden bij stap 3 op sectorniveau.

Op het individuele niveau kunnen we met het maximale model 42.2% van de variantie verklaren in "welbevinden" (zie stap 7 in tabel 8.5.B). De stappen die hieraan de grootste bijdrage leveren zijn PSA (individuele scores: 18.4%), NA (9.6%), en individuele biografische kenmerken (5.9%). De individuele en geaggregeerde PSA-gegevens verklaren in totaal 21.6% van de variantie op individu-niveau (18.4% + 1.9% + 1.3%).

*Tabel 8.5.A: Stapsgewijs verklaarde variantie (R-kwadraat)
per niveau - niet cumulatief*

Y: Welbevinden			
Niveau:	Individu	Afdeling	Sector
stap 1 : biografische kenmerken	5.9%	18.0%	37.3%
stap 2a : NA	9.6%	13.9%	15.4%
stap 2b : EO	1.2%	2.4%	-6.0%
stap 2c : confounders samen	10.8%	17.2%	10.7%
stap 3 : PSA individu	18.4%	14.7%	-22.8%
stap 4 : kenmerken afdeling	3.7%	17.8%	43.9%
stap 5 : PSA afdeling	1.9%	7.7%	15.7%
stap 6 : kenmerken sector	0.1%	0.7%	0.7%
stap 7 : PSA sector	1.3%	5.8%	4.3%

*Tabel 8.5.B: Stapsgewijs verklaarde variantie (R-kwadraat)
per niveau - cumulatief*

Y: Welbevinden			
Niveau:	Individu	Afdeling	Sector
stap 1 : biografische kenmerken	5.9%	18.0%	37.3%
stap 2a : NA	15.5%	31.9%	52.8%
stap 2b : EO	7.2%	20.4%	31.4%
stap 2c : confounders samen	16.7%	35.2%	48.1%
stap 3 : PSA individu	35.1%	49.9%	25.3%
stap 4 : kenmerken afdeling	38.8%	67.6%	69.2%
stap 5 : PSA afdeling	40.7%	75.3%	85.0%
stap 6 : kenmerken sector	40.8%	76.0%	85.6%
stap 7 : PSA sector	42.2%	81.8%	90.0%

Tabel 8.5.C: Significantie-toets per stap (over alle drie de niveaus gezamenlijk)

Y: Welbevinden				
	Dev.	Chi	Df	Sign.
stap 0 : leeg model	1126.95	nvt	nvt	nvt
stap 1 : biografische kenmerken	1113.02	13.93	5	.016
stap 2a : NA	1074.10	38.92	1	<.001
stap 2b : EO	1108.58	4.44	1	.035
stap 2c : confounders samen	1070.20	42.82	2	<.001
stap 3 : PSA individu	964.82	105.38	4	<.001
stap 4 : kenmerken afdeling	959.08	5.74	1	.017
stap 5 : PSA afdeling	951.89	7.19	4	.126
stap 6 : kenmerken sector	951.45	0.44	1	.507
stap 7 : PSA sector	944.90	6.55	4	.162

Een relatief groot deel van de ruwe variantie bij "welbevinden" zit op het afdelingsniveau, zo werd duidelijk in paragraaf 8.5.1. Hier gaat dan ook onze speciale belangstelling naar uit. Op dit niveau kunnen we 81.8% van de variantie verklaren. De opbouw van deze variantie laat zien, dat individuele biografische kenmerken, kenmerken van de afdeling, PSA (individuele scores), NA, en PSA (afdelingsscores) allen een belangrijke bijdrage hieraan leveren. De PSA-gegevens gezamenlijk verklaren 28.2% van de variantie op afdelingsniveau (14.7% + 7.7% + 5.8%).

Het sectorniveau, waar 90% van de variantie kan worden verklaard met het maximale model, laat ongeveer dezelfde resultaten zien als het afdelingsniveau. Vreemd is de negatieve R-kwadraat bij PSA (individuele scores).

Omdat het patroon van de verklaarde variantie per niveau nogal verschillend is, is het totale resultaat bij "welbevinden" wat moeilijker te overzien dan bij "spanning". In tabel 8.5.C staan de gegevens van de toetsing van de verklaarde variantie over de drie niveaus gezamenlijk, aan de hand van de deviance-maat. PSA (individuele scores) en NA blijken ook bij "welbevinden" de belangrijkste voorspellers te zijn. Ditmaal is de bijdrage van PSA (individuele scores) echter duidelijk het belangrijkste. Ook de biografische kenmerken, EO, en de afdelingskenmerken zijn in lichte mate significant. De geaggregeerde PSA-gegevens voegen, net als bij "spanning" weinig toe.

Samengevat levert de multi-level analyse van de variantiecomponenten het beeld op van "spanning" als een meer individugebonden fenomeen, waarbij confounding door NA inderdaad een belangrijk probleem is voor onderzoek naar de relatie tussen PSA en "spanning". Desalniettemin levert PSA (in de vorm van de individuele scores) bij "spanning" een flinke eigen bijdrage aan de verklaarde variantie. "Welbevinden" daarentegen komt naar voren als een in sterkere mate organisatiegebonden fenomeen, waarbij confounding door NA ook een rol speelt, maar een kleinere. De variantie in "welbevinden" kan vooral worden verklaard op basis van PSA (opnieuw in de vorm van de individuele scores).

8.5.3 T-Waarden

Bij het maximale model zijn de t-waarden berekend voor de 21 onafhankelijke variabelen, het bijbehorende aantal vrijheidsgraden, en de significantie. Deze rapporteren we nu voor "spanning" en "welbevinden" afzonderlijk. Gaf de vorige paragraaf inzicht in de bijdrage van de afzonderlijke stappen aan de vergelijking, nu kunnen we dit toespitsen tot de afzonderlijke onafhankelijke variabelen.

De t-waarden worden berekend door de schattingen voor de regressie-coëfficiënten (Est.) van de onafhankelijke variabelen te delen door hun standaardmeetfout (St.Err). De significantie-grenzen zijn hierbij bepaald op $p < .10$ (*), $p < .05$ (**), en $p < .01$ (***). Eerst komen in tabel 8.6 de t-waarden van "spanning" aan bod.

Tabel 8.6: T-waarden spanning

X:	Est.	St.Err.	T	Df	Sign.
geslacht	-.014	.097	-.14	399	ns
dienstjaren	-.095	.045	-2.11	399	**
leeftijd	.010	.048	2.09	399	**
uren werk per week	.104	.046	2.28	399	**
functieniveau	.099	.124	.79	399	ns
NA	.380	.043	8.95	399	***
EO	.025	.043	.59	399	ns
psychologische eisen ind.	.312	.058	5.39	399	***
veelzijdigheid ind.	-.042	.059	-.73	399	ns
regelmogelijkheden ind.	-.035	.048	-.74	399	ns
sociale ondersteuning ind.	.108	.050	2.15	399	**
type afdeling	.213	.201	1.06	38	ns
psychologische eisen afd.	-.039	.105	-.37	38	ns
veelzijdigheid afd.	-.011	.102	-.11	38	ns
regelmogelijkheden afd.	.139	.096	1.45	38	ns
sociale ondersteuning afd.	-.114	.085	-1.34	38	ns
type sector	-.213	.155	-1.37	2	ns
psychologische eisen sector	-.008	.126	-.06	2	ns
veelzijdigheid sector	-.379	.260	-1.46	2	ns
regelmogelijkheden sector	.354	.240	1.48	2	ns
sociale ondersteuning sector	-.014	.111	-.13	2	ns

Tabel 8.6 maakt duidelijk dat "arbeidsgerelateerde spanningsverschijnselen" voornamelijk kunnen worden begrepen vanuit NA en "psychologische taakeisen" op individueel niveau. Andere variabelen die significantie bereiken zijn: dienstjaren, leeftijd, uren per week, en "sociale ondersteuning" op individueel niveau. Vervolgens worden in tabel 8.7 de t-waarden gepresenteerd voor "welbevinden".

Tabel 8.7 laat een heel ander beeld zien voor "welbevinden" dan voor "spanning". De belangrijkste voorspeller is nu de "veelzijdigheid" van het werk (individueel niveau), gevolgd door respectievelijk: "sociale ondersteuning" op individueel niveau, NA, en dienstjaren. Verder leveren ook het type afdeling en het afdelingsoordeel over "regelmogelijkheden" een significante bijdrage.

Tabel 8.7: *T-waarden welbevinden*

X:	Est.	St.Err.	T	Df	Sign.
geslacht	-.132	.098	-1.35	399	ns
dienstjaren	-.130	.046	-2.86	399	***
leeftijd	-.039	.048	-.81	399	ns
uren werk per week	.035	.046	.76	399	ns
functieniveau	.091	.126	.72	399	ns
NA	.184	.043	4.32	399	***
EO	.034	.043	.80	399	ns
psychologische eisen ind.	.010	.058	.17	399	ns
veelzijdigheid ind.	.390	.059	6.66	399	***
regelmogelijkheden ind.	-.004	.048	-.09	399	ns
sociale ondersteuning ind.	.225	.050	4.47	399	***
type afdeling	.445	.214	2.08	38	**
psychologische eisen afd.	-.125	.110	-1.14	38	ns
veelzijdigheid afd.	-.178	.108	-1.64	38	ns
regelmogelijkheden afd.	.170	.101	1.68	38	*
sociale ondersteuning afd.	-.023	.089	-.26	38	ns
type sector	-.030	.168	-.18	2	ns
psychologische eisen sector	.047	.134	.35	2	ns
veelzijdigheid sector	-.532	.280	-1.90	2	ns
regelmogelijkheden sector	.558	.261	2.14	2	ns
sociale ondersteuning sector	-.16	.12	-1.35	2	ns

De resultaten van de t-waarden (tabel 8.6 en 8.7) stemmen volledig overeen met de resultaten van de deviance-maat, zoals per analysestap gerapporteerd in tabel 8.4.C en 8.5.C. De t-waarden geven informatie over de bijdrage van de afzonderlijke variabelen in de vergelijking, wat derhalve meer detail oplevert dan de deviance.

Wat betreft de biografische kenmerken valt vooral het aantal dienstjaren op: zowel bij "spanning" als bij "welbevinden" is deze variabele significant. NA speelt eveneens een rol bij zowel "spanning" als "welbevinden". Bij de individuele PSA-scores levert alleen "sociale ondersteuning" een bijdrage aan zowel "spanning" als "welbevinden".

De PSA-factoren "psychologische taakeisen" en "veelzijdigheid" (individuele scores) hangen gedifferentieerd samen met "spanning" en "welbevinden". De enige variabelen die verder nog met "spanning" samenhangen zijn de biografische kenmerken leeftijd en aantal uren werk per week. Alleen bij "welbevinden" worden significanties gevonden bij geaggregeerde variabelen, namelijk bij het type afdeling en de afdelingsscore van "regelmogelijkheden".

8.6 Samenvatting en conclusie multi-level analyse

Samengevat levert de multi-level analyse de volgende resultaten op. Er is meer variantie op boven-individueel niveau in "welbevinden" dan in "spanning". De verklaarde variantie in "spanning" en "welbevinden" met behulp van multi-level analyse, op grond van gegevens van dertien verschillende onafhankelijke variabelen, bedraagt iets meer dan 40% op individueel niveau. Op boven-individueel niveau ligt de verklaarde variantie telkens boven de 75%. Bij zowel "spanning" als "welbevinden" blijft er uiteindelijk nagenoeg alleen onverklaarde variantie over op individueel niveau. Bijna alle variantie op het niveau van afdelingen en sectoren kan worden verklaard. Opmerkelijk is dat bij "spanning" ook de boven-individuele variantie voor het grootste deel wordt verklaard door kenmerken en vragenlijstscores van individuele medewerkers. Bij "welbevinden" verklaren de onafhankelijke variabelen op afdelings- en sectorniveau (groepskenmerken en geaggregeerde vragenlijstgegevens) een groter deel van de boven-individuele variantie.

De belangrijkste voorspellers van "spanning" en "welbevinden" zijn NA en PSA (individuele scores). Bij "spanning" is NA de belangrijkste voorspeller, bij "welbevinden" is dat PSA (individuele scores).

Confounding door NA is inderdaad een belangrijk probleem voor onderzoek naar de relatie tussen PSA en "spanning". Confounding door NA is voor onderzoek naar de relatie tussen PSA en "welbevinden" een minder groot probleem. Voor een confounding-effect van EO is geen evidentie gevonden.

PSA levert bij "spanning" en bij "welbevinden" een belangrijke, eigen bijdrage aan de verklaarde variantie, bovenop de variantie die wordt verklaard door biografische kenmerken en confounders. Bij "spanning" is deze bijdrage - op het individuele verklaringsniveau - iets meer dan 10%, wat in de verwachte orde van grootte is. Voor "welbevinden" is deze bijdrage hoger dan verwacht: iets meer dan 20%.

Van de PSA-dimensies hangt "psychologische taakeisen" vooral samen met "spanning" en "veelzijdigheid" vooral met "welbevinden". "Sociale ondersteuning" hangt samen met zowel "spanning" als "welbevinden".

Teruggebracht tot twee zinnen: Uit de multi-level analyse komt "spanning" naar voren als een meer individugebonden fenomeen, waarbij, naast de individuele emotionele toestand van de werknemer, de individuele beleving van "psychologische taakeisen" een belangrijke rol speelt.

"Welbevinden" blijkt in vergelijking met "spanning" een meer groeps- en organisatiegebonden fenomeen te zijn, waarbij de "veelzijdigheid" van het werk een sleutelrol speelt, en waarbij naast de individuele beleving van werknemers ook de collectieve beleving van groepen werknemers van belang is.

Wat betekenen deze resultaten voor de aan het begin van dit hoofdstuk gestelde vragen?

Confounding

Het blijkt dat vooral NA een belangrijke confounder is, die inderdaad een verstorende rol kan spelen bij onderzoek naar de relatie tussen PSA en werkstress (Watson en Pennebaker, 1989; Brief et al, 1988). Deze verstorende rol is groter bij maten voor "spanning" dan bij maten voor "welbevinden". Dit is helemaal overeenkomstig het resultaat van Parkes et al. (1994). Bovenop de personalia leeftijd, geslacht en functietype verklaarde neuroticisme in deze studie 9% in een maat voor arbeidssatisfactie, en 19% in een maat voor psychosomatische gezondheidsklachten. Dit stemt redelijk overeen met het hier gevonden resultaat. Brief et al. (1988) vonden een analoog resultaat, namelijk dat NA een sterker verstrend effect heeft op de relatie van stressoren met gezondheidsklachten en emoties, dan op de relatie van stressoren met arbeidssatisfactie. Ook de resultaten van Chen en Spector (1991) bevestigen dit.

Het confounding-effect van NA werd hier aangetoond, ondanks het feit dat NA meer als "state" dan als "trait" is geoperationaliseerd in dit onderzoek. Ook de emotionele toestand van een werknemer (in tegenstelling tot de emotionele dispositie) kan dus een gelijktijdige beoordeling van PSA en werkstress met vragenlijsten vertekenen.

Opvallend is dat EO ten aanzien van beloning, ondanks een vroege entry in de vergelijking, slechts weinig bijdraagt aan de verklaarde variantie. Dit is niet conform de verwachting die we hadden op grond van studies zoals bijvoorbeeld gedaan naar de relatie tussen EO, arbeidsbelasting en burnout (Buunk en Schaufeli, 1993; Van Dierendonk en Schaufeli, 1993; Van Gorp et al., 1993; Schaufeli et al., 1994).

Een belangrijk resultaat voor het type onderzoek waarmee we hier bezig zijn is, dat PSA zowel bij "spanning" als bij "welbevinden" een belangrijke en unieke component aan de verklaarde variantie bijdraagt, bovenop de door confounders verklaarde variantie. Confounding treedt dus wel op, maar onderzoek naar de invloed van PSA op werkstress blijft z'n eigen betekenis houden, zowel ten aanzien van "spanning" als ten aanzien van "welbevinden".

De limiet van 10% à 15% verklaarde variantie

Voor individuele, arbeidsgerelateerde "spanning" blijkt de stelling van Semmer, Zapf en Greif (1995) dat 10% tot 15% verklaarde variantie op basis van PSA het theoretisch haalbare maximum is, redelijk op te gaan. Voor individueel, arbeidsgerelateerd "welbevinden" komen we in dit onderzoek echter hoger uit, namelijk rond de 20%. Het resultaat bij "welbevinden" komt overeen met hetgeen Semmer, Zapf en Greif (1995) vinden met hun concept "shared well-being": de verklaarde variantie komt hierbij in de buurt van de 20%.

Een verwante claim in het artikel van Semmer, Zapf en Greif (1995) is, dat met relatief weinig arbeids- en organisatievariabelen het merendeel van de boven-individuele variantie kan worden verklaard (in hun terminologie: "gedeelde variantie"). Deze stelling wordt door dit onderzoek bevestigd: nagenoeg alle onverklaarde variantie zit op individueel niveau.

Een opmerking is hier op zijn plaats wat betreft het boven-individuele niveau. De analyse bevat hier weinig cases (44 afdelingen en 8 sectoren) in verhouding tot het aantal variabelen (5 voor het afdelingsniveau, 5 voor het sectorniveau). Vooral voor het sectorniveau geldt dat er zoveel variabelen zijn, dat die door toeval (kanskapitalisatie) wel bijna alle variantie zullen verklaren. Dat op sectorniveau zo goed als alle verklaarbare variantie wordt verklaard, is derhalve op statistische gronden te verwachten vanwege de verhouding tussen cases en variabelen.

Multi-level analyse is een veelbelovende techniek voor onderzoek in bedrijven en instellingen. Het biedt de mogelijkheid om zonder het schenden van assumpties multivariate analyses te doen naar de relatie tussen PSA en werkstress. De multi-level techniek is eleganter dan alternatieve methoden zoals de "shared well-being" methode van Semmer, Zapf en Greif (1995), omdat de analyse toelaat de bijdrage van verschillende verklaringsniveaus tegelijkertijd en ten opzichte van elkaar te onderzoeken. Juist deze mogelijkheid heeft in dit onderzoek de verschillen tussen de relatie van PSA met "spanning" en de relatie van PSA met "welbevinden" aan het licht gebracht.

Uit dit onderzoek komt tenslotte naar voren, dat het tegelijkertijd invoeren in een analyse van individuele PSA-scores en geaggregeerde versies van dezelfde scores op afdelings- en organisatieniveau, weinig additionele verklaarde variantie oplevert.

Het is, zowel bij "spanning" als bij "welbevinden", vooral de specificatie van de geneste gegevensstructuur zelf, die additionele analyse- en verklaringsmogelijkheden biedt.

Eén of meer werkstress-mechanismen?

Uit de resultaten van dit hoofdstuk is af te leiden, dat er eerder sprake is van een differentiatie in werkstress-mechanismen, dan van één algemeen geldig mechanisme. "Spanning" is vooral een individugebonden fenomeen, waarbij NA een grote rol speelt, naast "psychologische taakeisen". "Welbevinden" is in vergelijking met "spanning" een meer organisatiegebonden fenomeen, waarbij NA een kleinere rol speelt, en waarbij vooral de "veelzijdigheid" van het werk van belang is. Dit resultaat van de multi-level analyse ondersteunt het vitamine-model van Warr (1987; 1994), en sluit minder goed aan bij het demand-control model van Karasek (1979).

De resultaten in dit hoofdstuk zijn op dit punt vergelijkbaar met de resultaten die in paragraaf 7.8 werden gevonden. Opvallend is dat noch in paragraaf 7.8, noch in dit hoofdstuk, een duidelijke relatie van de dimensie "regelmogelijkheden" met "spanning" of "welbevinden" is gevonden. De enige uitzondering hierop is dat bij "welbevinden" de afdelingsscores van "regelmogelijkheden" een significante t-waarde hebben, zij het slechts op een niveau van $<.10$. De resultaten op het gebied van "regelmogelijkheden" zijn onverwacht, mede omdat de analyse van de structuur van PSA en werkstress (hoofdstuk 6) op dit gebied niet meer of minder overtuigend was dan op andere gebieden. Dit resultaat sluit noch bij de theorie van Warr, noch bij de theorie van Karasek goed aan.

De zelfbeoordelingsmethode en de objectiviteit-subjectiviteit controverse

De multi-level analyse levert het inzicht op dat "spanning" en "welbevinden" qua verdeling van de variantie over de geneste verklaringsniveaus individu, afdeling, en organisatie verschillend zijn. Ook de aard van de voorspellers op de diverse verklaringsniveaus is verschillend. Dit resultaat heeft behalve inhoudelijke implicaties ook methodologische implicaties, vooral voor de objectiviteit-subjectiviteit controverse.

We hebben laten zien dat met behulp van een zelfbeoordelingsmethode verschillen aan het licht zijn te brengen tussen onderwerpen die zich vooral afspelen in de individuele belevingssfeer van werknemers, en onderwerpen waarover binnen groepen binnen organisaties tot op zekere hoogte consensus bestaat. De zelfbeoordelingsmethode mag door critici dan wel als "subjectief" worden beschouwd, dit onderzoek laat zien dat dit allerminst mag worden geïnterpreteerd als "onsystematisch". Er worden namelijk systematische verschillen gevonden, afhankelijk van het soort onderwerp, in de mate van subjectiviteit die met een zelfbeoordelingsmethode wordt aangetroffen.

De vele aandacht voor de objectiviteit-subjectiviteit controverse in de PSA- en werkstressliteratuur kan dan ook beter gericht worden op de vraag hoe zelfbeoordelingen tot stand komen en welke systematiek hierin valt te ontdekken. Als deze systematiek helder wordt, is de keuze voor een zelfbeoordelingsmethode meer een

kwestie van de aard van de toepassing (past een zelfbeoordelingsmethode bij het te meten onderwerp?), dan een kwestie van principe (objectieve methoden zijn per definitie beter/slechter). Het resultaat van de multi-level-analyse ondersteunt op dit punt de stellingnames van Spector (1992) en Zapf (1989) (zie 4.1.1).

8.7 Lineaire en non-lineaire effecten

Nu de theorie van Warr (1987; 1994) in de multi-level analyse enige bevestiging heeft gekregen op het punt van de differentiatie in meerdere parallelle werkstress-mechanismen, willen we aan het slot van dit hoofdstuk een ander onderwerp aansnijden dat binnen deze theorie een centrale plek inneemt: de aard van de verbanden tussen PSA en werkstress (zie 2.10).

Warr stelt namelijk dat er drie CE-factoren zijn, die alleen lineaire effecten hebben op werkstress, en waarbij een afvlakking optreedt van het effect bij extreem hoge waarden. De tweede groep van zes factoren die Warr onderscheidt zijn de AD-factoren, waarvan hij non-lineaire relaties veronderstelt met werkstress. Bij hoge waarden treden specifieke "toxische reacties" op. Lage waarden van zowel CE-als AD-factoren gaan gepaard met algemene, negatieve werkstress-effecten.

De drie PSA-dimensies "psychologische taakeisen", "veelzijdigheid", en "regelmogelijkheden" behoren tot de categorie AD-factoren. Een tekort van elk van deze drie factoren zou volgens Warr aanleiding zijn tot allerlei algemene stressreacties. Bij een "overdosis" voorspelt hij specifieke reacties per dimensie. Teveel "psychologische taakeisen" en teveel "regelmogelijkheden" leiden volgens hem tot overbelasting en stress. In de hier ontwikkelde terminologie: zij leiden tot "spanning". Deze veronderstellingen van het vitamine-model zijn goed te onderzoeken met de ontwikkelde structuur voor PSA en werkstress.

Voor de dimensie "veelzijdigheid" is het moeilijk om een eenduidige vertaling te maken van Warr's model naar het in dit proefschrift ontwikkelde begrippenkader. Volgens Warr zal een te groot beroep op iemands vaardigheden (in ons geval te interpreteren als teveel "veelzijdigheid") leiden tot overbelasting en stress. Aldus is er op basis van het vitamine-model een relatie te leggen tussen "veelzijdigheid" en "spanning".

Teveel afwisseling daarentegen (in ons geval opnieuw te interpreteren als teveel "veelzijdigheid") zou volgens Warr leiden tot lage concentratie, lage prestatie, en weinig ontwikkeling in de functie.

Op deze manier is op basis van het vitamine-model ook een relatie te leggen tussen "veelzijdigheid" en "welbevinden". Welke van deze twee aspecten ook het beste aansluit bij onze dimensie "veelzijdigheid", het gaat in beide gevallen om AD-factoren.

In beide gevallen is er sprake van een non-lineaire dosis-effect relatie. Alleen de aard van het door Warr veronderstelde effect is verschillend. Met het beschikbare materiaal kan worden onderzocht of "veelzijdigheid" vooral samenhangt met "spanning" dan wel met "welbevinden". Dit geeft een indicatie over de manier waarop de in dit proefschrift onderscheiden dimensie "veelzijdigheid" in het kader van het vitamine-model moet worden geïnterpreteerd.

De vierde PSA-dimensie "sociale ondersteuning" is evenmin eenduidig te plaatsen in Warr's begrippenkader. Deze dimensie heeft, net als "veelzijdigheid", inhoudelijk verwantschap met twee aspecten die door Warr (1987; 1994) in zijn vitamine-model worden onderscheiden. In het geval van "sociale ondersteuning" zijn deze twee aspecten echter niet van hetzelfde type. De een is een CE-factor, de ander een AD-factor.

In de eerste plaats is er het aspect sociale waardering, dat door Warr wordt gerangschikt onder de CE-factoren. Een tekort aan sociale waardering zou gepaard gaan met algemene negatieve stressgevolgen. Er treden geen "toxische reacties" op wanneer er sprake is van veel sociale waardering.

In de tweede plaats kan "sociale ondersteuning" in verband gebracht worden met het aspect "interpersoonlijk contact", een AD-factor. Ook een tekort hieraan gaat gepaard met algemene stress-gevolgen. Bij een overdosis interpersoonlijk contact treden volgens Warr specifieke "toxische" reacties op, namelijk: verlies van de persoonlijke controle en overbevolkingsreacties.

De specifieke toxische effecten van interpersoonlijk contact zijn binnen de dimensie-structuur van dit proefschrift niet geoperationaliseerd. Er kan derhalve alleen worden bekeken of er algemene negatieve gevolgen zijn van een tekort aan interpersoonlijk contact, net als voor sociale waardering.

Van de dimensie "sociale ondersteuning" is, gezien het voorgaande, niet te verwachten dat deze een non-lineaire relatie heeft met "spanning" of "welbevinden". De beschikbare gegevens kunnen geen indicatie geven over de exacte typering van onze dimensie "sociale ondersteuning" in het kader van het vitamine-model. Wel kan worden onderzocht of het door Warr veronderstelde algemene, negatieve effect van een tekort aan interpersoonlijk contact of sociale waardering wordt bevestigd.

Een volledige test van Warr's vitamine-model is, gezien de beperkte overlap met het in dit proefschrift ontwikkelde begrippenkader, niet mogelijk. Wel kunnen Warr's veronderstellingen over het bestaan van lineaire en non-lineaire relaties tussen PSA en werkstress worden onderzocht voor een belangrijk deel van zijn begrippenkader. Deze paragraaf kan dan ook beter worden beschouwd als een onderzoek naar de aard van de verbanden tussen PSA en werkstress, waarbij het model van Warr fungeert als leidraad.

Alvorens aan de analyse te beginnen is het nodig om eerst stil te staan bij de te hanteren terminologie en de aanpak van de analyse. De terminologie die Warr gebruikt bij het beschrijven van de effecten van CE- en AD-factoren is enigszins verwarrend. Geen van beide typen factoren heeft namelijk een puur lineair effect. Voor CE-factoren wordt een afvlakking beschreven bij hoge waarden: dit is een non-lineair effect. Voor AD-factoren wordt een kromlijning verband beschreven, wat eveneens een non-lineair effect betekent. Het verschil tussen de CE- en AD-factoren zit in het wel of niet monotoon zijn van het non-lineaire effect. Bij de CE-factoren is het effect non-lineair en monotoon stijgend. Bij de AD-factoren is het effect non-lineair en niet-monotoon. In het vervolg van deze paragraaf worden drie typen effect onderzocht: lineair (gelijke toename dosis en effect), non-lineair monotoon, en non-lineair niet-monotoon.

De hier uit te voeren analyse bestaat uit twee onderdelen. Eerst wordt onderzocht voor welke relaties tussen PSA en werkstress sprake is van lineaire en/of non-lineaire relaties (8.7.1). Daarna wordt voor de gevonden non-lineaire relaties nagegaan of er sprake is van monotone dan wel niet-monotone relaties (8.7.2).

8.7.1 Multi-level analyse non-lineaire effecten

In het eerste analyse-onderdeel is exact dezelfde multi-level analyse uitgevoerd, als is beschreven in paragraaf 8.4, maar met een aantal modificaties. Deze modificaties houden in, dat de analyse zich in eerste instantie beperkt tot stap 1 tot en met 3: individuele biografische kenmerken, individuele scores van de confounders, en individuele PSA-scores. Er worden geen groepskenmerken of geaggregeerde gegevens ingevoerd in de analyse. Bij de derde stap worden, net als beschreven in 8.4, de lineaire termen van de PSA-dimensies in de vergelijking ingevoerd.

Vervolgens worden als vierde stap de kwadratische termen van de PSA-dimensies in de vergelijking ingevoerd. Op deze manier wordt geprobeerd bij stap 4 non-lineaire effecten aan te tonen, bovenop de variantie die bij stap 3 reeds is verklaard door de normale, lineaire effecten. Deze werkwijze is afgeleid van Warr (1990a).

De resultaten voor de eerste drie stappen zijn uiteraard identiek aan die in paragraaf 8.5, en worden derhalve hier niet opnieuw gerapporteerd. Voor de vierde analyse-stap worden de volgende resultaten (zie tabel 8.8) gevonden in de deviance-maat (analoog aan de tabellen 8.4.C en 8.5.C):

De vier kwadratische termen leveren gezamenlijk een significante bijdrage aan de verklaarde variantie in zowel "spanning" als "welbevinden". Er is dus sprake van non-lineaire effecten, bovenop de eerder aangetoonde lineaire effecten. Bij "welbevinden" is het effect sterker dan bij "spanning".

Tabel 8.8: Deviance in stap 4 ten opzichte van stap 3

Variabele	Dev.	Chi	Df	Sign.
spanning	935.27	12.29	4	.015
welbevinden	941.38	23.44	4	.000

Hoe is de bijdrage aan de verklaarde variantie verdeeld over de drie verklaringsniveaus? De resultaten van de verandering in de R-kwadraat voor stap 4 ten opzichte van stap 3, analoog aan de tabellen 8.4.A en 8.5.A, is weergegeven in tabel 8.9:

Tabel 8.9: Variantiecomponenten per niveau voor stap 4 ten opzichte van stap 3

Variabele	Individu	Afdeling	Sector
spanning	1.7%	0.7%	0.2%
welbevinden	3.1%	-1.2%	4.5%

Bij "spanning" wordt de belangrijkste bijdrage van de kwadratische termen gevonden op het individuele verklaringsniveau. Bij "welbevinden" leveren de kwadratische termen een bijdrage zowel op individueel niveau als op sectorniveau. De negatieve R-kwadraat bij "welbevinden" op afdelingsniveau is zo klein, dat deze vermoedelijk als een toevalseffect kan worden afgedaan. De gevonden percentages zijn niet erg hoog in vergelijking met andere waarden, zoals die zijn opgenomen in tabel 8.4.A en 8.5.A.

Vervolgens zijn de t-waarden berekend, en is de significantie van deze t-waarden beoordeeld (het aantal vrijheidsgraden bedraagt bij alle onafhankelijke variabelen 399). Aangezien de belangstelling uitgaat naar stap 4 (lineaire en non-lineaire effecten PSA), maar de bijdrage van deze stap vooral ook geïnterpreteerd moet worden tegen de achtergrond van de resultaten uit stap 3 (alleen lineaire effecten PSA), worden in tabel 8.10 de resultaten van stap 3 en stap 4 naast elkaar gerapporteerd.

Tabel 8.10: T-waarden stap 4 versus stap 3

X	Spanning				Welbevinden			
	Stap 3		Stap 4		Stap 3		Stap 4	
	T	Sign.	T	Sign.	T	Sign.	T	Sign.
geslacht	-0.41	ns	-0.66	ns	-1.39	ns	-1.37	ns
dienstjaren	-2.11	**	-1.97	**	-2.74	***	-2.28	**
leeftijd	2.12	**	1.52	ns	-0.61	ns	-1.28	ns
uren werk/week	2.01	**	1.92	*	0.68	ns	0.71	ns
functieniveau	2.58	***	2.40	**	1.76	*	1.49	ns
NA	9.23	***	9.20	***	4.31	***	4.22	***
EO	0.42	ns	0.72	ns	0.83	ns	1.35	ns
psy. eisen	6.66	***	-1.81	*	0.02	ns	-2.72	***
veelzijdigheid	-0.85	ns	1.13	ns	6.68	***	-2.02	**
regelmogelijkheden	-0.26	ns	-1.80	*	0.38	ns	0.95	ns
sociale steun	2.31	**	-0.05	ns	5.18	***	0.09	ns
psy. eisen kwadr.			2.83	***			2.74	***
veelzijdigheid kwadr.			-1.34	ns			3.36	***
regelmog. kwadr.			1.75	*			-0.94	ns
sociale steun kwadr.			0.47	ns			0.94	ns

Bij de biografische kenmerken valt op dat de resultaten ongeveer hetzelfde zijn als in paragraaf 8.5 bij het maximale model. De enige afwijking wordt gevonden bij de variabele functieniveau, waar nu hogere t-waarden worden gevonden. Kennelijk wordt deze t-waarde in het maximale model te niet gedaan door het invoeren van groepskenmerken en geaggregeerde vragenlijstgegevens in de vergelijking.

De verschillen tussen stap 3 en stap 4 qua t-waarden voor de biografische kenmerken zijn betrekkelijk klein. Bij stap 4 gaan de t-waarden over het algemeen licht achteruit door het invoeren van de kwadratische termen voor de PSA-dimensies in de vergelijking.

Bij de confounders is het resultaat vergelijkbaar met hetgeen in paragraaf 8.5 werd gevonden voor het maximale model. Het verschil in t-waarden tussen stap 3 en stap 4 is klein.

Dit brengt ons bij de PSA-dimensies. Wanneer wordt gekeken naar de hiervoor gerapporteerde veranderingen in R-kwadraat en deviance, dan lijkt het effect van het invoeren van de kwadratische termen in de vergelijking mee te vallen. Uit een analyse van de t-waarden blijkt dat dit allesbehalve waar is.

Bij "spanning" is bij stap 3 het resultaat vergelijkbaar met hetgeen werd gevonden in paragraaf 8.5 voor het maximale model. Dit wil zeggen dat er bij "spanning" een sterk lineair effect is van "psychologische taakeisen" en een minder sterk lineair effect van "sociale ondersteuning". Bij "welbevinden" is het resultaat van stap 3 eveneens in overeenstemming met hetgeen in paragraaf 8.5 werd gevonden voor het maximale model: er zijn sterke lineaire effecten van zowel "veelzijdigheid" als "sociale ondersteuning" op "welbevinden".

De resultaten voor de lineaire en kwadratische termen van de PSA-factoren bij stap 4 moeten in relatie tot elkaar worden geïnterpreteerd. Bij "spanning" worden voor "psychologische taakeisen" en "regelmogelijkheden" zowel significante lineaire als kwadratische effecten gevonden. Er is hier dus inderdaad sprake van non-lineaire effecten. Bij deze dimensies is het van belang nader te onderzoeken of er sprake is van monotone of niet-monotone effecten. De effecten van "psychologische taakeisen" bij stap 4 zijn vrij sterk. Voor "regelmogelijkheden" zijn de effecten zwakker. Toch is dit laatste resultaat opmerkelijk: bij stap 3 was immers geen significant effect voor "regelmogelijkheden" aanwezig. Kennelijk heffen lineaire en kwadratische effecten van "regelmogelijkheden" elkaar op in het geval dat alleen een lineair effect wordt onderzocht.

De dimensie "veelzijdigheid" laat noch bij stap 3, noch bij stap 4 enig effect zien, en is derhalve niet gerelateerd aan "spanning".

Bij de dimensie "sociale ondersteuning" wordt bij stap 4 noch een lineair, noch een kwadratisch effect gevonden. Wanneer het lineair effect van "sociale ondersteuning" op "spanning" wordt onderzocht terwijl wordt gecontroleerd voor het kwadratische effect, dan blijkt geen significant lineair effect te worden gevonden. Aangezien er bij stap 3, zonder controle voor het kwadratisch effect, wel een duidelijk lineair effect was, lijkt "sociale ondersteuning" een puur lineaire relatie te hebben met "spanning". Dit effect is redelijk sterk. De schatting van de regressie-parameter uit de multi-level analyse bedraagt .106. Aangezien deze waarde positief is, betekent dit dat bij het toenemen van problemen met betrekking tot "sociale ondersteuning" de variabele "spanning" toeneemt.

Bij "welbevinden" levert stap 4 een opmerkelijk resultaat voor de dimensie "psychologische taakeisen" op. Bij stap 3 werd een t-waarde gevonden van nagenoeg 0 voor deze dimensie. Dit blijkt het gevolg te zijn van het niet onderscheiden van lineaire en kwadratische componenten: deze heffen elkaar op.

Bij stap 4 zijn namelijk zowel sterke lineaire als kwadratische effecten te zien. Ook hier geldt dat het nog de vraag is, of er sprake is van een monotoon dan wel niet-monotoon non-lineair verband.

Voor de dimensie "veelzijdigheid" geldt een sterk lineair effect op "welbevinden" bij stap 3. Bij stap 4 blijken zowel de lineaire term als de kwadratische term sterk significant te zijn. Ook deze relatie zal worden onderzocht op al dan niet monotoon zijn.

De dimensie "regelmogelijkheden" heeft noch in stap 3, noch in stap 4 een relatie met "welbevinden". Hier bestaat dus geen verband.

Tenslotte de relatie tussen "welbevinden" en "sociale ondersteuning". Hier wordt het resultaat gerepliceerd dat werd gevonden bij "spanning". Een sterk lineair effect bij stap 3 verdwijnt bij stap 4, en er wordt ook geen significante kwadrateterm gevonden. Het verband tussen "sociale ondersteuning" en "welbevinden" lijkt derhalve gewoon lineair te zijn, en is sterk. Ook bij "welbevinden" is de schatting van de regressie-parameter uit de multi-level analyse positief (.286), wat betekent dat naarmate de problemen op het gebied van "sociale ondersteuning" toenemen, het "onwelbevinden" toeneemt.

8.7.2 Nadere analyse non-lineaire effecten

Voor vier van de acht relaties tussen PSA-dimensies (x- oftewel onafhankelijke variabelen) en werkstress-dimensies (y- oftewel afhankelijke variabelen) zijn in de multi-level analyse significante combinaties gevonden van lineaire en kwadratische termen. Het gaat om de relaties van "psychologische taakeisen" en "regelmogelijkheden" met "spanning", en om de relaties van "veelzijdigheid" en "psychologische taakeisen" met "welbevinden". Om uit te zoeken wat de aard van de gevonden non-lineaire relaties is, is de volgende procedure gevolgd.

Voor elk van deze vier relaties zijn polynomiaal-vergelijkingen opgesteld van het type $y = ax^2 + bx$, waarbij a de schatting is van de regressie-parameter uit de multi-level analyse voor de kwadratische term, en b voor de lineaire term.

Om te bepalen of een polynomiaal-vergelijking monotoon of niet-monotoon is, wordt in eerste instantie berekend bij welke waarde de vergelijking een minimum of maximum heeft, met behulp van de afgeleide. Een vergelijking van het type zoals hierboven beschreven heeft een minimum of maximum bij de x-waarde $-b/2a$. Deze x-waarde wordt aangeduid met de term $x_{\min/\max}$. Als $x_{\min/\max}$ buiten de range van waarden ligt die x aan kan nemen, is er sprake van een monotoon non-lineair verband. Als $x_{\min/\max}$ binnen de range van waarden ligt die x aan kan nemen, dan is er sprake van een niet-monotoon non-lineair verband.

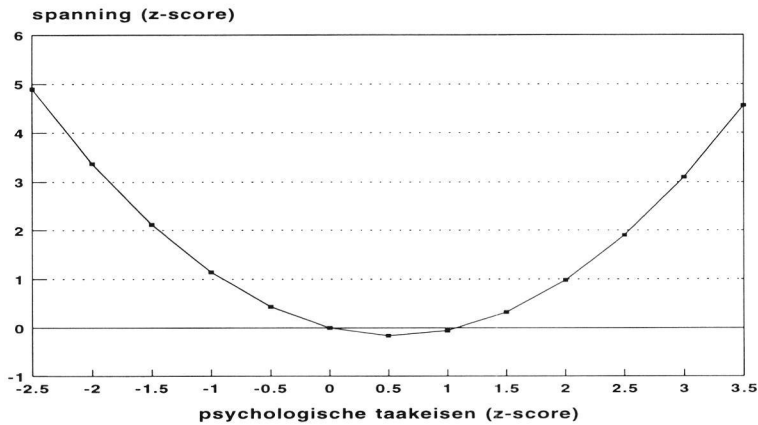
Als, in dit laatste geval, a negatief is betreft het een maximum, als a positief is betreft het een minimum.

Voor elk van de vier onderzochte non-lineaire relaties zijn in tabel 8.11 de waarden opgenomen van a, b, $x_{min/max}$, en de range van x-waarden.

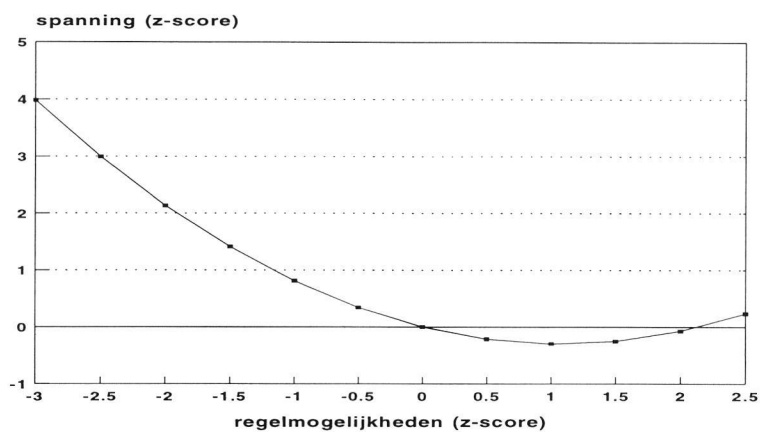
Tabel 8.11: Resultaten nadere analyse non-lineaire effecten

x	y	a	b	$x_{min/max}$	range x	
psychologische taakeisen	spanning	.54	-.60	.55	-2.78 t/m	3.71
regelmogelijkheden	spanning	.26	-.55	1.07	-3.01 t/m	2.89
psychologische taakeisen	welbevinden	.54	-.93	.86	-2.78 t/m	3.71
veelzijdigheid	welbevinden	.33	-.55	.84	-2.24 t/m	2.67

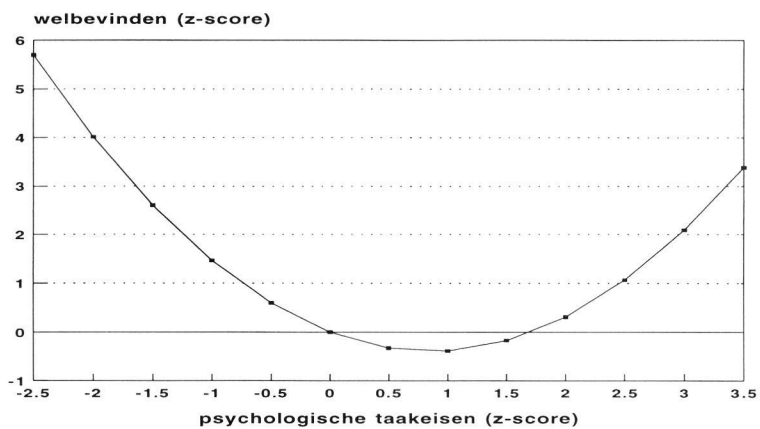
Zoals uit de tabel kan worden afgelezen valt $x_{min/max}$ in alle gevallen binnen de antwoordrange van x, en is a in alle gevallen positief. Dit betekent dat alle vier de polynomiaal-vergelijkingen niet-monotoon zijn, en dat ze alle vier een minimum hebben. In de figuren 8.2 t/m 8.5 staan de polynomiaal-vergelijkingen afgebeeld, over de volledige antwoordrange van x. Bij deze grafieken dient te worden opgemerkt dat het nulpunt op de y-as willekeurig is, omdat met de constante term in de vergelijking $y = ax^2 + bx$ geen rekening is gehouden. De afstanden op de y-as hebben wel betekenis.



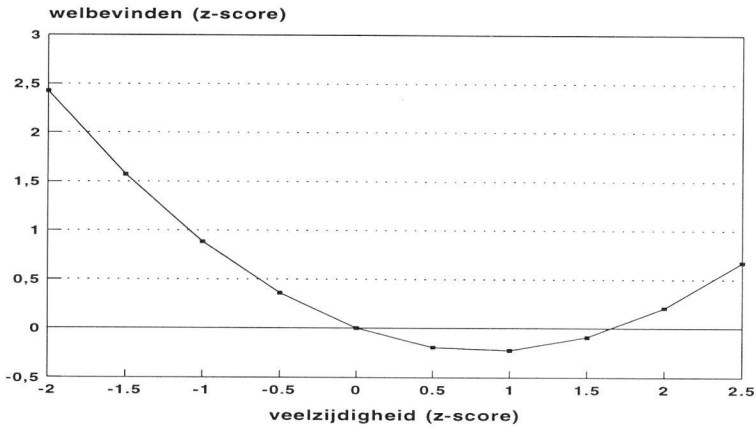
Figuur 8.2: Non-lineaire relatie tussen psychologische taakeisen en spanning



Figuur 8.3: Non-lineaire relatie tussen regelmogelijkheden en spanning



Figuur 8.4: Non-lineaire relatie tussen psychologische taakeisen en welbevinden



Figuur 8.5: Non-lineaire relatie tussen veelzijdigheid en welbevinden

De grafieken die de effecten van "psychologische taakeisen" op "spanning" en "welbevinden" weergeven, laten symmetrische figuren zien. Deze relaties zijn duidelijk kromlijng. Zowel veel als weinig "psychologische taakeisen" gaat gepaard met veel "spanning" en veel "onwelbevinden". De andere twee grafieken zijn scheef naar de negatieve kant. Alhoewel deze relaties eveneens niet-monotoon non-lineair zijn, zijn het vooral lage waarden op de dimensies "regelmogelijkheden" en "veelzijdigheid" die gepaard gaan met hoge waarden qua "spanning" respectievelijk "onwelbevinden". Het is belangrijk hierbij eraan te herinneren dat lage waarden op de schalen "regelmogelijkheden" en "veelzijdigheid" juist betekenen dat er veel "regelmogelijkheden" zijn en dat er veel "veelzijdigheid" is.

8.7.3 Het vitamine-model van Warr

Hoe verhouden de gevonden resultaten zich tot het vitamine-model van Warr (1987, 1994)? We bespreken de vier PSA-dimensies één voor één.

"Psychologische taakeisen" hebben een niet-monotone non-lineaire relatie met "spanning". Er is sprake van een optimum. Zowel bij lage als bij hoge "psychologische taakeisen" treedt veel "spanning" op. Dit is conform de predicties van Warr. Er wordt echter een vergelijkbaar effect gevonden van "psychologische taakeisen" op "welbevinden", wat niet door de Warr wordt voorspeld. Zowel de relatie van "psychologische taakeisen" met "spanning" als met "welbevinden" kunnen worden beschreven als sterk.

"Veelzijdigheid" heeft geen relatie met "spanning". Een interpretatie van deze dimensie in de richting van wat Warr noemt "een groot beroep op iemands vaardigheden", wordt hiermee onwaarschijnlijk. Een "overdosis" hiervan zou volgens hem immers gepaard moeten gaan met "spanning". Warr veronderstelt dat een tekort aan "veelzijdigheid", net als bij alle andere aspecten, gepaard zou moeten gaan met algemene, negatieve stress-reacties. Voor "spanning" wordt dit niet bevestigd.

Des te sterker is de relatie van "veelzijdigheid" met "welbevinden". Deze relatie is niet-monotoon non-lineair. Opnieuw geldt dat er sprake is van een optimum. De grafiek geeft aan dat vooral te veel "veelzijdigheid" gepaard gaat met "onwelbevinden". Dit strookt met de predicties die Warr doet aangaande "toxische" effecten van een "overdosis" van het aspect afwisseling. De dimensie "veelzijdigheid" moet kennelijk meer worden uitgelegd richting afwisseling dan richting beroep op iemands vaardigheden.

"Regelmogelijkheden" hebben geen relatie met "welbevinden", maar wel een zwakke niet-monotone non-lineaire relatie met "spanning". Zowel lage als hoge waarden van "regelmogelijkheden" hangen samen met veel "spanning".

De grafiek geeft aan dat vooral het hebben van veel "regelmogelijkheden" samenhangt met veel "spanning". Dit klopt helemaal met het vitamine-model.

Weinig "regelmogelijkheden" zou gepaard moeten gaan met een algemeen, negatief werkstress-effect. Voor "welbevinden" kon deze veronderstelling van Warr niet worden bevestigd.

Tenslotte de dimensie "sociale ondersteuning". Deze heeft een lineaire relatie met zowel "spanning" als "welbevinden". Hoe meer problemen met "sociale ondersteuning", des te meer problemen met werkstress, zowel qua "spanning" als qua "onwelbevinden". Dit resultaat strookt goed met de algemene, negatieve effecten die Warr veronderstelt van een gebrek aan interpersoonlijk contact en een gebrek aan sociale waardering.

Er werd geen afvlakking gevonden van het effect van "sociale ondersteuning" bij extreem veel sociale ondersteuning, zoals door Warr wordt verondersteld bij een CE-factor (monotoon non-lineair). Bij dit resultaat moet worden opgemerkt dat, theoretisch gezien, de kans bestaat dat in dit onderzoek een instelling is onderzocht waarin nergens extreem veel sociale ondersteuning voorkomt, zodat de voorspelde afvlakking van het effect ook niet kan worden waargenomen. De verkregen resultaten zijn daarom niet zonder meer te extrapoleren naar andere typen organisatie en/of arbeid.

De non-lineaire effecten die Warr veronderstelt van een "overdosis" interpersoonlijk contact konden in deze analyse niet worden onderzocht. Een echte uitspraak over de interpretatie van "sociale ondersteuning" als CE-factor (sociale waardering) of AD-factor (interpersoonlijk contact) is met het beschikbare materiaal niet mogelijk.

Voorlopig is de enige conclusie die kan worden getrokken dat de dimensie "sociale ondersteuning" zich in dit onderzoek gedraagt noch volgens het patroon van een CE-factor, noch volgens het patroon van een AD-factor. Met de toename van de "sociale ondersteuning" blijft het "welbevinden" toenemen, en de "spanning" afnemen. Een interpretatie in de richting van een CE-factor is in dit geval dus wel een stuk aannemelijker dan een interpretatie in de richting van een AD-factor.

8.7.4 Samenvatting en conclusie analyse non-lineaire effecten

De hier uitgevoerde analyse is geen complete test van het vitamine-model, met name niet op het gebied van "sociale ondersteuning". De resultaten pleiten grotendeels voor de geldigheid van het model. Wanneer de dimensie "veelzijdigheid" wordt beschouwd als een operationalisatie van het aspect afwisseling, zoals Warr dit beschrijft, dan worden alle "toxische reacties" bevestigd, die op grond van het vitamine-model konden worden voorspeld.

Een belangrijke toevoeging is het niet-monotone non-lineaire effect dat "psychologische taakeisen" heeft op "welbevinden", dat door Warr niet wordt voorspeld. Een soortgelijk non-lineair effect werd recentelijk gerapporteerd door De Jonge, Schaufeli en Furda (1995), waarbij "psychologische taakeisen" was geoperationaliseerd als werkdruk, en "welbevinden" als arbeidssatisfactie.

De voorspelde "toxische" reacties uit het model van Warr worden volledig bevestigd in dit proefschrift. Anders staat het met de algemene, negatieve gevolgen van tekorten aan de PSA-dimensies. Alleen voor "psychologische taakeisen" en "sociale ondersteuning" konden dergelijke effecten worden bevestigd. Hier gaan tekorten gepaard met zowel "spanning" als "onwelbevinden". Bij "veelzijdigheid" gaat een tekort alleen gepaard met "onwelbevinden". Bij "regelmogelijkheden" gaat een tekort alleen gepaard met "spanning".

De analyse in paragraaf 8.7 laat de belangrijkste conclusie uit de voorgaande paragrafen van dit hoofdstuk intact: "psychologische taakeisen" blijft een sterk effect houden op "spanning", "veelzijdigheid" houdt een sterk effect op "welbevinden". Bovendien zijn dit ook in de analyse van non-lineaire effecten bij stap 4 de sterkste effecten die zijn gevonden, voor zover het de PSA-dimensies betreft.

Uit paragraaf 8.7 wordt duidelijk hoe belangrijk het is om lineaire en kwadratische termen uit te splitsen waar dit nodig is. Alleen op deze manier worden effecten zichtbaar, zoals die zijn gevonden tussen "psychologische taakeisen" en "welbevinden", en tussen "regelmogelijkheden" en "spanning".

Het is niet zo dat de analyse van lineaire en non-lineaire effecten de theorie van Warr in alle opzichten bevestigt. De belangrijkste conclusie uit dit hoofdstuk is

dan ook niet zozeer de bevestiging van het vitamine-model van Warr, als wel het bestaan van een aantal, parallel opererende werkstress-mechanismen, elk met een eigen aantal determinanten en symptomen, en met een eigen dynamiek.

De analyse van non-lineaire effecten heeft tenslotte nog implicaties voor de discussie over confounding. Door critici van vragenlijstonderzoek is vaak als voorbeeld gegeven van confounding dat schalen over werkdruk en over spanningsverschijnselen overlappen in iteminhoud, wat hoge correlaties in de hand werkt. De grafiek over de relatie tussen "psychologische taakeisen" en "spanning" laat zien dat dit argument niet van toepassing is op deze schalen. Als de critici gelijk hadden, zou een lineair effect zichtbaar moeten zijn. Echter ook lage waarden van "psychologische taakeisen" gaan gepaard met "spanning", geheel volgens de theorie van Warr. Dit pleit voor de validiteit van de geconstrueerde schalen, en tegen confounding als argument om onderzoek naar de relatie tussen PSA en werkstress met vragenlijsten af te raden.

Hoofdstuk 9: De voorspellende waarde van PSA en werkstress voor ziekteverzuim

Weinig onderwerpen zijn op dit moment voor het toepassingsgebied arbeid en gezondheid zo relevant als het onderwerp ziekteverzuim. Alhoewel "spanning" en "welbevinden" in verband met het werk ook als op zichzelf staande gevolgen van PSA kunnen worden beschouwd, die (in het geval van negatieve bevindingen) al rechtvaardigen dat er actie wordt ondernomen, is het toch vaak het vermeende door PSA veroorzaakte ziekteverzuim, dat werkgevers weet te mobiliseren tot de aanpak van problemen op het gebied van PSA.

Dit is opmerkelijk, omdat de onderzoeksliteratuur over het oorzakelijk effect van PSA op ziekteverzuim allerm minst overtuigend in staat is om een dergelijke effect aan te tonen (Grosfeld, 1988; Schalk, 1989; Broersen, Vrijhof en Weel, 1990; Smulders en De Winter, 1993). Er speelt bij het aantonen van deze relatie namelijk eenzelfde fenomeen als we in het vorige hoofdstuk zagen in de relatie tussen PSA en werkstress (Semmer, Zapf en Greif, 1995): het multicausale karakter van ziekteverzuim zorgt dat er een limiet is aan de hoeveelheid verklaarde variantie die op basis van één oorzakelijke factor kan worden verklaard.

9.1 Het voorspellen van ziekteverzuim

PSA en ziekteverzuim

De resultaten die worden gepresenteerd door Smulders en De Winter (1993) kunnen dienen als voorbeeld van wat in het onderzoek naar de relatie tussen PSA en ziekteverzuim is gevonden. In dit onderzoek vond een afname plaats van de Vragenlijst Arbeid en Gezondheid (VAG: Grundemann, Smulders en De Winter, 1993) bij werknemers van twee grotere organisaties (een drukkerij in de particuliere sector: N=830, en een technisch onderhoudsbedrijf in de overheidssector: N=1830). Het verzuimpercentage en de verzuimfrequentie werden voor deze bedrijven gevolgd over een periode van vijf jaar na de vragenlijstafname. De werksituatievragen uit de VAG werden samengevoegd tot zeven schaalscores: taakinhoud, leiding, werkdruk, desorganisatie, werkomstandigheden, luchtomstandigheden, en waardering/beloning. De Cronbach's Alpha van deze zeven schalen ligt tussen .55 en .80, met een gemiddelde van .65. Vervolgens werd een multiële regressie-analyse uitgevoerd, waarbij deze zeven scores, samen met de leeftijd en het geslacht, werden gebruikt om de verzuimparameters te voorspellen. In tabel 9.1 volgen de R-kwadraten voor de totale vergelijkingen, voor het verzuimpercentage en de verzuimfrequentie afzonderlijk, en over vier opeenvolgende periodes na het vragenlijstonderzoek.

Tabel 9.1: Relatie van een éénmalige vragenlijstmeting van PSA met ziekteverzuim over vijf jaar na de vragenlijstmeting (ontleend aan Smulders en De Winter, 1993)

	verzuimpercentage	verzuimfrequentie
periode:		
hetzelfde jaar	.07	.14
1 jaar later	.05	.10
2 jaar later	.05	.09
3 jaar later	.06	.08

In nagenoeg alle door Smulders en De Winter (1993) gerapporteerde vergelijkingen, is de leeftijd de voornaamste voorspeller. Wanneer we leeftijd en geslacht uit de vergelijkingen zouden halen komen we tot een verklaarde variantie van rond de 5%, op basis van de zeven schaalcores gezamenlijk. Verder is te zien dat de kwaliteit van de voorspelling afneemt met de tijd na het vragenlijstonderzoek, zoals te verwachten was. Dit effect treedt vooral op bij de verzuimfrequentie.

Bij dit resultaat dienen we te bedenken, dat de onderwerpen in de werkbeleving die door Smulders en De Winter zijn onderzocht, breder zijn dan wat wij in het derde deel van dit proefschrift bekijken. PSA is in het derde deel van dit proefschrift beperkt tot arbeidsinhoud en -organisatie. In de analyse van Smulders en De Winter zitten ook scores op het gebied van fysisch-chemische arbeidsomstandigheden en arbeidsvoorwaarden. De scores van Smulders en De Winter strekken zich uit over de volle breedte van aspecten van de kwaliteit van de arbeid.

De resultaten van Smulders en De Winter zijn exemplarisch. Ook andere onderzoeken (Grosfeld, 1988; Schalk, 1989; Broersen, Vrijhof en Weel, 1990) komen tot vergelijkbare resultaten. De empirische basis voor de stelling dat PSA een oorzaak is van ziekteverzuim, is smal, ondanks de groeiende praktijk en dienstverlening op dit gebied (psychologie van arbeid en gezondheid). Eerder is besproken dat bij het verklaren van werkstress-variabelen door PSA een theoretische limiet geldt van 10 á 15% van de variantie (Semmer et al., 1995). Voor ziekteverzuim lijkt deze limiet nog iets lager te liggen op grond van de hier genoemde onderzoeken: 5 á 10% van de variantie lijkt realistisch.

Sociale vergelijking en ziekteverzuim

Het is - gezien het voorgaande - opmerkelijk dat er een onderzoekslijn is, die stelt dat ziekteverzuim wel degelijk voor een belangrijk deel te verklaren is op basis van psychische processen, en wel op basis van sociale vergelijkingprocessen (Chadwick-Jones et al., 1982; Nicholson en Johns, 1985; Geurts, 1994, 1995; Hopstaken, 1994). In paragraaf 3.2 is reeds ingegaan op de sociale vergelijkingstheorie.

Vooraf de billijkheidstheorie van Adams (1965) vindt op het gebied van ziekteverzuim toepassing. Een waargenomen scheefheid in de ruilrelatie tussen werknemer en organisatie qua kosten/baten wordt verondersteld een belangrijke oorzaak van verzuim te zijn. Verzuim wordt in deze optiek beschouwd als compensatiegedrag: een werknemer probeert op deze manier de scheve balans tussen hemzelf en de organisatie weer recht te trekken. Behalve ervaren onbillijkheid (EO) zijn ook de normen die werknemers hebben ten aanzien van verzuim (Hill en Trist, 1962) een belangrijk onderwerp van studie in deze onderzoekslijn. Bij het tot stand komen van deze normen worden sociale vergelijkingprocessen verondersteld eveneens een belangrijke rol te spelen. Een derde sociaal-psychologische theorie die op het terrein van verzuim toepassing begint te vinden, is de theorie van gepland gedrag van Ajzen (1991). Hierin spelen normen en opvattingen over de verwachte gevolgen van een bepaald gedrag een hoofdrol.

Geurts (1994) ontwikkelde op basis van de sociale vergelijkingstheorie een LISREL-model voor de verklaring van de verzuimfrequentie bij twee metaalbedrijven. De verzuimfrequentie werd hierbij prospectief gemeten, over het jaar na de vragenlijstafname. Met dit model is Geurts uiteindelijk in staat om 23% van de variantie te verklaren bij het ene bedrijf en 17% bij het andere. Bij buschauffeurs vindt zij op vergelijkbare wijze zelfs 30% verklaarde variantie. Het is niet mogelijk om op grond van haar presentatie van gegevens vast te stellen hoe groot het aandeel in deze verklaarde variantie is van EO, van de verzuimnorm, en van andere variabelen in het model.

Een complicatie is, dat in het model van Geurts (1994) ook ervaren gezondheidsklachten zijn opgenomen. Uit ander verzuimonderzoek weten we dat ervaren gezondheid een goede voorspeller van verzuim is (Grosfeld, 1989; Schalk, 1989; De Winter, 1991). Het eerder besproken onderzoek van Smulders en De Winter (1993) kan hier opnieuw als voorbeeld dienen. De R-kwadraat voor de voorspelling van het verzuimpercentage over het jaar van het vragenlijstonderzoek neemt toe van 7% tot 36%, wanneer we gezondheidsbelevingsvragen toevoegen aan de vergelijking. Bij de verzuimfrequentie neemt ditzelfde cijfer toe van 14% tot 44%.

Hiermee rekening houdend, en afgaand op de regressiecoëfficiënten die Geurts (1994) rapporteert, draagt EO zo'n 5 tot 10% bij aan de variantie in de verzuimfrequentie, wat aardig overeenkomt met de theoretische limiet van Semmer, Zapf en Greif (1995).

Vraagstellingen

Onderzoek naar de relatie tussen PSA, werkstress en verzuim is relevant voor de arbo-praktijk. De bestaande evidentie voor een dergelijke relatie is niet sterk. Het is aannemelijk op grond van de literatuur dat sociale vergelijkingprocessen een rol spelen bij het ontstaan van verzuim.

In het voorgaande is een structuur ontwikkeld voor het meten van PSA en werkstress met vragenlijsten (hoofdstuk 7), en is de relatie tussen PSA en werkstress uitgewerkt (hoofdstuk 8). Hierbij bleek dat de relatie tussen PSA en "spanning" in sterke mate vertekend kan worden door negatieve affectiviteit (NA). Ook bleek dat PSA wel degelijk een eigen bijdrage levert aan de verklaarde variantie in "spanning" en "welbevinden". In dit hoofdstuk proberen we onze tot hier aan toe opgebouwde kennis over PSA en werkstress te koppelen aan het onderwerp ziekteverzuim.

De vraagstelling voor dit hoofdstuk is de volgende: kunnen we ziekteverzuim voorspellen op basis van PSA en werkstress, rekening houdend met het verstorende effect van NA, en boven op het te verwachten effect van EO op ziekteverzuim?

9.2 Voorbereiding van de analyse

Voor alle personen in het gegevensbestand, zoals gebruikt in hoofdstuk 8, zijn individuele ziekteverzuimgegevens verzameld over het eerste half jaar na het vragenlijstonderzoek. Het gaat dus om prospectieve verzuimgegevens. De periode van een half jaar is gekozen omdat het aannemelijk is dat de voorspellende waarde van een éénmalige vragenlijstmeting voor ziekteverzuim afneemt naarmate we het tijdvak na de vragenlijstmeting groter maken (De Winter, 1991; Smulders en De Winter, 1993). Er zijn twee verzuimparameters verzameld: de verzuimfrequentie (het aantal nieuwe ziektegevallen in de periode) en de verzuimduur (het totaal aantal verzuimdagen in de periode). De ziekteverzuimgegevens zijn bewerkt met een logaritmische transformatie alvorens ze te analyseren, om te corrigeren voor de scheve verdeling.

Omdat bij de logaritmische transformatie een probleem ontstaat bij nulverzuim ($\log(0) = \text{min oneindig}$), is bij de oorspronkelijke verzuimwaarden eerst de waarde 1 opgeteld, zodat nulverzuim ook na de transformatie een waarde heeft gelijk aan 0 ($\log(1) = 0$). Ondanks deze transformatie blijft ziekteverzuim een parameter die niet-normaal is verdeeld, zodat de resultaten toch met enige voorzichtigheid dienen te worden geïnterpreteerd.

Ook in dit hoofdstuk wordt gewerkt met de begrippen NA en EO. Deze zijn op dezelfde wijze geoperationaliseerd als in hoofdstuk 8.

Verder is opnieuw gewerkt met de vier dimensiescores ter operationalisatie van PSA; "psychologische taakeisen", "veelzijdigheid", "regelmogelijkheden" en "sociale ondersteuning". Ook gebruiken we opnieuw de twee dimensiescores ter operationalisatie van werkstress: "spanning" en "welbevinden".

Het lag in de bedoeling om een vergelijkbare multi-level analyse, zoals beschreven in het vorige hoofdstuk, uit te voeren met individuele verzuimgegevens als afhankelijke variabele. Er werd daarom gestart met eenzelfde multi-level analyse

als is beschreven in paragraaf 8.4, met dezelfde onafhankelijke variabelen, dezelfde analysestappen, en dezelfde nestingstructuur. Hierbij trad een complicatie op. De bruto verdeling van de variantie over de drie verklaringsniveaus (zie paragraaf 8.5.1) was als volgt (zie tabel 9.2).

Tabel 9.2: Verdeling van de variantie in het lege model

Variabele	Individu	Afdeling	Sector
verzuimfrequentie	96.0%	4.0%	0.0%
verzuimduur	99.2%	0.8%	0.0%

De hoeveelheid variantie op boven-individueel niveau blijkt gering te zijn. De verzuimfrequentie kent nog een beperkte hoeveelheid "gedeelde variantie" (4%), bij de verzuimduur is dit echter nagenoeg nihil (minder dan 1%). Gezien hetgeen hiervoor is opgemerkt over de theoretische limiet, is de gevonden "gedeelde" variantie niet onmiddellijk te bestempelen als klein, met name voor de verzuimfrequentie. Als voor de analyse grote bestanden beschikbaar zouden zijn, met daarin veel verschillende soorten afdelingen en soorten organisaties, dan zou een multi-level analyse zinvol zijn. De analyse zou zich in dat geval kunnen beperken tot twee niveaus, afdeling en organisatie, door de individuele verzuimcijfers en vragenlijstgegevens eerst te aggregeren tot afdelingsniveau.

De multi-level analyse komt in het onderhavige geval echter neer op het verklaren van de variantie in individuele verzuimcijfers, verzameld bij één enkele organisatie met een beperkt aantal afdelingen en sectoren. Dit was aanleiding om ervoor te kiezen de vraagstelling naar de relatie tussen PSA, werkstress en verzuim uit te werken met LISREL, en niet met multi-level-analyse: die heeft in dit geval relatief weinig meerwaarde.

Op hetzelfde databestand als gebruikt in hoofdstuk 8 is derhalve een aantal LISREL-modellen onderzocht ter voorspelling van het ziekteverzuim. Alle gegevens in deze LISREL-analyse zijn gegevens van individuele werknemers. De steekproefomvang bedraagt 411.

We bouwen in dit hoofdstuk verder op het eerder ontwikkelde model uit paragraaf 7.8. Het is belangrijk om er opnieuw uitdrukkelijk op te wijzen, dat in deze LISREL-analyse uitsluitend wordt gewerkt met geobserveerde variabelen, in tegenstelling tot de confirmatorische factoranalyse in hoofdstuk 7, waar met latente en geobserveerde variabelen werd gewerkt.

Bij de conversie van de gegevens naar LISREL zijn alle variabelen gestandaardiseerd, en is gebruik gemaakt van een zogenaamde PM-matrix (Jöreskog en Sörbom, 1993). Dit wil zeggen dat voor elk paar van variabelen in de matrix een associatie-maat is berekend, die past bij het meetniveau van de desbetreffende twee variabelen. Drie variabelen hebben een ordinaal meetniveau, te weten: de verzuimfrequentie, NA en EO. De overige variabelen zijn in de analyse meegenomen op intervalniveau.

In hoofdstuk 8 bleek EO geen belangrijke confounder te zijn. Om EO in de hiernavolgende analyse toch nog te gebruiken als confounder is niet zinvol. Vanuit de literatuur is wel aannemelijk te maken, dat er een rechtstreeks effect is van EO op verzuim (Geurts, 1994). Dit wordt in het onderstaande onderzocht.

Watson en Pennebaker (1989) claimen dat NA wel gerelateerd is aan subjectieve parameters van gezondheid en welbevinden, maar niet aan objectieve gegevens zoals doktersbezoek, medische consumptie en verzuim. Dit past bij de rol van NA als confounder in werkstress-onderzoek. Watson en Pennebaker presenteren een grote hoeveelheid onderzoek, waarin hun stellingname wordt bevestigd. Alvorens NA te gebruiken als confounder zullen we echter ook het directe effect van NA op verzuim onderzoeken.

Eerst worden nu twee LISREL-modellen onderzocht waarbij de plaatsing van NA en EO verschilt. De specificaties van deze twee LISREL-modellen zijn opgenomen in bijlage 2.1 en 2.2.

In model 1 beschouwen we NA en EO als zelfstandige voorspellers van verzuim. We postuleren verder dat verzuim voorspeld kan worden uit "spanning" en "welbevinden", die op hun beurt weer verklaard kunnen worden door PSA. Dit model sluit aan bij de gedachtengang achter de herziene belasting-belastbaarheid modellen en achter het Michigan-model, over een oorzakelijkheidsketen van belasting via belastingsverschijnselen naar belastingsgevolgen. De vergelijkingen in model 1 zien er als volgt uit:

model 1 PSA = spanning
 PSA = welbevinden
 spanning + welbevinden + NA + EO = verzuim

In model 2 nemen we de resultaten uit hoofdstuk 8 ter harte, en beschouwen we NA als voorspeller van "spanning" en "welbevinden". Deze plaatsing is meer conform de visie op NA als "nuisance factor" in werkstress-onderzoek. Uit hoofdstuk 8 weten we ook dat een dergelijke invloed van EO niet te verwachten valt. Voor EO handhaven we het rechtstreekse effect op verzuim in het model.

De vergelijkingen behorend bij model 2 zijn de volgende.

model 2 PSA + NA = spanning
 PSA + NA = welbevinden
 spanning + welbevinden + EO = verzuim

9.3 Resultaten model 1 en 2

In de eerste plaats kijken we naar de passingsmaten van deze twee modellen voor de verzuimfrequentie en de verzuimduur afzonderlijk (zie tabel 9.3). Model 1 voldoet alleen qua RMSR aan de criteria voor een goed passend model (RMSR <.05), terwijl model 2 qua AGFI en RMSR hieraan voldoet (AGFI >.90 en RMSR <.05). Model 2 is duidelijk beter dan model 1. De plaatsing van NA als "nuisance factor" in het model komt de passing ten goede.

Tabel 9.3: Passingsmaten model 1 en model 2

Variabele	Model	Chi	Df	p	AGFI	RMSR
verzuimfrequentie	1	61.91	9	.000	.84	.049
	2	22.61	8	.004	.94	.033
verzuimduur	1	60.27	9	.000	.84	.046
	2	18.79	8	.016	.95	.026

In de tweede plaats kijken we naar de t-waarden van de onafhankelijke variabelen in de eerste twee vergelijkingen (zie tabel 9.4) . Bij significante t-waarden staat tussen haakjes de mate van significantie vermeld. De kritieke grenzen hierbij zijn .10 (*), .05 (**) en .01 (***). Onder aan de kolom met t-waarden is de verklaarde variantie (R-kwadraat) voor de hele vergelijking opgenomen. De resultaten voor de verzuimfrequentie en de verzuimduur zijn gelijk, dus we presenteren ze hier slechts één keer.

De bovenstaande resultaten zijn vergelijkbaar met die uit paragraaf 7.8 en die van de multi-level analyse uit paragraaf 8.5.3. "Spanning" hangt samen met "psychologische taakeisen" en "sociale ondersteuning". "Welbevinden" hangt samen met "veelzijdigheid" en "sociale ondersteuning". De score voor "regelmogelijkheden" heeft noch met "spanning", noch met "welbevinden" enig verband. De differentiatie in werkstress-mechanismen wordt opnieuw bevestigd.

Tabel 9.4: T-waarden model 1 en model 2 (vergelijking 1 en 2)

Variabele	Spanning		Welbevinden	
	model 1	model 2	model 1	model 2
psychologische eisen	9.22 (***)	7.21 (***)	1.47	0.58
veelzijdigheid	0.34	-0.88	5.75 (***)	5.13 (***)
regelmogelijkheden	-0.53	-0.53	1.00	1.02
sociale steun	4.24 (***)	3.92 (***)	6.88 (***)	6.70 (***)
NA	----	6.01 (***)	----	2.62 (***)
R-kwadraat	.28	.33	.26	.27

Het effect van NA als confounder lijkt wat milder dan bij de multi-level analyse. Alleen bij "spanning" is er bij model 2 een duidelijke toename van de R-kwadraat ten opzichte van model 1. We zien dat de introductie van NA in de vergelijking gepaard gaat met een reductie van de t-waarde voor "psychologische taakeisen" als voorspeller van "spanning". Dit is geheel volgens verwachting.

Vervolgens komt de derde vergelijking in model 1 en model 2 aan bod, die betrekking heeft op de voorspelling van verzuim. De gepresenteerde gegevens in tabel 9.5 zijn qua opzet vergelijkbaar met die in tabel 9.4.

Tabel 9.5: T-waarden model 1 en model 2 (vergelijking 3)

Variabele	Verzuimfrequentie		Verzuimduur	
	model 1	model 2	model 1	model 2
NA	1.80 (*)	----	0.84	----
EO	0.19	0.18	3.38 (***)	3.38 (***)
Spanning	1.35	2.01 (**)	2.70 (***)	3.04 (***)
Welbevinden	2.24 (**)	2.51 (**)	0.54	0.67
R-kwadraat	.03	.03	.06	.06

Als eerste valt op dat de verklaarde variantie (R-kwadraat) laag is. De verzuimduur (6%) kan beter worden voorspeld dan de verzuimfrequentie (3%). Het al dan niet in de vergelijking opnemen van NA doet er kennelijk niet zoveel toe, want er is geen verschil in verklaarde variantie tussen model 1 en 2.

Desalniettemin blijken alle vergelijkingen significante voorspellers te bevatten: Bij de verzuimfrequentie is "welbevinden" de beste voorspeller, zowel in model 1 als in model 2. In model 1 is NA in lichte mate een voorspeller van de verzuimfrequentie. EO doet ten aanzien van de verzuimfrequentie niets: noch bij model 1, noch bij model 2. Als we NA uit de vergelijking halen (model 2) blijkt "spanning" plotseling wel een significante voorspeller te zijn, waar dat bij model 1 nog niet zo was.

Bij de verzuimduur wordt bevestiging gevonden voor de voorspellende waarde van EO. Met of zonder NA in de vergelijking, is dit de beste voorspeller. "Spanning" blijkt een tweede significante voorspeller van de verzuimduur te zijn. Dit effect treedt op, ongeacht of NA al dan niet in de vergelijking is opgenomen. NA heeft in model 1 geen direct effect op de verzuimduur. "Welbevinden" heeft noch in model 1, noch in model 2 een effect op de verzuimduur.

Als we de informatie over passingsmaten en t-waarden combineren, dan is model 2 duidelijk beter dan model 1. De hoeveelheid verklaarde variantie in ziekteverzuim valt echter tegen. In model 1 en 2 wordt in beide gevallen er van uitgegaan dat ziekteverzuim alleen indirect samenhangt met PSA, via "spanning" en "welbevinden".

Deze indirecte samenhang is conform de gedachte van een oorzakelijkheidsketen van belasting via belastingsverschijnselen naar belastingsgevolgen (Caplan et al., 1975; Meijman, 1989; Van Dijk et al., 1990; Kompier, 1991). We kunnen onderzoeken of de gepostuleerde oorzakelijkheidsketen in dit onderzoek kan worden aangetoond, door model 2 te vergelijken met een model waarin ook een rechtstreeks effect van PSA op verzuim wordt opgenomen. Aldus komen we op model 3, met de volgende vergelijkingen:

model 3 PSA + NA = spanning
 PSA + NA = welbevinden
 PSA + spanning + welbevinden + EO = verzuim

De specificatie van model 3 is te vinden in bijlage 2.3.

9.4 Resultaten model 3

In deze paragraaf zetten we model 3 telkens af tegen model 2. We bekijken weer eerst de passingsmaten in tabel 9.6.

Tabel 9.6: Passingsmaten model 2 en model 3

Variabele	Model	Chi	Df	p	AGFI	RMSR
verzuimfrequentie	2	22.61	8	.004	.94	.033
	3	11.00	4	.027	.94	.017
verzuimduur	2	18.79	8	.016	.95	.026
	3	9.37	4	.052	.95	.016

Qua modelpassing levert model 3 een verbetering op ten opzichte van model 2. Model 3 voldoet aan de criteria AGFI >.90 en RMSR <.05. Het criterium Chi-kwadraat met een p-waarde >.05 wordt bij de verzuimduur eveneens gehaald. De eerste twee vergelijkingen zijn bij model 2 en 3 gelijk, en hoeven we hier niet te herhalen: zie de vorige paragraaf. We gaan derhalve verder met de derde vergelijking: het voorspellen van verzuim. De t-waarden van model 2 en 3 zijn opgenomen in tabel 9.7.

Tabel 9.7: T-waarden model 2 en model 3 (vergelijking 3)

Variabele	Verzuimfrequentie		Verzuimduur	
	model 2	model 3	model 2	model 3
EO	0.18	-0.08	3.38 (***)	2.97 (***)
Psychologische eisen	----	0.09	----	1.39
Veelzijdigheid	----	2.04 (**)	----	2.03 (**)
Regelmogelijkheden	----	1.14	----	1.56
Sociale steun	----	1.01	----	-0.90
Spanning	2.01 (**)	2.05 (**)	3.04 (***)	2.66 (***)
Welbevinden	2.51 (**)	0.62	0.67	-0.24
R-kwadraat	.03	.06	.06	.08

De totale verklaarde variantie blijkt door deze ingreep te verbeteren. Model 3 is hierin beter dan model 2, zij het dat de R-kwadraten altijd nog onder de .10 blijven.

Bij de verzuimfrequentie is te zien we dat EO geen effect heeft. Verder zien we dat de significantie van "welbevinden", die we vonden bij model 2, verdwijnt na het opnemen van PSA in de vergelijking (model 3). Nu is "veelzijdigheid" een belangrijke voorspeller geworden. "Spanning" daarentegen behoudt zijn effect op de verzuimfrequentie in model 3, ondanks de introductie van PSA in de vergelijking. De PSA-factoren "psychologische taakeisen", "regelmogelijkheden" en "sociale ondersteuning" hebben geen directe effecten op de verzuimfrequentie.

Bij de verzuimduur behoudt EO in model 3 een zelfstandig effect, zoals dat ook werd gevonden bij model 2. Hetzelfde geldt voor het effect van "spanning". "Welbevinden" doet niets: noch in model 2, noch in model 3. Van de PSA-factoren heeft alleen "veelzijdigheid" in model 3 een direct effect op de verzuimduur. Dit is overeenkomstig de situatie bij de verzuimfrequentie.

9.5 Samenvatting en conclusie over het voorspellen van ziekteverzuim

Uit de hiervoor gepresenteerde resultaten is een reeks van conclusies te trekken, te beginnen met de centrale vraag van dit hoofdstuk.

PSA, sociale vergelijking en verzuim

Wat betreft de rol van EO ten aanzien van verzuim hebben we deze kunnen bevestigen voor de verzuimduur. Bij de verzuimduur is additioneel de voorspellende waarde van werkstress (vooral "spanning") en PSA (vooral "veelzijdigheid") aangetoond. Bij het voorspellen van de verzuimfrequentie is alleen het belang van PSA ("veelzijdigheid") en werkstress ("spanning") aangetoond. Er is geen effect van EO op de verzuimfrequentie gevonden. Verzuim komt uit dit onderzoek derhalve niet naar voren als een gegeven dat alleen maar een kwestie van sociale vergelijking is. Ook factoren op het gebied van PSA en werkstress hebben een eigen voorspellende waarde ten aanzien van verzuim.

Het opmerkelijke van deze bevindingen is, dat Geurts (1994) haar effecten van EO aantoonde op prospectieve gegevens van de verzuimfrequentie. Bij de verzuimfrequentie wordt in dit onderzoek juist geen effect gevonden. Gegeven de R-kwadraten van de vergelijkingen in model 3 die verzuim voorspellen, kan op basis van dit onderzoek niet worden bevestigd dat EO zo'n 5% á 10% van de variantie verklaart in ziekteverzuim. Dit resultaat van Geurts (1994) wordt dus eveneens niet bevestigd. De beperkte operationalisatie van EO in het onderhavige onderzoek is hier wellicht debet aan. Een andere mogelijke verklaring is het verschil in onderzoekspopulatie.

Geurts deed haar onderzoek bij metaalbedrijven en buschauffeurs. In het onderzoek dat in dit hoofdstuk is gerapporteerd, is gewerkt met een populatie uit de gezondheidszorg.

Niveau van verklaarde variantie in ziekteverzuim

Ondanks dit alles komt het niveau van verklaarde variantie in ziekteverzuim op grond van PSA, werkstress, NA en EO in dit onderzoek niet boven de 8% uit. Ziekteverzuim blijkt een multicausaal fenomeen te zijn, waarvoor een limiet lijkt te bestaan van 5 á 10% verklaarde variantie op basis van één enkele categorie determinanten. In de vorm van model 3 hebben we echter wel een model gevonden, dat de relatie van een éénmalige vragenlijstmeting van PSA, persoonlijkheid (NA), sociale vergelijking (EO), en werkstress met ziekteverzuim goed beschrijft.

De verklaarde variantie in ziekteverzuim, op basis van acht variabelen met betrekking tot psychosociale factoren, loopt in de analyse op tot maximaal 8%. Dit lijkt laag, maar is in de orde van grootte die we, op grond van onderzoek naar de voorspelbaarheid van ziekteverzuim in Nederland, mogen verwachten (Grosfeld, 1988; Schalk, 1989; Broersen, Vrijhof en Weel, 1990; Smulders en De Winter, 1993).

Om inhoudelijke redenen is de analyse hier beperkt tot de invloed van psychosociale factoren op verzuim, terwijl er veel meer gegevensmateriaal beschikbaar was. In deze populatie had wellicht een groter gedeelte van het verzuim verklaard kunnen worden, indien ook de beschikbare gegevens over factoren als ervaren gezondheid, ergonomie, fysisch-chemische arbeidsomstandigheden en arbeidsvoorwaarden in de analyse waren betrokken. Aan de verklaarbare variantie in individuele verzuimgegevens zit echter een plafond, zo rond de 25% van de variantie (Grosfeld, 1988). Zo bezien is 8% op basis van psychosociale factoren alleen niet eens weinig.

Het is belangrijk vast te stellen dat de voorspelbaarheid van verzuim op grond van "spanning" en "welbevinden" lager uitvalt dan verwacht. Uit de eerder genoemde Nederlandse studies is bekend, dat verzuim redelijk kan worden voorspeld op basis van algemene, actuele gezondheidsklachten. Wanneer we - in plaats hiervan - meer arbeidsgerelateerde maten van "spanning" en "welbevinden" gebruiken, gaat de voorspelbaarheid van verzuim omlaag.

Het lijkt erop dat de arbeidsgerelateerde manier van vragen (in tegenstelling tot het vragen naar de algemene gezondheidsbeleving) gepaard gaat met een verlies aan voorspellend vermogen ten aanzien van verzuim wat een enigszins onverwacht effect is. Het is derhalve zaak om voor de verdere validering van de PSA- en werkstress-schalen te zoeken naar meer werkgebonden uitkomstvariabelen, bijvoorbeeld: productiviteit, verloop, succes van reorganisaties, of snelheid van aanpassing aan veranderingen (flexibiliteit).

De vraag, naar de mate waarin de diverse onafhankelijke variabelen en het onverklaarde deel van de variantie organisatie- dan wel individu-gebonden zijn, blijft in dit onderzoek onbeantwoord. Multi-level analyse op basis van grote bestanden met geaggregeerde data van vele afdelingen binnen vele organisaties is gewenst, omdat dit met name inzicht geeft in het voor organisaties het meest rechtstreeks beïnvloedbare deel van het ziekteverzuim.

Opnieuw de structuur van en relatie tussen PSA en werkstress

Een aantal bevindingen in dit hoofdstuk is relevant in het licht van de vraagstellingen en resultaten van de hoofdstukken 7 en 8, over de structuur van en relatie tussen PSA en werkstress. Met een bespreking hiervan sluiten we dit hoofdstuk af.

De rol van NA als confounder, met name in de relatie tussen PSA en "spanning", wordt opnieuw bevestigd. Ook in deze LISREL-analyse blijkt echter dat PSA een eigen additieve invloed heeft op werkstress. PSA heeft eveneens een eigen directe, additieve invloed op verzuim.

Het model van Warr (1987; 1994) vindt opnieuw ondersteuning. Met uitzondering van zijn claim aangaande het effect van "regelmogelijkheden" op "spanning", worden zijn predicties over de "toxische reacties" tussen de onderzochte PSA- en werkstress-dimensies opnieuw bevestigd (zie paragraaf 8.7). Warr's stelling dat de factoren "psychologische taakeisen", "veelzijdigheid", en "regelmogelijkheden" non-lineair samenhangen met werkstress-uitkomsten, is voor de ziekteverzuimparameters niet onderzocht.

"Sociale steun" en "veelzijdigheid" blijken ook in dit hoofdstuk samen te gaan met "welbevinden" in het werk. "Sociale steun", maar vooral "psychologische taakeisen", blijken samen te gaan met "arbeidsgerelateerde spanningsverschijnselen". Dit pleit opnieuw voor een differentiatie in werkstress-mechanismen, in tegenstelling tot het redeneren vanuit één algemeen geldig mechanisme.

Het is opmerkelijk dat het effect van "spanning" op verzuim prominent is in model 2 en 3. Bij "welbevinden" is de situatie duidelijk anders. Ten aanzien van de verzuimduur is nergens een effect gevonden. Bij de verzuimfrequentie verdwijnt het effect wanneer directe effecten van PSA op verzuim in de vergelijking worden opgenomen. Dit alles bevestigt een interpretatie van "spanning" als een belastingsverschijnsel. "Welbevinden" laat zich volgens deze resultaten minder goed interpreteren als een belastingsverschijnsel. Dit is, naast het hebben van verschillende PSA-voorspellers, naast het verschil in gevoeligheid voor NA, en naast het verschil in de mate van "gedeelde variantie", de vierde aanwijzing voor het bestaan van differentiatie in meer dan één werkstressmechanisme.

Hoofdstuk 10: Discussie

Dit proefschrift gaat over de waarde van een zelfbeoordelingsmethode voor het onderzoek van psychosociale arbeidsbelasting (PSA) en werkstress. In de eerste plaats is een reeks vragenlijstschalen ontwikkeld voor het meten van PSA en werkstress. Hierbij is veel aandacht besteed aan technische criteria waaraan een goede zelfbeoordelingsmethode dient te voldoen. Er is gekeken naar unidimensionaliteit, schaalbaarheid, betrouwbaarheid en validiteit.

Met deze reeks vragenlijstschalen als hulpmiddel zijn we ons daarna bezig gaan houden met een aantal inhoudelijke vragen op het gebied van PSA en werkstress. In het derde deel van dit proefschrift zijn zowel theoretische als methodologische aspecten aan de orde geweest, die van belang zijn voor het onderzoek van PSA en werkstress met vragenlijsten. Dit tiende en laatste hoofdstuk gaat in op de resultaten uit het derde deel van dit proefschrift, en plaatst deze tegen de achtergrond van de literatuur die is besproken in het eerste deel. De inhoudelijke vraagstellingen in het derde deel van dit proefschrift kunnen worden geordend in drie hoofdgroepen, die in het vervolg één voor één worden besproken:

- Wat is de structuur van PSA en werkstress? (paragraaf 10.1)
- Hoe is de relatie tussen PSA en werkstress? (paragraaf 10.2)
- In welke mate kan ziekteverzuim worden voorspeld op basis van PSA en werkstress? (paragraaf 10.3)

De resultaten van hoofdstuk 7, 8 en 9 houden onderling sterk verband met elkaar en vullen elkaar aan. Dit vraagt om een integratie van de resultaten (paragraaf 10.4). Hierna komen de beperkingen van het onderzoek aan bod (paragraaf 10.5), en de betekenis van het onderzoek voor de praktijk (paragraaf 10.6).

10.1 De structuur van PSA en werkstress

Geen algemene onvrede maar een gedifferentieerde structuur

Dit proefschrift levert een spaarzame structuur voor PSA en werkstress. Deze structuur kan worden gebruikt om PSA en werkstress te operationaliseren in vragenlijstonderzoek. Uitgaande van meer dan 25 aspecten, die op basis van een literatuurinventarisatie zijn geïdentificeerd, en die vervolgens zijn geoperationaliseerd als Mokkenschalen, wordt een model ontwikkeld met vier kernaspecten voor PSA en twee voor werkstress. Bij PSA onderscheiden we: "psychologische taakeisen", "veelzijdigheid", "regelmogelijkheden" en "sociaal-organisatorische aspecten". Bij werkstress kunnen we onderscheid maken in "arbeidsgerelateerde spanningsverschijnselen" en "arbeidsgerelateerd welbevinden".

Voor alle kernaspecten geldt, dat deze kunnen worden geoperationaliseerd met veel of weinig afzonderlijke schalen: de structuur blijft overeind. Het "sociaal-organisatorische aspect" blijkt intern redundant te zijn. Bij dit aspect bestaat het risico, dat veel afzonderlijke items en schalen allemaal informeren naar hetzelfde.

Uit dit onderzoek komt naar voren dat het herleiden van de geconstrueerde schalen tot één onderliggende dimensie van "algemene onvrede" (het nulmodel in hoofdstuk 7) slechter past bij de gegevens, dan het herleiden tot het genoemde zestal kernaspecten.

Vaak wordt na een werkbelevingsonderzoek een interpretatie van de resultaten in termen van "algemene onvrede" gehanteerd. Hoe bruikbaar dit ook is voor de signalering van probleemgroepen, het staat het gebruik van zelfbeoordelingsinstrumenten als hulpmiddel voor probleemdiagnose in de weg. Een ongedifferentieerde probleemanalyse heeft als risico, dat er geen aanwijzingen zijn voor een gedifferentieerde probleemoplossing. Het gebruik van een interpretatie in termen van "algemene onvrede" draagt, in tegenstelling tot het argument dat hiervoor vaak wordt gebruikt, niet bij aan de eenvoud van de interpretatie maar vooral aan de vaagheid ervan.

Deze conclusie wordt versterkt door het gegeven, dat de analyse van de structuur van PSA en werkstress (hoofdstuk 7) beperkt is gebleven tot aspecten op het gebied van arbeidsinhoud en -organisatie. Aspecten op het gebied van arbeidsvoorwaarden en arbeidsomstandigheden zijn bijvoorbeeld buiten beschouwing gelaten. Dit is enerzijds gedaan op inhoudelijke gronden, anderzijds ook op grond van een exploratieve factoranalyse op de schaalscores. Het toevoegen van goede schalen op deze (en andere) terreinen aan zelfbeoordelingsmethoden voor PSA en werkstress zal het totale vragenlijstonderzoek gedifferentieerder maken, waardoor het "algemene onvrede"-argument nog minder van toepassing wordt.

Beslissingsruimte: de som is kleiner dan de delen

In de theorie van Karasek (1979) neemt het begrip "beslissingsruimte" een centrale plaats in. Karasek stelt dit begrip samen uit twee deelbegrippen, namelijk "skill discretion" en "task authority".

De geconstrueerde vragenlijstschalen die met deze concepten corresponderen, zijn: "taakvariatie" en "taakautonomie". In hoofdstuk 7 hebben we (in het beknopte model) gezien, dat "taakvariatie", samen met "leermogelijkheden", de dimensie "veelzijdigheid van het werk" vormt. "Taakautonomie" gaat samen met "inspraakmogelijkheden". De hieruit resulterende dimensie hebben we "regelmogelijkheden" genoemd.

Volgens Karasek (1979) is het niet mogelijk om deze twee begrippen in de praktijk van elkaar te scheiden: het hebben van veelzijdig werk biedt regelmogelijkheden,

het hebben van eenzijdig werk biedt weinig regelmogelijkheden. Deze stellingname wordt door dit onderzoek niet bevestigd. Integendeel: alle resultaten in deel III van dit proefschrift wijzen op het belang van het onderscheiden van "veelzijdigheid" en "regelmogelijkheden".

In de eerste plaats de exploratieve en confirmatorische factoranalyse. Noch in de exploratieve, noch in de confirmatorische factoranalyse vinden we aanwijzingen om "veelzijdigheid" en "regelmogelijkheden" te combineren.

Er is onderzocht in hoeverre de operationalisatie van "veelzijdigheid" thuishoort onder "psychologische taakeisen", en niet onder "regelmogelijkheden". Dit zou het Karasek-model intact laten, zij het dat de "decision latitude"-dimensie dan versmalt tot "task authority". Een model dat "veelzijdigheid" combineert met "psychologische taakeisen" komt echter tot slechtere passingsmaten bij de confirmatorische factoranalyse, dan een model dat "veelzijdigheid" en "psychologische taakeisen" onderscheidt.

Dit wordt later bevestigd door de zeer verschillende relaties die deze drie dimensies hebben met werkstress en ziekteverzuim. Dit bevestigt de validiteit van de gevonden structuur voor PSA en werkstress.

De conclusie dat "veelzijdigheid" en "regelmogelijkheden" als losse begrippen dienen te worden gezien betekent, dat het hieruit samengestelde begrip "decision latitude" (één van de elementen in het door Karasek veronderstelde demand-control interactie-effect) aan kracht verliest. Ter onderbouwing van deze stelling worden hier de correlaties weergegeven tussen de twee "veelzijdigheid"-schalen en de twee "regelmogelijkheden"-schalen (zie tabel 10.1). De correlaties worden apart gepresenteerd voor de ontwikkelgroep (O) (N=409) en de testgroep (T) (N=407), die werden gebruikt in hoofdstuk 7.

Tabel 10.1: Correlaties tussen de schalen uit de dimensies "veelzijdigheid" en "regelmogelijkheden"

variabele	taakvariatie		leermogelijkheden	
	O	T	O	T
gebrek taakautonomie	-.40	-.41	-.36	-.30
gebrek inspraak	-.43	-.54	-.48	-.49

Dat de claim van Karasek enige geldigheid heeft, wordt door deze correlaties bevestigd. De correlaties zijn allemaal redelijk hoog en in de verwachte richting. De gedeelde variantie (R-kwadraat) ligt bij dit niveau van correlaties echter ergens tussen 10% en 25%. Dit is onvoldoende om het introduceren van een samengesteld begrip als "beslissingsruimte" te rechtvaardigen.

10.2 De relatie tussen PSA en werkstress

Confounding

In dit onderzoek zijn twee mogelijke confounders onderzocht: negatieve affectiviteit (NA) en ervaren onbillijkheid (EO). De resultaten zijn bijzonder helder waar het gaat om het confounding effect van deze variabelen. Er zijn weinig aanwijzingen gevonden dat EO optreedt als confounder in onderzoek naar de relatie tussen PSA en werkstress. Dit valt tegen, omdat op grond van de literatuur werd verwacht dat EO zowel samenhangt met symptomen van werkstress als met PSA.

Voor NA daarentegen geldt, dat het risico van confounding erg groot is. Dit geldt met name voor onderzoek naar de relatie tussen PSA en "spanning", maar in mindere mate ook voor onderzoek naar de relatie tussen PSA en "onwelbevinden". Dit resultaat is analoog aan wat werd verwacht op basis van de literatuur (Brief et al., 1988, Parkes, 1990; Chen en Spector, 1991; Parkes et al., 1994). Op grond van dit onderzoek kan gesteld worden dat onderzoek naar de relatie van PSA met "arbeidsgerelateerde spanningsklachten", maar ook met psychische en psychosomatische gezondheidsklachten, zinloos is zonder rekening te houden met NA (Brief et al., 1988).

Een interessante mogelijkheid vanuit de symptoomperceptie-theorie (Pennebaker, 1982) is, dat NA als persoonskenmerk een eigen bijdrage levert aan het ontstaan van gezondheidsproblemen, via de mechanismen van competitie van stimuli en het selectieve zoekproces. Parkes (1990) doelt hierop, wanneer zij spreekt over NA als factor die de reactiviteit beïnvloedt op sommige werkstressoren, onafhankelijk van het contaminerende effect op de meting van stressoren en stress-uitkomsten. Interessant is ook haar hypothese dat sociaal-organisatorische factoren vooral te maken krijgen met een methodologisch contaminatie-effect van NA, terwijl kwantitatieve taakeisen vooral te maken krijgen met het reactiviteit-effect van NA.

Aansluitend bij Parkes (1990), lijkt het inderdaad al te simpel om het grote effect van NA op relaties tussen PSA-factoren en "spanning" puur te interpreteren als methodologische ruis. Het gegeven dat NA zoveel van de variantie verklaart in "spanning" bij de multi-level analyse, wijst eveneens in de richting van factoren in het individu, die bepalend zijn voor de perceptie van spanningsverschijnselen. De symptoomperceptie-hypothese verdient dan ook meer aandacht in de toekomst bij onderzoek naar "spanningsverschijnselen".

Het voornaamste resultaat uit het derde deel van dit proefschrift wat betreft het onderwerp confounding is, dat er inderdaad sprake is van confounding, maar dat PSA boven op de confounders een unieke eigen bijdrage levert aan de verklaring van werkstress en ziekteverzuim. Er is geen reden om op grond van dit onderzoek te twijfelen aan de waarde van een zelfbeoordelingsmethode voor onderzoek naar de relatie tussen PSA en werkstress.

In paragraaf 8.7 bleek dat de relatie tussen PSA- en werkstress-dimensies in een aantal gevallen kromlijng is. Dat dit fenomeen kon worden aangetoond met vragenlijstgegevens, pleit tegen het optreden van de "triviality trap". Indien er bijvoorbeeld bij het meten van "psychologische taakeisen" en "spanning" sprake zou zijn van een sterke mate van confounding door item-overlap en/of inhoud-overlap, dan is niet makkelijk te verklaren waarom zowel hoge als lage waarden van de schaal "psychologische taakeisen" samenhangen met hoge waarden op de schaal "spanning".

Individu versus organisatie

Met behulp van multi-level analyse is in dit proefschrift een materie aangeroerd, die betrekkelijk nieuw is. Hoe vreemd het ook klinkt: het merendeel van het onderzoek op het gebied van arbeid en organisatie heeft tot nu toe weinig rekening gehouden met het gegeven dat individuele werknemers niet random zijn verdeeld over de organisatie. Uitzonderingen hierop zijn onderzoek dat plaatsvindt op geaggregeerde gegevens van beroepen of afdelingen (Karasek en Theorell, 1990) en onderzoek naar werkstress-uitkomsten als latente begrippen die worden verklaard uit afzonderlijke metingen van meer individuen en/of gegevensbronnen ("shared well-being": Semmer et al., 1995).

In feite gaat het bij arbeids- en organisatie-onderzoek altijd om individuen die werken binnen groepen (afdelingen, functies), die weer zijn te plaatsen binnen organisaties. Een dergelijke geneste gegevensstructuur is het uitgangspunt voor multi-level analyse. Deze techniek kan dan ook vragen beantwoorden, die met andere technieken niet zijn te beantwoorden. Vooral de mogelijkheid om verklaarde variantie op meer verklaringsniveaus tegelijkertijd en ten opzichte van elkaar te onderzoeken biedt interessante mogelijkheden. Een aantal van deze vragen is aan de orde gekomen in hoofdstuk 8 van dit proefschrift. We resumeren hier kort deze vragen en de gevonden resultaten.

Uit de stand der techniek van onderzoek op het terrein van arbeid en organisatie is duidelijk, dat er een limiet is in het maximaal haalbare, verklaarbare deel van de variantie in "spanning" en "welbevinden". Als de te verklaren variantie echter wordt gesplitst in niveaus: hoeveel hiervan ligt dan bij het individu en hoeveel bij de organisatie?

Het antwoord op deze vraag is, dat zowel voor "spanning" als voor "welbevinden" het grootste deel van de te verklaren variantie ligt op individuniveau. Opmerkelijk is echter, dat het boven-individuele deel bij "welbevinden" (19%) twee maal zo groot is als bij "spanning" (9%). Er is meer consensus tussen werknemers binnen groepen binnen organisaties over de beleving en beoordeling van "welbevinden", dan van "spanning". Dit is een belangrijk resultaat.

Een volgende vraag is, hoeveel variantie verklaarbaar is op elk van de niveaus. Hier levert het onderzoek eveneens een opmerkelijk resultaat.

Nagenoeg alle variantie op boven-individueel niveau (zowel bij "spanning" als bij "welbevinden") is verklaarbaar op basis van een beperkt aantal variabelen. Nagenoeg alle residu-variantie zit op het niveau van het individu. Dit is conform wat Semmer et al. (1995) vinden met hun concept van "shared well-being".

Belangrijk is, dat PSA van het boven-individuele deel van de variantie in werkstress-uitkomsten een groot deel verklaart. Dit geldt bij "arbeidsgerelateerde spanningsverschijnselen", maar vooral bij "arbeidsgerelateerd welbevinden". Dit betekent dat consensus tussen werknemers in groepen in organisaties qua beleving van welbevinden, relatief sterk samenhangt met consensus tussen dezelfde werknemers qua beoordeling van PSA. De resultaten zoals die in dit onderzoek zijn verkregen op het boven-individuele niveau, moeten overigens met enige voorzichtigheid worden geïnterpreteerd vanwege het betrekkelijk kleine aantal afdelingen en organisatie-onderdelen dat is onderzocht.

Dit onderzoek geeft aan dat praktijkonderzoek naar "spanningsverschijnselen" zich voornamelijk zou moeten bewegen op het niveau van de individuele werknemer. Bij "welbevinden" is het zinvoller dan bij "spanning", om naast het individuele niveau ook het organisatieniveau te onderzoeken.

Eén of meer werkstressmechanismen

In hoofdstuk 7, 8 en 9 zijn resultaten gepresenteerd, die ingaan op de vraag of er één algemeen werkstressmechanisme bestaat, dan wel meer dan één, parallel opererende mechanismen. De theorieën van Karasek en Warr zijn gebruikt als prototypisch voor deze twee mogelijkheden. Karasek (1979) veronderstelt dat de interactie tussen psychologische taakeisen en beslissingsruimte werkstress veroorzaakt. Er is in zijn theorie geen plek voor specifieke werkstress-effecten ten gevolge van specifieke PSA-factoren. Dit algemeen mechanisme is door Karasek en Theorell (1990) verbreed tot de factor sociale ondersteuning. Deze uitbreiding laat het uitgangspunt van één algemeen werkstress-mechanisme echter intact.

Warr's theorie (1987, 1994) is te beschouwen als een reeks van amendementen op de stellingnames van Karasek. In de eerste plaats is er volgens Warr sprake van differentiatie in typen oorzaken. Er zijn typen PSA die een tekort kunnen vertonen, maar nooit een overdosis (CE-factoren: beloning, veiligheid, sociale waardering). Er zijn echter ook factoren die een tekort kunnen vertonen, maar die bij een overdosis leiden tot "toxische reacties" (AD-factoren: hoge taakeisen, weinig regelmogelijkheden et cetera). In de tweede plaats is er bij de "toxische reacties" sprake van differentiatie in gevolgen. Er zijn AD-factoren die vooral overbelasting en stress oproepen: hoge prestatie-eisen, te groot beroep op vaardigheden, en gebrek aan regelmogelijkheden. Er zijn echter ook AD-factoren die vooral leiden tot lage arbeidsmotivatie en -satisfactie. Dit zijn factoren als: teveel duidelijkheid en teveel afwisseling in het werk.

De resultaten uit hoofdstuk 7, 8 en 9 wijzen allemaal in de richting van een differentiatie in werkstress-mechanismen. Dit ondersteunt de theorie van Warr, en sluit minder goed aan bij de theorie van Karasek. De resultaten van dit proefschrift komen op hoofdlijnen overeen met de resultaten van ander recent Nederlands onderzoek op dit gebied (Furda, 1995; De Jonge, 1995; De Jonge en Furda, 1995).

Het ontstaan van "spanning" lijkt te verlopen volgens een ander mechanisme dan het ontstaan van "onwelbevinden".

In de eerste plaats blijkt de bruto verdeling van de variantie over de drie verklaringsniveaus, die in hoofdstuk 8 zijn onderzocht (werknemer, afdeling, organisatie), verschillend te zijn voor "spanning" en "welbevinden".

In de tweede plaats hebben "spanning" en "welbevinden" andere PSA-voorspellers. Hierbij maakt het verschil of uitsluitend wordt gekeken naar lineaire effecten, of naar een combinatie van lineaire en non-lineaire effecten. Wanneer alleen lineaire verbanden worden onderzocht, is het resultaat dat "psychologische taakeisen" sterk samenhangt met "spanning", en "veelzijdigheid" sterk samenhangt met "welbevinden". "Sociale steun" hangt samen met zowel "spanning" als "welbevinden", zij het sterker met "welbevinden" dan met "spanning". "Regelmogelijkheden" hangt met geen van beide werkstress-dimensies samen.

Wanneer zowel lineaire als non-lineaire effecten worden onderzocht, worden opnieuw alle relaties gevonden die in het voorgaande zijn beschreven. Er zijn echter twee aanvullingen. "Regelmogelijkheden" vertoont nu een zwakke samenhang met "spanning". "Psychologische taakeisen" heeft nu een sterke samenhang met "welbevinden". Met name dit laatste effect is onverwacht, en komt niet voor in het vitamine-model van Warr.

Wanneer de sterkte van de gevonden relaties wordt betrokken bij de interpretatie, dan is het verschil tussen de voorspellers van "spanning" en "welbevinden" het beste als volgt weer te geven. Voor "spanning" wordt slechts één duidelijke determinant gevonden: "psychologische taakeisen". "Welbevinden" vertoont een duidelijke samenhang met zowel "veelzijdigheid", als "psychologische taakeisen", als "sociale ondersteuning". Kort geformuleerd: "Welbevinden" heeft meer PSA-determinanten dan "spanning".

Ten derde zijn "spanning" en "welbevinden" in verschillende mate vatbaar voor confounding door NA. Geheel zoals van een individugebonden variabele als NA verwacht mag worden, is het effect bij "spanning" groter dan bij "welbevinden". Bij "spanning" is immers de ruwe variantie op individuniveau groter dan bij "welbevinden".

Ten vierde blijkt bij het onderzoek van de relaties van PSA en werkstress met ziekteverzuim, dat bij "spanning" een relatie met verzuim blijft bestaan terwijl deze relatie bij "welbevinden" verdwijnt, wanneer directe verbanden tussen PSA en verzuim worden opgenomen in het model.

De PSA-voorspellers van "spanning" lijken dus vooral een indirect effect op verzuim te hebben, via "spanning". De voorspellers van "welbevinden" hebben een direct effect op verzuim, niet gemedieerd door "welbevinden".

Dit onderzoek levert derhalve vier aanwijzingen op dat het mechanisme bij het ontstaan van "spanning" verschilt van dat bij het ontstaan van "welbevinden". Er zijn op basis van de gevonden dimensie-structuur twee afzonderlijke werkstress-mechanismen te onderscheiden. Onduidelijk is, of er naast "spanning" en "welbevinden" nog andere werkstress-mechanismen bestaan, zoals Warr (1987; 1994) suggereert.

Opmerkelijk is de bevinding bij "regelmogelijkheden". Er is geen lineair effect van "regelmogelijkheden" gevonden: noch op "spanning", noch op "welbevinden". Er is wel een non-lineair niet-monotoon effect gevonden van "regelmogelijkheden" op "spanning", maar dit effect is toch redelijk zwak. Er is een aantal interpretaties mogelijk van deze bevinding. De eerste mogelijkheid is dat regelmogelijkheden geen determinant zijn van werkstress. Er is echter een grote hoeveelheid onderzoek die deze mogelijkheid afwijst (Ganster en Fusilier, 1989; Furda en Meijman, 1992). Een andere mogelijkheid is, dat het aspect "regelmogelijkheden" niet valide wordt gemeten met de hiervoor geconstrueerde schalen. Hiertegen pleiten de gunstige en eenduidige resultaten van de schaal- en factoranalyses. Het is niet eenvoudig om voor één van deze twee verklaringen te kiezen. Wellicht dat de zwakte van de dimensie "regelmogelijkheden" erin schuilt, dat er geen differentiatie is aangebracht naar de aard van de "regelmogelijkheden". Gegeven de overige resultaten van het proefschrift op het gebied van de dimensionalisering van PSA, lijkt het van belang om "regelmogelijkheden" in ieder geval te onderscheiden in twee soorten: "regelmogelijkheden ten aanzien van psychologische taakeisen" en "regelmogelijkheden ten aanzien van de veelzijdigheid van het werk". Deze suggestie sluit aan bij en wordt verder aangevuld door het volgende citaat van Furda en Meijman (1992, p. 144):

"Uit de eerdere besprekingen van de facetten die aan het sturingsbegrip zijn te onderscheiden moet naar onze mening geconcludeerd worden dat het werkstress-onderzoek aan kracht zou winnen wanneer sturingsmogelijkheden met behulp van een multidimensioneel instrument zouden worden onderzocht. Bovendien zou meer aandacht gegeven moeten worden aan de betekenis van het doel waarop de sturing is gericht. Tal van aspecten in werksituaties zijn eenvoudigweg nauwelijks te beïnvloeden, en mensen zijn in hun beoordeling daarvan meestal zeer realistisch. Dat wil zeggen dat meer aandacht gegeven zou moeten worden aan het feit dat mensen op de verschillende aspecten van hun arbeidstaak en werksituatie verschillende niveaus van invloed hebben en/of zouden willen hebben".

Op het gebied van "regelmogelijkheden" en de relatie hiervan met "spanning" en "welbevinden", is vervolgonderzoek en verdere instrumentontwikkeling nodig.

Lineaire en non-lineaire effecten

In de theorie van Warr (1987; 1994) spelen non-lineaire effecten een belangrijke rol. In paragraaf 8.7 is onderzocht in hoeverre dergelijke effecten optreden voor de vier PSA-dimensies in relatie tot "spanning" en "welbevinden".

De relaties van "psychologische taakeisen" en "regelmogelijkheden" met "spanning" blijken non-lineair en niet-monotoon te zijn. De relatie tussen "sociale steun" en "spanning" is lineair. "Veelzijdigheid" heeft geen effect op "spanning". Met "welbevinden" blijken "veelzijdigheid" en "psychologische taakeisen" non-lineair en niet-monotoon samen te hangen. "Sociale ondersteuning" heeft een lineaire relatie met "welbevinden". "Regelmogelijkheden" heeft geen relatie met "welbevinden".

Uit deze resultaten is de conclusie te trekken dat er voor "psychologische taakeisen", "veelzijdigheid" en "regelmogelijkheden" optimale niveaus bestaan in relatie tot werkstress. Zowel te weinig als te veel van deze factoren roept negatieve effecten op. Bij "sociale steun" is aangetoond dat een tekort een algemeen, negatief effect heeft op werkstress.

Deze resultaten zijn van groot belang. De aard van de verbanden tussen PSA-factoren en werkstress-uitkomsten is een onderwerp van discussie in de literatuur. Het is moeilijk om tussen alle argumenten die worden aangereikt voor additieve effecten, interactie-effecten en non-lineaire effecten een keus te maken. Onderzoeksresultaten die door de één voortdurend worden gerepliceerd (bijvoorbeeld het demand-control interactie-effect: Karasek, 1979; Karasek en Theorell, 1990), worden door anderen niet gevonden (Spector, 1987; Warr, 1990a; Parkes, 1990; Fletcher en Jones, 1993).

In dit onderzoek is de keuze gemaakt om in eerste instantie geen interactie-effecten te onderzoeken. De keuze is gemaakt om additieve en non-lineaire effecten te onderzoeken. Het model van Warr fungeerde hierbij als leidraad en werd, op die delen die onderzocht konden worden, grotendeels bevestigd.

Een opvallende afwijking van het vitamine-model werd gevonden voor "psychologische taakeisen" in relatie tot "welbevinden". Hiertussen blijkt een non-lineair niet-monotoon verband te bestaan, dat niet door Warr wordt verondersteld.

Meer onderzoek is nodig om het bestaan van de door Warr gepostuleerde lineaire en non-lineaire relaties te toetsen in uiteenlopende populaties. Wanneer op deze manier goed inzicht is verkregen in de relatie tussen afzonderlijke (soorten) PSA-factoren en afzonderlijke (soorten) werkstress-uitkomsten, rijst vervolgens de vraag hoe diverse PSA-factoren eventueel samenwerken in het ontstaan van bepaalde specifieke werkstress-uitkomsten.

Hoe deze samenwerking er uit ziet (drempel-effect, optelsom, interactie-effect) is dan te onderzoeken. De lijn volgend die Warr heeft uitgezet, is het ook mogelijk dat de manier waarop bepaalde PSA-factoren met elkaar samen een werkstress-effect bewerkstelligen, verschilt tussen diverse combinaties van factoren. De manier waarop "psychologische taakeisen" interacteert met "veelzijdigheid" hoeft bijvoorbeeld niet hetzelfde te zijn als de manier waarop "psychologische taakeisen" interacteert met "regelmogelijkheden".

10.3 De voorspelling van ziekteverzuim

In hoofdstuk 9 is onderzocht in hoeverre ziekteverzuim (geregistreerd gedurende een half jaar na de vragenlijstafname) voorspeld kan worden op basis van vragenlijstgegevens over PSA en werkstress. Tevens is onderzocht of NA en EO een rol spelen bij het verklaren van ziekteverzuim.

Met behulp van LISREL is een model ontwikkeld dat - zowel voor de verzuimfrequentie als voor de verzuimduur - voldoet aan de criteria van een AGFI $>.90$ en een RMSR van $<.05$. Voor de verzuimduur geldt tevens dat dit model voldoet aan het criterium van een Chi-kwadraat met een p-waarde $>.05$. Dit model kan als volgt worden omschreven:

PSA + NA = spanning

PSA + NA = welbevinden

PSA + spanning + welbevinden + EO = verzuim

De eerste twee vergelijkingen zijn gebaseerd op de kennis die in hoofdstuk 7 en 8 is opgebouwd over de structuur van en relatie tussen PSA en werkstress. Ook in dit LISREL-model blijkt "spanning" vooral samen te hangen met "psychologische taakeisen" en "sociale steun". "Welbevinden" blijkt opnieuw vooral samen te hangen met "veelzijdigheid" en "sociale steun". NA levert in dit LISREL-model opnieuw een belangrijke bijdrage in de relatie tussen PSA en "spanning". Hetzelfde geldt voor NA met betrekking tot de relatie tussen PSA en "welbevinden", zij het in mindere mate. Ook bij dit LISREL-model wordt gevonden dat "regelmogelijkheden" noch met "spanning", noch met "welbevinden" een duidelijke samenhang heeft.

De derde vergelijking in het model gaat in op de voorspelling van ziekteverzuim. De resultaten voor de verzuimfrequentie en de verzuimduur zijn enigszins verschillend:

Verzuimfrequentie

De totale verklaarde variantie in de verzuimfrequentie op basis van de derde vergelijking is 6%. Er worden significante effecten gevonden voor "spanning" en

"veelzijdigheid". Voor "welbevinden" wordt geen effect op de verzuimfrequentie gevonden. Vóór het invoeren van de vier PSA-factoren in de vergelijking (Model 2 in hoofdstuk 9) werd wél een effect van "welbevinden" gevonden op de verzuimfrequentie. Dit effect verdwijnt bij de introductie van PSA in de vergelijking, ten bate van een significant effect van "veelzijdigheid". Hierin verschillen "welbevinden" en "spanning": "spanning" behoudt wel een eigen effect op verzuim bij introductie van PSA in de vergelijking. Deze bevindingen over PSA en "welbevinden" in relatie tot de verzuimfrequentie interpreterend tegen de achtergrond van het herziene model belasting-belastbaarheid, kan worden gezegd dat "welbevinden" meer een belastingsgevolg lijkt te zijn dan een belastingsverschijnsel.

Verzuimduur

De totale verklaarde variantie op basis van de derde vergelijking bedraagt 8%. Ook bij de verzuimduur zijn er significante effecten van "spanning" en "veelzijdigheid". Het effect van "veelzijdigheid" heeft dezelfde consequenties voor de bijdrage van "welbevinden", als bij de verzuimfrequentie. Bij de verzuimduur blijkt ook EO een significant effect te hebben, in tegenstelling tot bij de verzuimfrequentie. EO is bij de verzuimduur zelfs de sterkste predictor in de vergelijking.

De resultaten ten aanzien van EO zijn in tegenspraak met de resultaten van Geurts (1994). Zij vond namelijk sterke effecten van EO op de verzuimfrequentie bij medewerkers van metaalbedrijven en bij buschauffeurs. Wij vinden deze effecten niet terug. De claim dat de ziekteverzuimfrequentie vooral wordt gedetermineerd door sociale vergelijkingsprincipes, wordt niet bevestigd door dit onderzoek. Voor PSA en werkstress zijn dergelijke effecten op de verzuimfrequentie wél gevonden. Bij de verzuimduur is weliswaar een effect van EO gevonden, maar hier leveren factoren op het gebied van PSA en werkstress een additionele bijdrage, bovenop de bijdrage van EO. Ook voor de verzuimduur kan de claim dat "verzuim vooral dient te worden gezien als compensatiegedrag", derhalve niet worden bevestigd. Op grond van dit onderzoek is het beter om te stellen dat bij een deel van de verzuimgevallen de duur van de gevallen wordt beïnvloed door compensatiegedrag van de werknemers.

Bij de resultaten ten aanzien van verzuim dient als voornaamste kanttekening te worden geplaatst dat nergens meer dan 8% van de variantie kon worden verklaard, waarbij de minimale verklaarde variantie 3% was.

Dit resultaat is geheel in lijn met wat werd verwacht op basis van eerder onderzoek (Spector en Jex, 1991; Smulders en De Winter, 1993). Voor ziekteverzuim geldt het argument van Semmer et al. (1995) omtrent de limiet aan de verklaarbare variantie in werkstress op basis van PSA alleen. Ziekteverzuim is een multicausaal fenomeen. Het zoeken naar oorzaken van verzuim binnen één bepaald domein (bijvoorbeeld PSA), is beperkt tot een maximaal haalbaar verklaringsniveau van circa 5% à 10% van de variantie.

Wellicht is het beter om aan de slag te gaan met het verklaren van "gedeeld ziekteverzuim", dat wil zeggen: het verzuim dat werknemers binnen een bepaalde groep (afdeling, beroep) gemeenschappelijk hebben. Dit zou betekenen dat een voorspelling van het totale verzuim van individuele medewerkers wordt vervangen door een voorspelling van specifieke soorten verzuim van groepen medewerkers, bijvoorbeeld het aantal gevallen van overspannenheid, de lengte van verzuimgevalen door rugproblemen et cetera.

Op een dergelijk gegevensbestand is er voor oorzakelijke verzuimfactoren, onder andere ook voor PSA, meer eer te behalen. Ook alleen op een dergelijk bestand is de vraag op te lossen welke oorzakelijke factoren voor welk type verzuim het meest bepalend zijn. Het is jammer dat dergelijke bestanden in Nederland op dit moment ontbreken, terwijl de maatschappelijke behoefte aan dit type onderzoek groot is.

Ook het onderzoek van non-lineaire verbanden tussen PSA, werkstress en verzuim is wellicht een optie voor toekomstig onderzoek.

10.4 Integratie: een multi-level model voor de beleving en beoordeling van PSA en werkstress

De resultaten van dit onderzoek geven aan dat "arbeidsgerelateerde spanning" een sterk individugebonden fenomeen is. De belangrijkste voorspellers hiervan zijn "psychologische taakeisen" en "regelmogelijkheden". Zowel "psychologische taakeisen" als "regelmogelijkheden" hebben een kromlijngige relatie met "spanning". Zowel te veel als te weinig van deze factoren gaat gepaard met "spanning". Het kromlijngig verband van "regelmogelijkheden" met "spanning" is betrekkelijk zwak. Naast deze effecten heeft "sociale ondersteuning" een lineair effect op "spanning", maar niet zo overtuigend als het effect van "psychologische taakeisen". "Spanning" heeft derhalve hoofdzakelijk één determinant: "psychologische taakeisen".

NA speelt een belangrijke rol bij het interpreteren van "psychologische taakeisen" en hieraan gerelateerde "spanningsverschijnselen". Wanneer we dit resultaat bekijken vanuit de belasting-belastbaarheid traditie, dan vinden we hierin een bevestiging van het begrip "verwerkingsvermogen", zoals dat is geformuleerd door Van Dijk et al. (1990) en Meijman (1989).

Met al de aandacht in de media en in het vakgebied voor het onderwerp werkdruk is dit een belangrijk resultaat. Het is niet zozeer het gegeven dat werkdruk een individugebonden proces is waarbij persoonsgebonden factoren een belangrijke rol spelen, dat van belang is. Het is belangrijker te constateren hoe klein de relatie tussen PSA en "spanning" is op organisatieniveau: het niveau van verklaarde variantie is hier dusdanig laag, dat een collectieve werkstress-aanpak op dit gebied weinig aansluit bij de werkelijkheid. Het zal er altijd om gaan het individu te bereiken via zijn/haar persoonlijke beleving en beoordeling.

Op het gebied van "welbevinden" gaan de conclusies van dit onderzoek juist de andere kant uit. Een aantal zaken versterkt elkaar. Een groter deel van de te verklaren variantie zit op collectief niveau. De mate van confounding door NA is kleiner. De bijdrage van PSA aan de verklaarde variantie is groter. Dit geldt op individueel niveau, maar vooral ook op collectief niveau. "Veelzijdigheid" en "psychologische taakeisen" zijn de belangrijkste voorspellers. Ook deze verbanden zijn kromlijng. Zowel te veel als te weinig van deze factoren gaat dus gepaard met "onwelbevinden". Naast deze twee effecten wordt ook een relatief sterk lineair effect gevonden van "sociale ondersteuning" op "welbevinden". "Welbevinden" heeft, met uitzondering van "regelmogelijkheden", alle PSA-dimensies als determinant.

"Welbevinden" lijkt, in vergelijking tot "spanning", een meer groeps- en organisatiegebonden fenomeen te zijn. Alhoewel in dit onderzoek niet veel aandacht is besteed aan het begrip "organisatiecultuur", is dit wel een begrip dat in gedachten komt bij het zien van deze resultaten. Het is bij de motivatie en satisfactie van werknemers van belang hun gemeenschappelijke beleving en beoordeling van het werk te kennen. Deze constatering houdt in dat toepassing van de sociale vergelijkingstheorie juist op het gebied van "arbeidsgerelateerd welbevinden" een interessante bijdrage zou kunnen leveren. Hierbij is niet alleen van belang wat de sociale vergelijkingstheorie heeft te zeggen over hoe individuen binnen groepen hun sociale werksituatie construeren, maar ook hoe groepen werknemers elkaar in dit opzicht beïnvloeden.

De in dit proefschrift gebruikte operationalisatie van sociale vergelijking, namelijk EO qua beloning, leverde niet het resultaat op dat ervan werd verwacht. Meer onderzoek is wenselijk naar de rol van sociale vergelijking bij het ontstaan van "arbeidsgerelateerd welbevinden".

In tegenstelling tot het gestelde bij "spanning" lijkt een collectieve werkstress-aanpak op het gebied van "welbevinden" wèl zinvol te kunnen zijn, bovenop een individugerichte benadering.

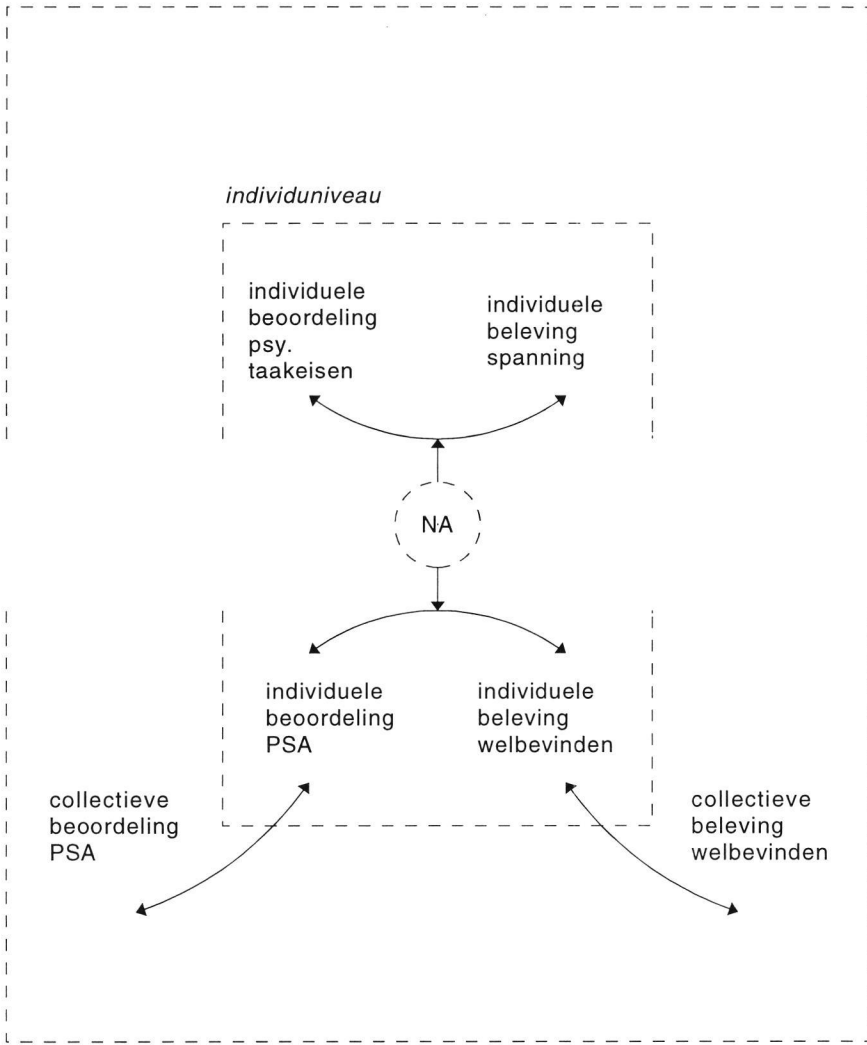
De integratie van de onderzoeksresultaten in dit proefschrift is weergegeven in figuur 10.1. De figuur beperkt zich tot het in één model plaatsen van de meest belangrijke resultaten, die in deel III van dit proefschrift zijn verkregen. De resultaten uit hoofdstuk 8 wegen hierbij het zwaarst.

De terminologie volgend die in deze tekst is gehanteerd, kunnen we de onderstaande figuur een multi-level model noemen voor de beleving en beoordeling van PSA en werkstress.

Figuur 10.1 geeft de verschillen weer in de mechanismen die leiden tot "spanning" en "welbevinden". Het "spanning"-mechanisme speelt zich uitsluitend af op individueel niveau, het "welbevinden"-mechanisme speelt zich af op zowel individueel als collectief niveau. De determinanten van "spanning" zijn beperkt tot één dimensie, namelijk "psychologische taakeisen".

"Welbevinden" wordt gedetermineerd door een breed scala aan PSA-dimensies. NA speelt bij beide mechanismen een intermediaire rol.

organisatieniveau



Figuur 10.1: Multi-level model voor de beleving en beoordeling van PSA en werkstress

10.5 Beperkingen van het onderzoek

Het onderzoek, zoals het in dit proefschrift is gerapporteerd, heeft een aantal beperkingen. In deze paragraaf worden deze beperkingen besproken.

Een eerste beperking is, dat alle maten van PSA, werkstress en confounders berusten op zelfbeoordelingen. Alhoewel de grootste zorg is besteed aan de schaalconstructie en aan het voorkomen van item-overlap, blijven de resultaten kwetsbaar voor de "mono-method" kritiek. Dit wordt bevestigd door de resultaten die zijn verkregen met de enige niet-vragenlijstgegevens in dit onderzoek, namelijk: de ziekteverzuimgegevens. De correlaties van de vragenlijstmaten met verzuim liggen over het algemeen stukken lager dan de correlaties tussen de vragenlijstmaten onderling. Dit fenomeen is niet anders dan op basis van de literatuur te verwachten was (Watson en Pennebaker, 1989; Spector en Jex, 1991; Spector, 1992).

De tweede beperking zit hierin, dat alle vragenlijstgegevens die in dit onderzoek zijn gebruikt cross-sectioneel zijn verzameld. Op geen enkele wijze is gewerkt met herhaalde metingen. Alle uitspraken over relaties tussen PSA en werkstress in dit proefschrift, dienen dan ook begrepen te worden als uitspraken over samenhangen, niet over oorzakelijke verbanden. Een uitzondering hierop vormt hoofdstuk 9 over verzuim. Hier is gewerkt met prospectief verzamelde verzuimgegevens, zodat de mate van voorspelbaarheid van verzuim daadwerkelijk kon worden onderzocht.

Longitudinaal onderzoek met aanvullende gegevens als verzuim, verloop en productiviteit zou een goede test bieden van de validiteit en praktijkwaarde van de in dit proefschrift beschreven schalen, dimensies en de relaties daartussen.

Een derde beperking heeft betrekking op de inhoudelijke dekking in het derde deel van het proefschrift. We hebben aan het begin hiervan expliciet gekozen om ons te beperken tot arbeidsinhoud, arbeidsorganisatie, en "arbeidsgerelateerde werkstressverschijnselen". Het gebied PSA en werkstress is hierdoor grotendeels, maar zeker niet uitputtend, afgedekt.

Het is aannemelijk dat terreinen als "arbeidsvoorwaarden" meerdere dimensies bevatten (bijvoorbeeld beloning, werk- en rusttijden, loopbaanperspectief, toekomstonzekerheid), die bij het ontstaan van werkstress een rol spelen. Het is ook aannemelijk dat werkstress doorwerkt in meer algemene uitkomstmaten, zoals gezondheidsklachten en medische consumptie. De resultaten van dit proefschrift (zie paragraaf 10.4) mogen dan ook niet zonder meer vertaald worden naar de praktijk, zonder ervoor te waarschuwen dat in concrete arbeidssituaties ook heel andere factoren een belangrijke rol kunnen spelen.

In paragraaf 10.2 werd gevonden dat een belangrijk deel van de ruwe variantie in werkstress-uitkomsten ("spanning" en "welbevinden") op individueel niveau zit.

In dit verband is het jammer dat in de onderzoeken die het materiaal hebben geleverd voor dit proefschrift weinig gegevens zijn verzameld over het verwerkingsvermogen van de werknemers.

Dit komt omdat bij de schaalconstructie is gestreefd naar een instrument dat zoveel mogelijk aansluit bij de stand der techniek voor het meten PSA en werkstress. In hoofdstuk 5 werd duidelijk dat een eenduidige operationalisatie van het begrip verwerkingsvermogen nog niet tot deze stand der techniek behoort. De bij het schaalconstructie-onderzoek verzamelde gegevens zijn later gebruikt voor het beantwoorden van inhoudelijke vragen, met als gevolg dat op het gebied van persoonskenmerken en coping weinig materiaal voor handen was. Een studie die primair zou zijn opgezet vanuit een inhoudelijke doelstelling, zou dit waarschijnlijk anders hebben aangepakt.

Dit brengt ons bij de variabelen NA en EO. In beide gevallen is getracht een operationalisatie te maken van deze begrippen, op basis van het uit het schaalconstructie-onderzoek beschikbare materiaal (zie paragraaf 8.1 en 8.2). Voor NA lijkt dit aardig te zijn gelukt. Het confounding-effect is bevestigd, evenals de hypothese dat dit vooral zou optreden bij maten voor "spanning" en minder bij maten voor "welbevinden".

Voor EO is het resultaat minder duidelijk. Noch de verwachting dat dit begrip zou optreden als confounder, noch de verwachte relatie met de verzuimfrequentie werd in dit onderzoek teruggevonden. De resultaten, verkregen op het gebied van EO, dienen dan ook met enige voorzichtigheid te worden geïnterpreteerd.

De operationalisatie van EO met slechts twee items kan debet zijn aan het uitblijven van de verwachte resultaten. Anderzijds is het heel gangbaar om in onderzoek in deze traditie te werken met operationalisaties van EO, die gebruik maken van slechts één of twee items. Een algemeen aanvaarde schaal met goede kwaliteiten ontbreekt nog.

Een andere mogelijke verklaring voor het achterblijven van de verwachte resultaten is, dat het vragen naar de ruilrelatie tussen werknemer en organisatie qua beloning, toch andere resultaten oplevert dan wanneer de aard van de ruilrelatie wordt opengelaten ("ik stop meer in de organisatie dan ik ervoor terugkrijg"). In onderzoek zoals uitgevoerd door Geurts (1994), is met een dergelijke brede operationalisatie van EO gewerkt. Voor vervolgonderzoek naar EO in relatie tot PSA en werkstress, verdient een brede operationalisatie de voorkeur.

Het is verder de vraag in hoeverre de in dit proefschrift gevonden dimensiestructuur en de gevonden relaties tussen de dimensies, kunnen worden gegeneraliseerd naar andere bedrijven. Dit probleem speelt met name voor de resultaten uit hoofdstuk 8 en 9, die immers zijn verkregen bij slechts één organisatie, namelijk een gezondheidszorginstelling.

Bij hoofdstuk 8 dient nog een extra opmerking te worden gemaakt. Bij de onderzochte organisatie is sprake van een relatief beperkt aantal afdelingen en sectoren.

Hierdoor moeten de conclusies over het organisatiegebonden deel van de variantie met enige terughoudendheid worden geïnterpreteerd.

Het is wenselijk om op een groot gegevensbestand, waarin allerlei soorten organisaties en typen arbeid zijn vertegenwoordigd, te onderzoeken of het resultaat van de multi-level analyse klopt, dat "spanning" een meer individugebonden proces is, en "welbevinden" een meer organisatiegebonden proces.

In hoofdstuk 8 is een aantal voorspellingen uit het model van Warr (1987; 1994) getoetst. Hierbij is de analyse beperkt tot die relaties tussen stressoren en uitkomsten, die beschikbaar waren binnen de in dit proefschrift ontwikkelde dimensie-structuur. Het model van Warr is breder dan dit. Zo veronderstelt hij ook een relatie tussen rolonduidelijkheid en een gevoel van controle en/of groei. Dit is in hoofdstuk 8 niet aan bod geweest. Ook factoren als beloning en veiligheid hebben een plek in zijn model, en zijn hier niet getoetst. Verder is er de beperking dat de concepten tussen het vitamine-model en de in dit proefschrift ontwikkelde dimensie-structuur niet altijd perfect aansluiten. Dit geldt met name voor de dimensies "veelzijdigheid" en "sociale ondersteuning". Deze zijn beide te interpreteren in termen van meer dan één aspect uit het vitamine-model.

De resultaten in hoofdstuk 8, met name in paragraaf 8.7, mogen dan ook niet worden opgevat als een volledig bewijs voor de theorie van Warr, ook al steunen de resultaten de theorie op hoofdlijnen. Het zal duidelijk zijn dat de theorie van Warr, door de hoeveelheid variabelen en onderlinge relaties daartussen, ook niet eenvoudig in zijn volledigheid te toetsen is. Gegeven de resultaten die in hoofdstuk 8 zijn gevonden, verdient de theorie van Warr meer aandacht in de werkstressliteratuur. Een meer volledige test van zijn theorie, inclusief lineaire en non-lineaire verbanden, is wenselijk.

Al deze beperkingen ten spijt levert dit proefschrift een aantal interessante, wetenschappelijke antwoorden en vervolgvragen op. Dit proefschrift werd echter geïnitieerd vanuit een praktijksituatie. Het is dan ook van belang om te evalueren wat de opbrengst is van dit proefschrift voor de praktijk van advisering over PSA en werkstress.

10.6 Betekenis voor de praktijk

In deel II is een verzameling schalen ontwikkeld voor onderzoek naar PSA en werkstress met behulp van een zelfbeoordelingsmethode. Deze verzameling van schalen is een flexibel instrument, dat naar eigen inzicht kan worden aangepast (keuze van schalen per onderzoek) en/of aangevuld (bijvoorbeeld op het gebied van persoonskenmerken en coping). De beschikbaarheid van een dergelijk instrument maakt praktijkonderzoek naar PSA en werkstress mogelijk. Behalve deze instrumentele opbrengst, levert dit proefschrift voor de praktijk echter ook inhoudelijk en methodologisch een aantal inzichten op.

De vorm van vragenlijsten voor onderzoek naar PSA en werkstress

De multi-level analyse heeft een resultaat opgeleverd dat voor de onderzoekspraktijk op het gebied van PSA en werkstress van groot belang is. Voor zover de variantie in "spanning" en "welbevinden" van werknemers "gedeeld" is, dat wil zeggen: voor zover die variantie op een boven-individueel niveau zit, is deze goed te verklaren met een beperkt aantal predictoren. Dit beeld kan eventueel wat geflatteerd zijn door het betrekkelijk kleine aantal afdelingen en sectoren in de analyse.

Dit betekent voor de praktijk, dat het deel van de problematiek in "spanning" en "welbevinden" dat voor bedrijven het meest beïnvloedbaar is (het niet puur individuele stuk), goed te onderzoeken is met een qua onderwerpen beperkt instrument. Dit instrument dient dan wel goede meetkwaliteiten te bezitten. In feite hebben we niet veel meer nodig dan de schalen uit het beknopte model (zie paragraaf 7.5) en NA, om voor het grootste deel de variantie op het boven-individuele niveau te verklaren in "spanning" en "welbevinden".

Op grond van deze resultaten wil ik hier dan ook pleiten voor praktijkonderzoek naar PSA en werkstress met een beperkt aantal goede vragenlijstschalen, in plaats van met een lange reeks van enkelvoudige items, zoals bijvoorbeeld gebruikelijk is in het klassieke Periodiek Bedrijfsgezondheidkundig Onderzoek (PBGGO). Methoden met veel losse items kunnen op een drietal manieren de plank mis slaan.

Ten eerste: een groot deel van de resultaten is onderling redundant. Dit leidt tot niet-efficiënte nabesprekingen, waarin telkens in andere termen dezelfde punten worden herhaald. Het telkens herhalen van negatieve punten in andere bewoordingen, komt bij bedrijven en instellingen over als een koude douche en werkt frustrerend. Hierdoor wordt een goede follow-up belemmerd.

Ten tweede: het vermogen om positieve punten te signaleren is beperkt, waardoor de discussie bij nabesprekingen naar de negatieve kant trekt. Opdrachtgevers willen ook positieve punten terughoren, en willen beloond worden voor geleverde inspanningen om werkstressproblemen aan te pakken. Ook dit belemmert een goede follow-up.

Ten derde: het signaleren van vele, afzonderlijke punten wekt ten onrechte de indruk dat deze punten onafhankelijk van elkaar kunnen worden opgelost en aangepakt. Dit leidt bij de vertaling van onderzoeksresultaten naar actiepunten tot allerlei problemen. Het is verstandiger om vanaf het begin te zoeken naar een beperkt aantal kernproblemen, en deze in onderlinge relatie tot elkaar op te lossen c.q. bij te stellen.

Aan deze inhoudelijke bezwaren tegen het werken met losse items kunnen ook nog methodologische bezwaren worden toegevoegd (Evers, 1995).

Het is verleidelijk om het resultaat van de verklaring van een groot deel van de boven-individuele variantie met weinig variabelen door te trekken naar het gebied van interventie.

Dit zou betekenen dat met ingrepen in een zeer beperkt aantal kernaspecten (werkdruk, werkinhoud, werksfeer), nagenoeg alles bereikt kan worden op het gebied van reductie van "spanning" en verbetering van het "welbevinden" in het werk, dat bereikbaar is op een hoger niveau dan het niveau van de individuele werknemer.

Negatieve affectiviteit

Voor de praktijk is uit dit proefschrift een belangrijk advies af te leiden waar het gaat om NA. Wanneer vragenlijstonderzoek wordt gedaan naar de relatie van spanningsverschijnselen, psychische klachten of psychosomatische gezondheidsklachten met factoren in de arbeid of de arbeidssituatie, is het van groot belang een maat voor NA toe te voegen (Brief et al., 1988). Op deze manier kunnen relaties tussen stressoren en klachten worden onderzocht, terwijl wordt gecorrigeerd voor NA. Geflatteerde correlaties worden hierdoor voorkomen, en relaties tussen diverse stressoren en uitkomstvariabelen worden hierdoor beter onderling vergelijkbaar.

Op grond van dit onderzoek kan worden voorspeld dat in ons land veel gebruikte instrumenten als

- de Vragenlijst Over de Ervaren Gezondheidstoestand (VOEG: Dirken, 1967);
- de psychische klachtenschaal uit de Vragenlijst OrganisatieStress-D (VOS-D: Bergers et al., 1986);
- de gezondheidsvragen uit de basismodule voor Periodiek ArbeidsGezondheidskundig Onderzoek (basismodule PAGO: Weel, Broersen en Rozema, 1993);
- de emotionele uitputtingsschaal van de Maslach Burnout Inventory (MBI-NL: Schaufeli en Van Dierendonck, 1994);

uitermate gevoelig zijn voor vertekening door NA. Op individueel niveau is het zelfs de vraag, of het wel de specifieke gezondheidsklachten van een werknemer zijn, die worden gemeten met de afzonderlijke items in dit soort vragenlijsten (vaak moe zijn, hoofdpijn, rugklachten et cetera), dan wel dat hiermee voor een belangrijk deel slechts iemands mate van NA wordt gemeten. Bij personen met een hoge mate van NA, kan getwijfeld worden aan de specificiteit van eventuele signalen over moeheid, maagpijn, spierpijn, hoofdpijn et cetera. Voorzichtigheid bij de begeleiding en behandeling is dan ook geboden in dergelijke gevallen.

Voor vragenlijstonderzoek naar onderwerpen op het gebied van "arbeidsgerelateerd welbevinden" (zoals arbeidssatisfactie, arbeidsmotivatie, commitment, verlooptentatie en dergelijke), is het minder van belang om rekening te houden met vertekening door NA, wanneer we relaties willen onderzoeken met factoren in de arbeid of de arbeidssituatie.

Een laatste, maar belangrijk punt is, dat dit onderzoek laat zien dat vragenlijstonderzoek naar PSA en werkstress zinvol is, ondanks dat er - zoals critici suggereren - inderdaad confounding optreedt. De mate van confounding door NA is niet zo groot, dat elke vorm van samenhang tussen PSA en werkstress hierdoor wegverklaard kan worden. PSA levert een eigen bijdrage aan de verklaring van de variantie in "spanning" en "welbevinden", bovenop de door NA verklaarde variantie.

Algemene onvrede

Op grond van de bevindingen in dit proefschrift zou ik iedereen die onderzoek doet op het gebied van PSA en werkstress willen adviseren om nooit meer interpretaties van resultaten te maken c.q. van opdrachtgevers te accepteren in termen van algemene onvrede. Algemene onvrede is een doodoener, die een echte oplossing van problemen in de weg staat.

Bij onderzoek naar PSA en werkstress kunnen de problemen van individuen, afdelingen, en organisaties het beste worden begrepen in termen van de zes dimensies, die zijn gevonden in hoofdstuk 7. Wellicht zijn aanvullingen nodig op gebieden als "fysieke belasting", "arbeidsvoorwaarden" en "gezondheidsklachten" et cetera. Ook richting maatregelen is het van belang in termen van dit begrippenkader te denken. Het denken in termen van differentiatie in typen oorzaken en differentiatie in typen gevolgen is een voorwaarde om te komen tot een gedifferentieerde aanpak van werkstress. Dit proefschrift laat zien dat zelfbeoordelingsgegevens geschikt zijn om een dergelijke differentiatie aan te brengen. Dit pleit tegen de scepsis die velen a-priori tegenover zelfbeoordelingen lijken te hebben als onderzoeks- en beleidsinstrument.

Ziekteverzuim

PSA en werkstress, zoals gemeten met vragenlijstschalen, kunnen individueel ziekteverzuim in beperkte mate voorspellen. Vooral "spanning" en "veelzijdigheid" voorspellen verzuim. EO is in dit onderzoek naar voren gekomen als een voorspeller van de individuele verzuimduur, maar niet van de individuele verzuimfrequentie. Aanpak van PSA en werkstress heeft dus mogelijk resultaat in termen van minder verzuim. Dit is een stimulans voor de (beginnende) praktijken van arbeid- en organisatiedeskundigheid en psychologie van arbeid en gezondheid binnen arbodiensten.

Toch is een relativiserende opmerking hier op zijn plek. Uit eerder onderzoek was reeds bekend dat ziekteverzuim een multicausaal fenomeen is. Als maximaal haalbare verklaarde variantie op basis van één type verklaringen, geldt bij ziekteverzuim 5 à 10% van de variantie. PSA en werkstress blijkt aan deze verwachting te voldoen: 3% à 8% van de variantie is verklaarbaar gebleken in dit onderzoek. Toch zijn deze percentages laag. Het is dan ook van belang om enerzijds te benadrukken dat aanpak van PSA en werkstress resultaat kan opleveren in termen van verzuimreductie, en anderzijds tegelijkertijd te waarschuwen tegen te hoge verwachtingen.

Ziekteverzuim vraagt om een aanpak op meerdere fronten tegelijkertijd: wetgeving, arbeid en organisatie, arbeidsomstandigheden, individuele problematiek enzovoort. Een aanpak van PSA en werkstress verdient hierin een normale plek, naast aandacht voor andere factoren.

Een laatste kanttekening aangaande de voorspelbaarheid van verzuim is, dat een deel van het ziekteverzuim onvermijdelijk is en tevens ongerelateerd aan factoren in de arbeid en in de arbeidssituatie (griepgolven, sportblessures). Als dit deel van het ziekteverzuim zou worden weggelaten, blijft het reëel door het bedrijf beïnvloedbare verzuim over. Als de analyses in hoofdstuk 9 beperkt zouden zijn tot dit beïnvloedbare verzuim, dan is het aannemelijk dat de verklaarde variantie ten gevolge van PSA en werkstress hoger zou zijn uitgevallen. Verder is te verwachten dat per organisatie de hoeveelheid van het beïnvloedbare verzuim, die verband houdt met PSA en werkstress, zal variëren. Het feit dat 3% à 8% van het totale verzuim wordt verklaard bij één specifieke organisatie, wil dus niet zeggen dat in andere organisaties door PSA en werkstress niet veel meer variantie kan worden verklaard in het beïnvloedbare verzuim.

Non-lineaire effecten

In hoofdstuk 8 zijn aanwijzingen gevonden dat voor "psychologische taakeisen", "veelzijdigheid" en "regelmogelijkheden" een optimaal niveau bestaat. Zowel bij te veel als te weinig van deze aspecten treden werkstressverschijnselen op. Wat als te veel en te weinig geldt, is hierbij afhankelijk van de stabiele kenmerken van de individuele werknemer, de variabele toestand van de individuele werknemer, en de verandering van deze werknemer over de tijd. Voor de praktijk heeft deze bevinding twee consequenties.

In de eerste plaats spreekt dit resultaat het bestaan van standaardoplossingen tegen. Veel projecten op het gebied van functieverbetering en organisatieverandering, waarbij werkstressaanpak mede een doel is, zijn gericht op verlaging van "psychologische taakeisen", meer "veelzijdigheid", en meer "regelmogelijkheden". De tegenovergestelde bewegingen kunnen echter ook nodig zijn!

In de tweede plaats is individuele normstelling lastig, terwijl het op populatieniveau wel degelijk mogelijk is om tot normstelling te komen. Dit wordt geïllustreerd aan de hand van de relatie tussen "psychologische taakeisen" en "spanning". Het moet mogelijk zijn om, door scores van diverse populaties onderling te vergelijken, een score te vinden op de dimensie "psychologische taakeisen", waarboven het risico op "spanningsverschijnselen" en de bijbehorende gezondheidsklachten op populatieniveau onevenredig toeneemt. Zo'n schaalscore kan een handvat bieden bij normstelling aangaande "psychologische taakeisen" op meso- en macro-niveau.

Tegelijkertijd toont dit onderzoek aan dat het vergelijken van een individuele medewerker met deze zelfde norm weinig zinvol is: de individuele variabiliteit is zo groot, dat een score boven de norm voor de ene werknemer overbelasting betekent, en voor de andere werknemer onderbelasting. Op micro-niveau hebben we derhalve niet zoveel aan een norm op het gebied van "psychologische taakeisen" om risico-personen aan te wijzen. Hier zullen juist de effect-maten ("spanning" en "welbevinden") moeten zorgen voor signalering van personen, die risico lopen ten aanzien van overspannenheid, burnout en verlies van arbeidsmotivatie.

Arbeidsgerelateerde spanningsverschijnselen

Voor de praktijk van advisering over werkstress is het resultaat van dit onderzoek van groot belang. Op het gebied van "arbeidsgerelateerde spanningsverschijnselen" blijkt de verklaarbare variantie vooral te zitten op het individuele niveau, NA speelt een grote rol, en "psychologische taakeisen" blijken de belangrijkste determinant te zijn.

"Spanning" dient vooral te worden begrepen als een individugebonden proces. De invloed van de arbeidsbelasting en de arbeidssituatie zelf, is relatief beperkt. Het verwerkingsvermogen en de eigen wijze van interpreteren van en omgaan met "psychologische taakeisen" zijn van essentieel belang. Daarnaast speelt het hebben van "regelmogelijkheden" een rol.

De resultaten van dit onderzoek zijn aanleiding tot enige scepsis ten aanzien van managementsystemen, die proberen om de taaklast op een zo objectief mogelijke manier gelijk te verdelen over medewerkers in een bepaalde groep. Van dergelijke systemen mag geen volledig doeltreffend effect worden verwacht in termen van spanningsreductie. Het risico bestaat dat dergelijke - goed bedoelde - systemen tot op zekere hoogte spanningsklachten veroorzaken, doordat met het individuele verwerkingsvermogen onvoldoende rekening wordt gehouden.

Voor de duidelijkheid: ik wil hier niet betogen dat het gelijk verdelen van de taaklast een slechte zaak is. Integendeel: dit lijkt mij een uitstekende zaak en een absolute voorwaarde om de arbeidsbelasting van een groep werknemers af te kunnen stemmen op de capaciteit van deze groep. In mijn ogen begint hier echter pas het echte probleem, waar het gaat om het ontstaan van "spanning". Hoe kunnen we - gegeven een genormeerde taaklast - zorgen, dat dit bij individuele werknemers niet leidt tot dusdanige (in-)spanningsverschijnselen, dat dit op lange termijn overspannenheid en gezondheidsproblemen zal veroorzaken.

In preventieve zin betekent dit enerzijds, dat aan individuele medewerkers "regelmogelijkheden" geboden moeten worden om de taaklast aan hun verwerkingsvermogen aan te kunnen passen. Anderzijds is het tegelijkertijd van belang werknemers duidelijk te maken hoe hun eigen manier van interpreteren van en omgaan met "psychologische taakeisen" bijdraagt aan het ontstaan van en in stand houden van "spanningsverschijnselen".

Voor interventie betekent dit, dat het signaleren van individuele "spanningsverschijnselen" door werknemers zelf, door hun directe werkomgeving, maar ook door arbodiensten, een belangrijke bijdrage kan leveren aan het voorkomen van erger. Tenslotte houdt dit in dat bij werknemers die reeds in een fase zijn van lange termijn-gevolgen (overspannenheid, ernstige psychische en/of psychosomatische aandoeningen) bij de reïntegratie opnieuw rekening moet worden gehouden met de individuele manier van interpreteren en omgaan met "psychologische taakeisen". Een tijdelijke aanpassing, puur in termen van taaklast alleen, zal niet altijd het gewenste effect opleveren.

Gelukkig zijn er aanzetten in de literatuur over psychologie van arbeid en gezondheid (Winnubst en Schrabracq, 1992; Schrabracq en Winnubst, 1993) en klinische arbeids- en organisatiepsychologie (De Moor, 1989; 1992) te vinden, die kunnen helpen om dit type dienstverlening bij arbodiensten op te bouwen. Dit proefschrift draagt niet zozeer bij aan kennis over wat dergelijke dienstverlening zou moeten inhouden, maar onderbouwt vooral het belang van het feit dat dit type dienstverlening er komt.

Arbodiensten kunnen een belangrijke bijdrage leveren aan het terugdringen en beperkt houden van werkstressproblematiek bij de Nederlandse beroepsbevolking, door het uitvoeren van periodiek arbeidsgezondheidskundig onderzoek (PAGO) op het gebied van werkstress. De in dit proefschrift ontwikkelde vragenlijst is hiervoor een geschikt middel.

Bij een dergelijk "PAGO werkstress" is aandacht nodig voor de individuele beleving en beoordeling van PSA en werkstress, waarbij in ieder geval de zes kernaspecten van de hier gevonden structuur aan bod moeten komen. Ook hiermee verband houdende persoonskenmerken en copingstijlen zouden in dit kader kunnen worden onderzocht. De bedrijfsarts, psycholoog en bedrijfsmaatschappelijk werker van de arbodienst kunnen een belangrijke rol vervullen door (in het kader van zo'n "PAGO werkstress") een risico-situatie en/of risico-persoon te signaleren, en dit bespreekbaar te maken binnen het bedrijf en/of bij de betrokken werknemer(s).

Arbeidsgerelateerd welbevinden

Plezier in het werk en je betrokken voelen bij het werk, de afdeling en de organisatie: dit is wat met "arbeidsgerelateerd welbevinden" wordt bedoeld. Van dit "welbevinden" is aangetoond, dat een relatief groot deel ervan "gedeeld" is. Dit betekent dat er een bepaalde mate van overeenstemming over is tussen werknemers binnen afdelingen binnen organisaties. De determinanten van "welbevinden" beslaan nagenoeg alle PSA-dimensies. Alleen voor "regelmogelijkheden" kon geen verband worden aangetoond.

Voor de praktijk suggereert dit onderzoek dat een collectieve aanpak (naast een individuele aanpak) zin heeft op het gebied van "welbevinden in het werk". Het is mogelijk het "welbevinden" van individuele en groepen werknemers te beïnvloeden, door het beïnvloeden van hun "gedeelde" mening over arbeid en organisatie. Groepsgewijze sessies in het kader van functieverbeteringen, veranderingen in de productie of dienstverlening, en organisatieveranderingen lijken een zinvol instrument te kunnen zijn. Op voorhand zijn hierbij weinig PSA-factoren uit te sluiten als zijnde onbelangrijk.

Het groeps- en organisatieniveau biedt een krachtig middel voor de beïnvloeding van individueel "welbevinden". De keerzijde van de medaille is, dat het ook collectief fout kan gaan. Als in een bepaalde afdeling bij een aantal individuele medewerkers een negatieve houding ontstaat tegenover de werkinhoud, als men hierdoor het plezier verliest in het werk, en de betrokkenheid van deze medewerkers bij de afdeling en de organisatie afneemt, dan steekt men hier anderen mee aan. Een negatieve spiraal kan het gevolg zijn.

In dit onderzoek is qua bedrijfsmatige gevolgen uitsluitend gekeken naar verzuim. In verband met de hier gepresenteerde redenering over de praktijkconsequenties van "onwelbevinden" lijkt het van belang nader onderzoek te doen naar de relatie tussen PSA, "arbeidsgerelateerd welbevinden" en productiviteit.

Als aangetoond kan worden dat afdelingen die problematisch scoren qua "welbevinden" minder productief zijn of minder flexibel zijn bij aanpassingen in de organisatie en/of productie, dan is hiermee een belangrijk bedrijfsmatig motief gegeven om aan PSA en werkstress iets te doen.

Bij arbodiensten is de afgelopen decennia een traditie ontstaan om bij het periodiek bedrijfsgezondheidskundig onderzoek (PBGO) te werken met groeps- en bedrijfsrapportages over vragenlijstgegevens (Jenniskens, 1982; Weel en Broersen, 1992). Juist voor het periodiek signaleren van de hier bedoelde problemen in het "welbevinden" van afdelingen, beroepsgroepen en organisaties als geheel, blijft dit een goed instrument.

Het is de hoop van de auteur dat deze traditie binnen arbodiensten niet teloor gaat, nu de praktijk zich - onder druk van de nieuwe wetgeving - stort op risico-inventarisatie en -evaluatie (RI&E) en periodiek arbeidsgezondheidskundig onderzoek (PAGO). Het zou beter zijn als de bestaande expertise en faciliteiten ten behoeve van groepsgewijs vragenlijstonderzoek werden gebruikt om nieuwe vormen van dienstverlening (zoals RI&E en PAGO) gestalte te geven. Daarnaast is er ook de hoop dat arbodiensten niet zullen blijven steken bij de signalering van problemen op het gebied van PSA en werkstress. Met goede instrumenten voor nader onderzoek (niet alleen zelfbeoordelingsinstrumenten, maar ook andere!), kan ook de stap worden gezet naar problemdiagnose, en van daaruit naar een probleemaanpak samen met het bedrijf en de betrokken medewerkers.

10.7 Slot deel III

PSA en werkstress zijn thema's, die aan het eind van de twintigste eeuw hoog op de agenda staan van diegenen, die verantwoordelijk zijn of werkzaam zijn voor de gezondheid van werkende mensen. Het belang van deze problematiek neemt de afgelopen decennia duidelijk toe.

De gevolgen van PSA en werkstress zijn niet alleen voelbaar voor werknemers zelf. Voor de afdelingen waar deze mensen werken betekent een verminderde prestatie bij één van de afdelingsleden vaak een verzwarende voor de overige afdelingsleden. Voor organisaties zijn de resulterende kosten merkbaar, niet in de laatste plaats doordat de kwaliteit en de kwantiteit van de productie c.q. dienstverlening vermindert. Ook de maatschappij als geheel ervaart de tol van PSA en werkstress, doordat zij de grote hoeveelheid personen, die in de huidige race niet (meer) mee kunnen doen, moet onderhouden.

Velen vragen zich af hoe het mogelijk is, dat PSA en werkstress zulke belangrijke thema's zijn, terwijl de moderne arbeid en arbeidssituatie in zoveel opzichten verbeterd lijken te zijn ten opzichte van vroeger.

Dit proefschrift levert op twee manieren een bijdrage aan dit maatschappelijk beladen onderwerp. In de eerste plaats wordt een verzameling van zelfbeoordelingsschalen aangereikt, die onderzoek op dit terrein mogelijk maakt. In de tweede plaats is, gebruik makend van de zelf ontwikkelde schalen, onderzocht welke dimensies op dit gebied zijn te onderscheiden, hoe deze dimensies zich onderling verhouden, en hoe deze dimensies zich verhouden tot het object waarover ze verondersteld worden informatie te geven: werknemers binnen afdelingen binnen organisaties.

Twee mechanismen kunnen worden onderscheiden op het gebied van werkstress: "spanning" en "welbevinden". Vooral het mechanisme van "spanning" en de daarbij behorende PSA-factoren zijn van belang in verband met de gezondheid van werknemers. Vooral deze factoren leiden tot de eerder beschreven negatieve gevolgen op het niveau van afdelingen, organisaties en de maatschappij als geheel. "Arbeidsgerelateerde spanningsverschijnselen" zijn - op deze manier bekeken - de achilleshiel van het huidige systeem van arbeid en organisatie. Hoe zou dit komen?

Uit dit onderzoek leren we dat vooral "psychologische taakeisen" en "regelmogelijkheden" verband houden met "spanning". Wat is op deze gebieden veranderd binnen organisaties in de afgelopen decennia? Belangrijke veranderingen op dit gebied zijn: de verandering in het soort arbeid (meer mentaal, minder fysiek), het verhogen van de productiviteit (uit concurrentie-oogpunt noodzakelijk), en het vergroten van de mate van interne controle (management-informatiesystemen, kwaliteitsnormen).

Echter niet alleen organisaties zijn in de afgelopen decennia veranderd. Ook werknemers zijn tegelijkertijd veranderd. Het gemiddelde opleidingsniveau is toegenomen, met alle bijkomende veranderingen in verwachtingen naar arbeid en arbeidssituatie toe. Werknemers hebben meer privé-zaken gekregen waaraan zij tijd en energie willen besteden, naast tijd en energie die zij willen besteden aan werken. Ook bij werknemers is de mate van controle toegenomen, binnen en buiten het werk: velen leven op het werk, maar ook in de privé-tijd met een agenda. Het wegvallen van veel zorgen over het tegemoet komen aan de primaire levensbehoeften heeft velen mogelijkheden gebracht om keuzes te maken aangaande allerlei manieren van het besteden van de verworven vrije tijd.

Werknemers en organisaties staan hierdoor tegenover elkaar. Voor organisaties blijkt de gemiddelde werknemer geen altijd flexibele, hard werkend, intelligente, energieke, (cliënt-)vriendelijke automaat te zijn. Voor werknemers blijkt de gemiddelde baan/organisatie niet de altijd flexibele, comfortabele, stimulerende, steunende, respectvolle situatie te zijn, die zij verwachten. Toch hebben werknemers en organisaties elkaar elke dag nodig.

Het lijkt alsof organisaties en werknemers anno 1996 strijden om de prijs, besteding en controle van hetzelfde goed: menselijke inspanning en activiteit. Veel organisaties zouden, bij gelijke kwaliteit en kwantiteit van de prestatie, mensen verruilen voor machines, ook als dit meer kost. Veel werknemers zouden, bij gelijke beloning en sociale zekerheid, werken verruilen voor niet werken, ook al verliezen ze hierdoor contacten en vaardigheden. In de onderhandelingen over de prijs, besteding en controle van menselijke inspanning en activiteit zijn de onderhandelingen hard. Dit is niet wezenlijk anders dan vroeger. Ook toen ging het om een ruil tussen organisaties en werknemers: geld voor inspanning. Waar vroeger de mate van fysieke inspanning en de hoeveelheid ongevallen aantoonde wat de zware beroepen waren, doen dit nu de cijfers over werkdruk, burnout en overspannenheid.

Literatuur

- Adams, J.S. (1965). Inequity in social exchange. In: Berkowitz, L. (Ed.), *Advances in experimental social psychology* (Vol. 2). New York: Academic Press.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50, 179-211.
- Algera, J.A. (1981). *Kenmerken van werk*. Lisse: Swets & Zeitlinger.
- Algera, J.A. (1983). Objective and perceived task characteristics as a determinant of reactions by task performers. *Journal of occupational psychology*, 56, 95-107.
- Algera, J.A. (1989). Taakkenmerken. In: Drenth, P.J.D., Thierry, Hk., Wolff, Ch. J. de (Red.), *Nieuw handboek arbeids- en organisatiepsychologie*. Deventer: Van Loghum Slaterus.
- Beehr, T.A., Newman, J.E. (1978). Job stress, employee health, and organizational effectiveness: a facet analysis, model, and literature review. *Personnel psychology*, 31, 665-699.
- Bentler, P.M., Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, 88, 588-606.
- Berg, S.A. van den, Evers, G.E., Smid, G.A.C., Vaas, S. (1993). *A&O-deskundige: taken, kwalificaties, overgangsbeleid, opleiding*. Amsterdam/Leiden: NIA/NIPG-TNO.
- Bergers, G.P.A., Marcelissen, F.H.G, Wolff, Ch. de (1986). *Vragenlijst organisatie stress-D. Handleiding*. (intern rapport psychologie van arbeid en organisatie), Katholieke Universiteit Nijmegen.
- Biessen, P.G.A. (1992). *Oog voor de menselijke factor: achtergrond, constructie en validering van de basisvragenlijst Amsterdam*. Amsterdam: Swets & Zeitlinger.
- Bijl, R., Bauduin, D. (1991). *Categorie V: arbeidsongeschikt wegens psychische stoornissen*. Utrecht: Nederlands Centrum Geestelijke Volksgezondheid.
- Bosker, R.J., Scheerens, J. (1989). Issues in the interpretation of the results of school-effectiveness research. *International journal of educational research*, 13, 741-751.
- Bosker, R.J., Snijders, T.A.B. (1990). Statistische aspecten van multi-niveau onderzoek. *Tijdschrift voor onderwijsresearch*, 15, 317-329.
- Brief, A.P., Burke, M.J., Robinson, B.S., Webster, J. (1988). Should negative affectivity remain an unmeasured variabele in the study of job stress? *Journal of applied psychology*, 73, 193-198.
- Broersen, J.P.J., Weel, A.N.H., Dijk, F.J.H. van (1989). *Periodiek bedrijfsgezondheidkundig onderzoek: middel en maatstaf*. Doetinchem/Amsterdam: BGD Oost-Gelderland/Coronel Laboratorium/Studiecentrum Arbeid en Gezondheid.
- Broersen, J.P.J., Vrijhof, B., Weel, A.N.H. (1990). *PBGO en ziekteverzuim: de voorspelbaarheid van individueel ziekteverzuim op grond van het PBGO*. Amsterdam: NIA.
- Brown, G.W. (1989). Life events and measurement. In: Brown, G.W., Harris, T.O. (Eds.), *Life events and illness*. London: The Guilford Press.
- Bryk, A.S., Raudenbusch, S.W. (1992). *Hierarchical linear models*. Newbury Park: Sage.
- Burger, G.C.E. (1959). De betekenis van kwantitatieve meting en functionele beoordeling van arbeidsbelasting en belastbaarheid voor de praktische bedrijfsarts. *TSG*, 37, 377-384.

- Buunk, A.P., Wolff, Ch.J. de (1988). Sociaal-psychologische aspecten van stress op het werk. In: Drenth, P.J.D., Thierry, Hk., Wolff, Ch.J. de (Red.), Nieuw handboek arbeids- en organisatiepsychologie. Deventer: Van Loghum Slaterus.
- Buunk, A.P., Vries, N.K. de (1991). Sociale vergelijking en sociale cognitie. In: Vries, N.K. de, Pligt, J. van der (Red.), Cognitieve sociale psychologie. Amsterdam: Boon.
- Buunk, A.P., Schaufeli, W.B. (1993). Burnout: a perspective from social comparison theory. In: Schaufeli, W.B., Maslach, C., Marek, T. (Eds.), Professional burnout: recent developments in theory and research. Washington: Taylor and Francis.
- Cannon, W.B. (1932). The wisdom of the body. New York: Norton.
- Caplan, R., Cobb, S., French, J., Harrison, R. van, Pinneau, S. (1975). Job demands and the worker. Survey Research Center, Institute for Social Research, University of Michigan.
- Chadwick-Jones, J.K., Nicholson, N., Brown, C. (1982). Social psychology of absenteeism. New York: Praeger.
- Chen, P.Y., Spector, P.E. (1991). Negative affectivity as the underlying cause of correlations between stressors and strains. *Journal of applied psychology*, 76, 398-407.
- Christis, J. (1988). Psychische arbeidsbelasting. *Gedrag en organisatie*, 1, 55-71.
- Christis, J. (1991). Sociotechniek en de arbo-wet. Geestelijke gezondheid op de werkplek. In: POST-groep, Onderweg naar nieuwe fabrieken en kantoren. Deventer: Kluwer.
- Cooper, C.L., Payne, R.L. (1991). Personality and stress: individual differences in the stress process. Chichester: Wiley.
- Dawis, R.V., Lofquist, L.H. (1984). A psychological theory of work adjustment. Minneapolis: University of Minnesota Press.
- Debets, P., Brouwer, E. (1989). MSP: a program for Mokken scale analysis for polychotomous items. Groningen: iec ProGAMMA.
- DeFrank, R.S. (1988). Psychometric measurement of occupational stress: current concerns and future directions. In: Hurrell, J.J., Murphy, L.R., Sauter, S.L. (Eds.), Occupational stress: issues and developments in research. New York: Taylor and Francis.
- Dewe, Ph. (1991). Measuring work stressors: the role of frequency, duration, and demand. *Work and stress*, 5, 77-91.
- Dhondt, S. (1993). D4: Interbeoordelaarsbetrouwbaarheid WEBA-methodiek. Leiden: NIPG-TNO.
- Diekstra, R.F.W., Heus, P. de, Schouten, M.H., Houtman, I.L.D. (1994). Werken onder druk. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Dierendonck, D. van, Schaufeli, W.B. (1993). Burnout en organisatiebetrokkenheid bij leerlingverpleegkundigen. De rol van billijkheid, sociale steun en probleemhantering. *TSG*, 71, 339-344.
- Dijk, F.J.H. van, Dormolen, M. van, Kompier, M.A.J., Meijman, T.F. (1990). Herwaardering model belasting-belastbaarheid, *TSG*, 68, 3-10.
- Dijkhuizen, N. van (1980). From stressors to strains. Research into their interrelationships (proefschrift Katholieke Universiteit Nijmegen). Lisse: Swets & Zeitlinger.
- Dijkstra, A., Grinten, M.P. van der, Schlatmann, M.J.Th., Winter, C.R. de (1981). Funktioneren in de arbeidssituatie. Leiden: NIPG/TNO.
- Dirken, J.M. (1967). Het meten van stress in industriële arbeidssituaties. Leiden: NIPG/TNO.
- Dohrenwend, B.S., Dohrenwend, B.P. (1974). Stressful life events: their nature and effects. New York: Wiley.

- Drenth, P.J.D., Sijtsma, K. (1990). *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten/Antwerpen: Bohn, Stafleu & Van Loghum.
- Ekkers, C.L., Sanders, A.F. (1987). *Stress in de arbeidssituatie*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Elliot, G.R., Eisdorfer, C. (1982). *Stress and human health*. New York: Springer.
- Elliott, T.R., Chartrand, J.M., Harkins, S.W. (1994). Negative affectivity, emotional distress, and the cognitive appraisal of occupational stress. *Journal of vocational behavior*, 45, 185-201.
- Ettema, J.H. (1973). Het model belasting en belastbaarheid. *TSG*, 51, 44-54.
- Evers, A. (1995). *Meetinstrumenten voor arbeidsomstandigheden, stress en welzijn*. Lisse: Swets & Zeitlinger.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Fletcher, B.C., Jones, F. (1993). A refutation of Karasek's demand-discretion model of occupational stress with a range of dependent measures. *Journal of organizational behavior*, 14, 319-330.
- French, J.R.P., Jr., Caplan, R.D., Harrison, R. van (1982). *The mechanisms of job stress and strain*. New York: Wiley.
- Frese, M. (1985). Stress at work and psychosomatic complaints: a causal interpretation. *Journal of applied psychology*, 70, 314-328.
- Frese, M., Zapf, D. (1988). Methodological issues in the study of work stress: objective vs. subjective measurement of work stress and the question of longitudinal studies. In: Cooper, C.L., Payne, R. (Eds.), *Causes, coping and consequences of stress at work*. Chichester: Wiley.
- Fried, Y., Ferris, G.R. (1986). The dimensionality of job characteristics: some neglected issues. *Journal of applied psychology*, 71, 419-426.
- Fried, Y., Ferris, G.R. (1987). The validity of the job characteristics model: a review and meta-analysis. *Personnel psychology*, 40, 287-322.
- Furda, J., Meijman, T. F. (1992). Druk en dreiging: sturing of stress. In: Winnubst, J., Schabracq, M. (red.), *Handboek arbeid- en gezondheidspsychologie*. Utrecht: LEMMA.
- Furda, J. (1995). *Werk, persoon en welzijn: een toetsing van het Job Demand-Control Model* (proefschrift Universiteit Utrecht). Enschede: Copyprint 2000.
- Ganster, D.C., Fusilier, M.R. (1989). Control in the workplace. In: Cooper, C.L., Robertson, I. (Eds.), *International review of industrial and organizational psychology*. Chichester: Wiley.
- Geurts, S.A., Buunk, A.P., Schaufeli, W.B. (1991). Ziekteverzuim vanuit een sociaal vergelijkingsperspectief. *Gedrag en organisatie*, 4, 385-398.
- Geurts, S.A. (1994). *Absenteeism from a social psychological perspective* (proefschrift Katholieke Universiteit Nijmegen). Nijmegen: Quickprint.
- Geurts, S.A. (1995). Verzuim: een pleidooi voor theorie-gestuurd onderzoek. *De psycholoog*, 30, 363-368.
- Gorp, K. van, Schaufeli, W.B., Hopstaken, L. (1993). Burnout en cognitief terugtrekgedrag vanuit sociaal uitwisselingsperspectief. *Gedrag en gezondheid*, 21, 274-285.
- Grosfeld, J.A.M. (1988). *De voorspelbaarheid van individuele verzuimduur*. Amsterdam: Swets & Zeitlinger.
- Grumbkov, J. von (1980). *Sociale vergelijking van salarissen*. Tilburg: Van Spaendonck.
- Gründemann, R.W.M., Smulders, P.G.W., Winter, C.R. de (1993). *Vragenlijst arbeid en gezondheid*. Handleiding. Lisse: Swets & Zeitlinger.

- Gutek, B., Repetti, R., Silver, D. (1988). Nonwork roles and stress at work. In: Cooper, C, Payne, R. (eds.), *Causes, coping and consequences of stress at work*. Chichester: Wiley.
- Haak, A.T. (1994). Dutch sociotechnical design in practice: an empirical study of the concept of the whole task group (proefschrift). Rijksuniversiteit Groningen.
- Hacker, W. (1973). *Allgemeine Arbeits- und Ingenieurspsychologie. Psychische Struktur und Regulation von Arbeitstätigkeiten*. Berlin: Deutscher Verlag der Wissenschaften.
- Hacker, W. (1986). *Arbeitspsychologie. Psychische Regulation von Arbeitstätigkeiten*. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Hackman, J.R., Oldham, G.R. (1974). The job diagnostic survey. Yale University.
- Hackman, J.R., Oldham, G.R. (1975). Development of the job diagnostic survey. *Journal of applied psychology*, 60, 159-170.
- Hackman, J.R., Oldham, G.R. (1980). *Work redesign*. Reading/London: Addison-Wesley.
- Hermkens, P., (1995). Sociale vergelijking en rechtvaardigheid van beloning: enkele bevindingen uit de literatuur. *Gedrag en organisatie*, 8, 359-371.
- Heus, P. de, Diekstra, R.F.W. (1992). *Stress aan het werk in Nederland: een onderzoek onder vijftienduizend werknemers*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Hill, J.M.M., Trist, E.L. (1962). *Industrial accidents, sickness and other absences*. London: Tavistock Publications.
- Hobfoll, S.E. (1988). *The ecology of stress*. Washington: Hemisphere.
- Hobfoll, S.E., Stokes, J.P. (1988). The process and mechanics of social support. In: Duck, S., Hay, D.F., Hobfoll, S.E., Ickes, W., Montgomery, B.M. (Eds.), *Handbook of personal relationships: theory, research, and interventions*. London: Wiley.
- Hobfoll, S.E. (1989). Conservation of resources. A new attempt at conceptualizing stress. *American psychologist*, 44, 513-524.
- Hobfoll, S.E., Freedy, J. (1993). Conservation of resources: a general stress theory applied to burnout. In: Schaufeli, W.B., Maslach, C., Marek, T. (Eds.), *Professional burnout: recent developments in theory and research*. Washington: Taylor and Francis.
- Hogervorst, P. (1992). *Welbevinden in het werk en de relatie met werkdruk en taakautonomie (scriptie psychologie van arbeid en gezondheid)*. Universiteit van Amsterdam.
- Holmes, T.H., Rahe, R.H. (1967). The social readjustment rating scale. *Journal of psychosomatic research*, 11, 213-218.
- Hopstaken, L.E.M. (1994). *Willens en wetens. Ziekmelden als beredeneerd gedrag (proefschrift)*. Rijksuniversiteit Groningen.
- House, J.S., Strecher, V., Metzner, H.L., Robbins, C.A. (1986). Occupational stress and health among men and women in the Tecumseh community health study. *Journal of health and social behavior*, 27, 62-77.
- House, J.S. (1987). Chronic stress and chronic disease in life and work: conceptual and methodological issues. *Work and stress*, 1, 129-134.
- Houtman, I.L.D., Bloemhoff, A., Kompier, M.A.J., Marcelissen, F.H.G. (1991). *Werkstressrisico's in bedrijf en beroep*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Houtman, I.L.D., Goudswaard, A., Dhondt, S., Grinten, M. P. van der, Hildebrandt, V.H., Kompier, M.A.J. (1995). *Evaluatie van de monitorstudie naar stress en lichamelijke belasting*. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.

- James, L.R., Tetrick, L.E. (1986). Confirmatory analytic tests of three causal models relating job perceptions to job satisfaction. *Journal of applied psychology*, 71, 77-82.
- Janssen, P.P.M., Buunk, A.P. (1990). De invloed van relatieve deprivatie op interne en externe verloopintentie in verschillende fasen van de loopbaan. *Gedrag en organisatie*, 3, 126-137.
- Jenniskens, P.A.H. (1982). *Modelleren van bedrijfsgezondheidszorg*. Breda: BGD West-Brabant.
- Johnson, J.V., Hall, E.M. (1988). Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the swedish working population. *American journal of public health*, 78, 1336-1342.
- Jong, R.D. (1987). *Sociale ondersteuning, spanning en stemming: een onderzoek naar functioneren en gezondheid bij ondernemers en managers* (proefschrift). Rijksuniversiteit Utrecht.
- Jonge, J. de, Schaufeli, W.B., Furda, J. (1995). Werkkenmerken: psychologische arbeidsvitamines? *Gedrag en organisatie*, 8, 231-248.
- Jonge, J. de (1995). *Job autonomy, well-being, and health: a study among dutch health care workers* (proefschrift Rijksuniversiteit Limburg). Maastricht: Datawyse.
- Jonge, J. de, Furda, J. (1995). Het Job Demand-Control Model. *De psycholoog*, 30, 496-501.
- Jöreskog, K.G., Sörbom, D. (1993). *LISREL VIII: user's reference guide*. Chicago: Scientific Software International.
- Kalimo, R., El-Batawi, M.A., Cooper, C.L. (1987). *Psychosocial factors at work and their relation to health*. Geneva: World Health Organisation.
- Karasek, R.A. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative science quarterly*, 24, 285-308.
- Karasek, R.A. (1985). *Job Content Instrument: questionnaire and user's guide*. Columbia University.
- Karasek, R.A., Theorell, T. (1990). *Healthy work: stress, productivity, and the reconstruction of working life*. New York: Basic Books.
- Kasl, S.V. (1978). Epidemiological contributions to the study of work stress. In: Cooper, C.L., Payne, R. (Eds.), *Stress at work*. New York: Wiley.
- Kasl, S.V. (1986). Stress and health in the workplace: a methodological commentary on the accumulated evidence. in: Cataldo, M., Coates, T. (Eds.), *Health promotion in industry: a behavioral medicine perspective*. New York: Wiley.
- Katz, D., Kahn, R.L. (1978). *The social psychology of organizations*. New York: Wiley.
- Klink, J.J.L. van der (1993). *Psychische problemen en de werksituatie*, Amsterdam: NIA.
- Knepper, S. (1991). *De epidemie van psychische arbeidsongeschiktheid: feiten en achtergronden*. *Gedrag en gezondheid*, 19, 315-319.
- Kompier, M.A.J. (1988). *Arbeid en gezondheid van stadsbuschauffeurs* (proefschrift Rijksuniversiteit Groningen). Eburon: Delft.
- Kompier, M.A.J., Marcelissen, F.H.G. (1990). *Handboek werkstress: een systematische aanpak voor de bedrijfspraktijk*. Amsterdam/Leiden: NIA/NIPG-TNO.
- Kompier, M.A.J. (1991). *Arbeid, stress en gezondheid*. *De psycholoog*, 26, 161-167.
- Kompier, M.A.J., Pot, F.D. (1993). Diagnostiek van stress en welzijn: benaderingen en instrumenten. In: Schabracq, M.J., Winnubst, J.A.M. (Red.), *Handboek arbeid en gezondheidpsychologie*. Deel 2: toepassingen, Utrecht: Lemma.

- König-Zahn, C., Furer, J.W., Tax, B. (1993-1995). Het meten van de gezondheidstoestand. Deel 1: algemene gezondheid; Deel 2: lichamelijke gezondheid, sociale gezondheid; Deel 3: psychische gezondheid. Assen: Van Gorcum.
- Kuipers, H., Amelsvoort, P. van (1990). Slagvaardig organiseren. Inleiding in de sociotechniek als integrale ontwerpleer. Deventer: Kluwer.
- Lazarus, R.S. (1966). Psychological stress and the coping process. New York: McGraw-Hill.
- Lazarus, R.S., Folkman, S. (1984). Stress, appraisal and coping. New York: Springer.
- Lazarus, R.S., DeLongis, A., Folkman, S., Gruen, R. (1985). Stress and adaptational outcomes: the problem of confounded measures. *American psychologist*, 40, 770-779.
- Lazarus, R.S. (1990). Theory-based stress measurement. *Psychological inquiry*, 1, 3-13.
- Leitner, K. (1993). Auswirkungen von Arbeitsbedingungen auf die psychosoziale Gesundheit. *Zeitschrift für Arbeitswissenschaft*, 47 (19 NF)/2, 98-107.
- Leventhal, H., Mosbach, P. (1983). Perceptual-motor theory. In: Caccioppo, J.T., Petty, R.E. (Eds.), *Social psychophysiology*. New York: Guilford.
- Leventhal, H., Nerenz, D.R., Steele, D. (1984). Disease representations and coping with health threats. In: Baum, A., Taylor, S.E., Singer, J. (Eds.), *Handbook of psychology and health* (Vol. 4). Hillsdale: Erlbaum.
- Leventhal, H., Tomarken, A. (1987). Stress and illness: perspectives from health psychology. In: Kasl, S.V., Cooper, C.L. (Eds.), *Stress and health: issues in research methodology*. New York: Wiley.
- Leventhal, H. (1992). Illness cognition: the study of the representation and coping with current and future health threats over the life span (Pieter de la Court-lezing). Rijksuniversiteit Leiden.
- Longford, N.T. (1993). VARCL: software for variance component analysis of data with nested random effects (maximum likelihood). Groningen: iec ProGAMMA.
- Maes, S., Veldhoven, M. van (1990). From health behaviour to health behaviour change, in: Kaptein, A., Ploeg, H. van der, Garssen, B., Schreurs, P., Beunderman, R. (Eds.), *Behavioral medicine: psychological treatment of somatic disorders*. Chichester: Wiley.
- Marcelissen, F.H.G. (1987). Gangmakers in het stressproces: de rol van type-A gedrag en sociale ondersteuning bij het stressproces in de werksituatie (proefschrift Katholieke Universiteit Nijmegen). Leiden: NIPG/TNO.
- Marsh, H.W., Balla, J.R., McDonald, R.P. (1988). Goodness-of fit indexes in confirmatory factor analysis: the effects of sample size. *Psychological bulletin*, 103, 391-410.
- McCrae, R.R. (1990). Controlling neuroticism in the measurement of stress. *Stress medicine*, 6, 237-241.
- Meijman, T.F., O'Hanlon, J.F. (1984). Arbeidsbelasting. Een inleidend overzicht van psychologische theorieën en meetmethoden. In: Drenth, P.J.D., Thierry, Hk., Willems, P., Wolff, Ch.J. de (Red.), *Handboek voor arbeids- en organisatiepsychologie*. Deventer: Van Loghum.
- Meijman, T.F., Vries-Griever, A. de, Vries, G.M. de, Kampman, R. (1985). The construction and evaluation of a onedimensional scale measuring subjective sleep quality. Rijksuniversiteit Groningen.
- Meijman, T.F. (1989). Belasting en herstel: een begrippenkader voor het arbeidspsychologisch onderzoek van werkbelasting. In: Meijman (Ed.), *Mentale belasting en werkdress: een arbeidspsychologische benadering*. Assen/Maastricht: Van Gorcum.

- Meijman, T.F. (1991). Over vermoeidheid (proefschrift Rijksuniversiteit Groningen). Amsterdam: Studiecentrum Arbeid en Gezondheid.
- Meijman, T.F., Mulder, G. (1992). Arbeidspsychologische aspecten van werkbelasting. In: Drenth, P.J.D., Thierry, Hk., Wolff, Ch.J. de (Red.), Nieuw handboek arbeids- en organisatiepsychologie. Deventer: Van Loghum Slaterus.
- Meulenbeld, C., Lingen, P. van (1993). Instrumenten op het terrein van arbeid en gezondheid. Amsterdam/ Leiden: NIA/NIPG-TNO.
- Mokken, R.J. (1971). A theory and procedure of scale analysis. The Hague: Mouton.
- Mokken, R.J., Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied psychological measurement*, 6: 417-430.
- Molenaar, I.W. (1982). Mokken scaling revisited. *Kwantitatieve methoden*, 3, 8, 145-164.
- Moor, W. de (1989). Stress- en conflictmanagement. Deventer: Van Loghum Slaterus.
- Moor, W. de (1992). Een constructivistische visie op de organisatie en enkele implicaties van deze visie. *M&O*, 458-470.
- Nicholson, N., Johns, G. (1985). The absence culture and the psychological contract: who's in control of absence? *Academy of management review*, 10, 397-407.
- Oort-Marburger, D., Schmidt, S.H., Berg, B.M.C. van den, Berg, D.M.Th. van den, Meijman, T.F., Oel, C.J. van (1993). Dimensies in het AVP. Amsterdam: Studiecentrum Arbeid en Gezondheid.
- Ouwerkerk, R.J. van, Meijman, T.F., Mulder, G. (1994). Arbeidspsychologische taakanalyse. Utrecht: Lemma.
- Parkes, K.R. (1990). Coping, negative affectivity, and the work environment: additive and interactive predictors of mental health. *Journal of applied psychology*, 75, 399-409.
- Parkes, K.R., Mendham, C.A., Rabenau, C. von (1994). Social support, and the demand-discretion model of job stress: test of additive and interactive effects in two samples. *Journal of vocational behavior*, 44, 91-113.
- Peeters, M.C.W., Meijer, S. (1995). Betrokkenheid bij de organisatie, de afdeling en het werk: een zinvol onderscheid? *Gedrag en organisatie*, 8, 153-165.
- Pennebaker, J.W. (1982). The psychology of physical symptoms. New York: Springer.
- Pennebaker, J.W. (1983). Physical symptoms and sensations: psychological causes and correlates. In: Cacioppo, J.T., Petty, R.E. (Eds.), *Social psychophysiology*. New York: Guilford.
- Pennebaker, J.W. (1984). Accuracy of symptom perception. In: Baum, A., Taylor, S.E., Singer, J. (Eds.), *Handbook of psychology and health* (Vol. 4). Hillsdale: Erlbaum.
- Ploeg, H.M. van der, Defares, P.B., Spielberger, C.D. (1980). Handleiding bij de Zelfbeoordelingsvragenlijst. Lisse: Swets & Zeitlinger.
- Pot, F., Christis, J., Fruytier, B., Kommers, H., Middendorp, J., Peeters, M., Vaas, S. (1989). Functieverbetering en organisatie van de arbeid. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.
- Reiche, H.M.J.K.I. (1982). Stress aan het werk. Over de effecten van de persoonlijkheid en sociale ondersteuning op strains (proefschrift Katholieke Universiteit Nijmegen). Lisse: Swets & Zeitlinger.
- Roberts, K.H., Glick, W.H. (1981). The job characteristics approach to task design: a critical review. *Journal of applied psychology*, 66, 193-217.

- Roe, R., Zijlstra, F. (1991). Arbeidsanalyse ten behoeve van (her)-ontwerp van functies: een handelingstheoretische invalshoek. In: Algera, A. (ed.), *Analyse van arbeid vanuit verschillende perspectieven*. Amsterdam: Swets & Zeitlinger.
- Salancik, G.R., Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative science quarterly*, 23, 224-253.
- Sarason, I.G. (1975). Test anxiety, attention, and the general problem of anxiety. In: Spielberger, C.D., Sarason, I.G. (Eds.), *Stress and anxiety* (Vol. 1). Washington: Hemisphere.
- Schabracq, M.J. (1992). Arbeid en gezondheidpsychologie en haar toekomst. In: Winnubst, J.A.M., Schabracq, M.J. (Red.), *Handboek Arbeid- en Gezondheidpsychologie*. Utrecht: Lemma.
- Schabracq, M.J., Winnubst, J.A.M. (1993). *Handboek arbeid- en gezondheidpsychologie. Deel 2: Toepassingen*. Utrecht: LEMMA.
- Schachter, S. (1959). *The psychology of affiliation*. Palo Alto: Stanford University Press.
- Schalk, M.J.D. (1989). Determinanten van veelvuldig kortdurend verzuim. Den Haag: Delwel.
- Schalk, M.J.D., Keunen, A., Meijer, T. (1995). Warr's schalen voor welzijn en mentale gezondheid: factorstructuur en betrouwbaarheid. *Gedrag en organisatie*, 8, 116-127.
- Schaufeli, W.B., Dierendonck, D. van (1993). The construct validity of two burnout measures. *Journal of organizational behavior*, 14, 631-647.
- Schaufeli, W.B., Dierendonck, D. van (1994). Burnout, een begrip gemeten. De Nederlandse versie van de Maslach Burnout Inventory (MBI-NL). *Gedrag en gezondheid*, 22, 153-172.
- Schaufeli, W.B., Eijnden, R.J.J.M. van den, Brouwers, H.M.G. (1994). Stress en burnout bij penitentiaire inrichtingswerkers. De rol van sociaal-cognitieve factoren. *Gedrag en organisatie*, 7, 216-224.
- Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal: Acta.
- Semmer, N. (1992). One man's meat, another man's poison: stressors and their cultural background. In: Cranach, M. von, Doise, W., Mugny, G. (Eds.), *Social representations and the social bases of knowledge*. Lewiston: Hogrefe & Huber.
- Semmer, N., Zapf, D., Greif, S. (1995). Shared well-being on the job: a new approach for assessing the validity of job stress measurements. *Journal of occupational and organizational psychology*, (in druk).
- Semmer, N. (1995). Individual differences, work stress and health. In: Schabracq, M.J., Winnubst, J.A.M., Cooper, C.L. (Eds.), *Handbook of work and health psychology*. Chichester: Wiley, (in druk).
- Sevastos, P., Smith, L., Cordery, J.L. (1992). Evidence on the reliability and construct validity of Warr's (1990) well-being and mental health measures. *Journal of occupational and organizational psychology*, 65, 33-49.
- Sijsma, K., Debets, P., Molenaar, I.W. (1990). Mokken scale analysis for polychotomous items: theory, a computer program, and an empirical application. *Quality and quantity*, 24, 173-188.
- Sitter, L.U. de (1981). *Op weg naar nieuwe fabrieken en kantoren*. Deventer: Kluwer.
- Sitter, L.U. de (1990). Gezond samenwerken in productiesystemen. *Gedrag en gezondheid*, 17, 149-154.
- Smulders, P.G.W., Bloemhoff, A. (1991). *Arbeid, gezondheid en welzijn in de toekomst*. Houten/Antwerpen: Bohn, Stafleu & van Loghum.

- Smulders, P.G.W., Winter, C.R. de (1993). De voorspellende waarde van arbeidssituatie en gezondheid voor ziekteverzuim en arbeidsongeschiktheid. In: Schabracq, M.J., Winnubst, J.A.M. (Red.), *Handboek arbeid en gezondheidspsychologie. Deel 2: toepassingen*. Utrecht: LEMMA.
- Snijders, T.A.B., Bosker, R.J. (1994). Modeled variance in two-level models. *Sociological methods and research*, 22, 342-363.
- Spector, P.E. (1987). Interactive effects of perceived control and job stressors on affective reactions and health outcomes for clerical workers. *Work and stress*, 1, 155-162.
- Spector, P.E., Dwyer, D.J., Jex, S.M. (1988). Relation of job stressors to affective, health and performance outcomes: a comparison of multiple data sources. *Journal of applied psychology*, 73, 11-19.
- Spector, P.E., Jex, S.M. (1991). Relations of job characteristics from multiple data sources with employee affect, absence, turnover intentions, and health. *Journal of applied psychology*, 76, 46-53.
- Spector, P.E. (1992). A consideration of the validity and meaning of self-report measures of job conditions. *International review of industrial and organizational psychology*, 7, 123-151.
- Spielberger, C.D. (1966). Theory and research in anxiety. In: Spielberger, C.D. (Ed.), *Anxiety and behavior*. New York: Academic Press.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R.E. (1970). *Manual for the state-trait anxiety inventory*. Palo Alto: Consulting Psychologists Press.
- Syroit, J.E.M.M. (1984). *Interpersonal injustice: a psychological analysis illustrated with empirical results (dissertatie)*. Tilburg: Katholieke Universiteit Brabant.
- Taylor, S.E. (1986). *Health Psychology*. New York: Random House.
- Thierry, Hk. (1969). *Arbeidsinstelling en prestatiebeloning*. Utrecht: Het Spectrum.
- Thierry, Hk. (1990). Beloning van de arbeid. In: Drenth, P.J.D., Thierry, Hk., Wolff, Ch.J. de (Red.), *Nieuw handboek arbeids- en organisatiepsychologie*. Deventer: Van Loghum Slaterus.
- Trist, E.L., Bamforth, K.W. (1951). Some social and psychological consequences of the long-wall method of coal-getting. *Human relations*, 4, 3-38.
- Tucker, L.R., Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.
- Turner, A.N., Lawrence, P.R. (1965). *Industrial jobs and the worker. An investigation of response to task attributes*. Boston: Harvard Graduate School of Business Administration.
- Veldhoven, M. van (1993a). PBGO-factor module psychosociale arbeidsbelasting: Tussenrapport 1, uitgangspunten en constructie van een itempool. Amsterdam: NIA.
- Veldhoven, M. van (1993b). PBGO-factor module psychosociale arbeidsbelasting: Tussenrapport 2, schaalconstructie-onderzoek en ontwikkeling van de vragenlijst beleving en beoordeling van de arbeid (VBBA). Amsterdam: NIA.
- Veldhoven, M. van, Meijman, T.F. (1994). *Het meten van psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van de arbeid (VBBA)*. Amsterdam: NIA.
- Verbaan, D. (1991). *Verlaag het ziekteverzuim: een praktische handleiding*. Groningen: Wolters-Noordhoff.

- Vliet, K. van, Gijbers van Wijk, C. (1990). Selectieve aandacht voor lichamelijke sensaties. Symptoomperceptie: een cognitief-psychologische benadering van somatisatie. *De psycholoog*, 25, 205-212.
- Volpert, W., Oesterreich, R., Jablenz-Kolakovic, S., Krogoll, T., Resch, P. (1983). Verfahren zur Ermittlung von Regulationsanfordernissen in der Arbeitstätigkeit. Köln: Verlag TUV Rheinland.
- Wall, T.D., Clegg, C.W., Jackson, P.R. (1978). An evaluation of the job characteristics model. *Journal of occupational psychology*, 51, 183-196.
- Warr, P.B. (1987). *Work, unemployment and mental health*. Oxford: Clarendon Press.
- Warr, P.B. (1990a). Decision latitude, job demands and employee well-being. *Work and stress*, 4, 285-294.
- Warr, P.B. (1990b). The measurement of well-being and other aspects of mental health. *Journal of occupational psychology*, 63, 193-219.
- Warr, P.B. (1994). A conceptual framework for the study of work and mental health. *Work and stress*, 8, 84-97.
- Watson, D., Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin*, 98, 219-235.
- Watson, D., Clark, L.A., Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54, 1063-1070.
- Watson, D., Pennebaker, J.W. (1989). Health complaints, stress, and distress: exploring the central role of negative affectivity. *Psychological review*, 96, 234-254.
- Weel A.N.H., Broersen, J.P.J. (1992). Signalen van problemen in werk en gezondheid: periodiek bedrijfsgesondheidskundig onderzoek bij groepen werkenden (proefschrift). Universiteit van Amsterdam.
- Weel, A.N.H., Broersen, J.P.J., Rozema, R. (1993). Eindverslag PBGO-basismodule, Amsterdam: SKB.
- Weel, A.N.H., Broersen, J.P.J. (1995). Verslag van het SKB-project PAGO-module werkstress. Amsterdam: SKB.
- Winnubst, J.A.M. (1986). Rolspanningen in het werk. In: Roe, R.A. (Red.), *Psychologie der arbeidsomstandigheden*. Assen/Maastricht: Van Gorcum.
- Winnubst, J.A.M., Schabracq, M.J. (1992). *Handboek arbeid- en gezondheidspsychologie*. Utrecht: LEMMA.
- Winter, C.R. de (1991). *Arbeid, gezondheid en verzuim als voorspellers van uitval uit het werk* (proefschrift Rijksuniversiteit Limburg). Leiden: NIPG-TNO.
- Ybema, J.F., Buunk, A.P. (1992). Affectieve gevolgen van sociale vergelijking bij stress. Een onderzoek onder arbeidsongeschikten. *Gedrag en gezondheid*, 20, 277-288.
- Yperen, N.W. van, Buunk, A.P., Schaufeli, W.B. (1992). Communal orientation and the burnout syndrome among nurses. *Journal of applied social psychology*, 22, 173-189.
- Yperen, N.W. van, Horn, J.E. van (1993). Stress, coping en sociale vergelijking bij arbeidsongeschikten. *Gedrag en organisatie*, 6, 293-309.
- Yperen, N.W. van, Hagedoorn, M., Geurts, S.A. (1994). Terugtrekgedrag van werknemers. Verloop en verzuim als reacties op onbillijkheid. *Gedrag en organisatie*, 7, 243-260.

- Yperen, N.W. van (1995). The pursuit of the fairly treated/adaptive worker. *De psycholoog*, 30, 97-101.
- Zapf, D. (1989). *Selbst- und Fremdbeobachtung in der psychologischen Arbeitsanalyse*. Göttingen: Hogrefe.
- Zijlstra, F., Mulder, G. (1989). Mentale belasting: theoretische gezichtspunten en overzicht van meetmethoden. In: Meijman (Ed.), *Mentale belasting en werkstress: een arbeidspsychologische benadering*. Assen/Maastricht: Van Gorcum.

Samenvatting

M. van Veldhoven. *Psychosociale arbeidsbelasting en werkstress*.

Het proefschrift gaat over psychosociale arbeidsbelasting en werkstress. Dit onderwerp wordt onderzocht met behulp van een in dit proefschrift ontwikkelde zelfbeoordelingsmethode.

In het eerste deel (hoofdstuk 2 t/m 4) worden drie verschillende invalshoeken gepresenteerd, die van belang zijn voor het vervolg: theorieën over psychosociale arbeidsbelasting, theorieën over het ontstaan van psychologische (werk-)stress, en de meettechnische voor- en nadelen van een zelfbeoordelingsmethode.

In het tweede deel (hoofdstuk 5 en 6) wordt een verzameling van zelfbeoordelingschalen geconstrueerd voor het meten van psychosociale arbeidsbelasting en werkstress.

Het derde deel (hoofdstuk 7 t/m 9) gebruikt de geconstrueerde schalen voor het beantwoorden van een aantal fundamentele vragen op het gebied van psychosociale arbeidsbelasting en werkstress: de structuur en de onderlinge relatie van deze onderwerpen, en de voorspellende waarde met betrekking tot ziekteverzuim.

Hoofdstuk 1 vormt de inleiding tot het proefschrift. Doel en belang van het onderwerp worden kort aangegeven, waarna de methodologische en theoretische vragen die in het proefschrift worden uitgediept worden geïntroduceerd. Vervolgens wordt een elementair begripkader gedefinieerd.

In **hoofdstuk 2** komen acht theorieën over psychosociale arbeidsbelasting aan bod. Achtereenvolgens wordt ingegaan op: de sociotechniek, de handelingstheorie, de taakkenmerken-benadering, het Michigan-model, het demand-control model, het vitamine-model, de belasting-belastbaarheid benadering, en het integraal model van Beehr en Newman. Bij elk model worden twee vragen gesteld die van belang zijn voor de in het proefschrift te construeren zelfbeoordelingsschalen.

De eerste vraag heeft betrekking op de opvattingen die de theorieën hebben over het voorkomen en bestrijden van werkstress. Hierbij blijkt er een tweedeling te bestaan. Enerzijds zijn er theorieën die kiezen voor een situatieve benadering. Het elimineren van risico's in de situatie is hierbij de primaire strategie voor het voorkomen en bestrijden van werkstress. Anderzijds zijn er theorieën die zich richten op de interactie tussen persoon en situatie. De strategie voor het voorkomen en bestrijden van werkstress richt zich bij deze theorieën op het bewaken of herstellen van de balans tussen de mogelijkheden van de persoon en de eisen van de situatie.

De tweede vraag heeft betrekking op wat en hoe er volgens een bepaalde theorie moet worden gemeten met betrekking tot psychosociale arbeidsbelasting. Alhoewel de te meten begrippen per theorie verschillend zijn, blijkt hierin wel degelijk een grootste gemene deler te vinden te zijn. Hiermee worden begrippen bedoeld die in nagenoeg elke theorie terugkomen, zij het vaak in andere bewoordingen.

Opvallend is dat een theoretische voorkeur voor een risicobenadering niet per definitie gepaard gaat met het afwijzen van een zelfbeoordelingsmethode. Zo maakt men in het demand-control model veelvuldig gebruik van vragenlijsten. Opvallend is ook dat een theoretische voorkeur voor een interactionistisch model niet per definitie gepaard gaat met het exclusief gebruiken van zelfbeoordelingsmethoden. Zo maakt men in de handelingstheorie veelvuldig gebruik van expert-beoordelingsmethoden en observaties, die parallel aan zelfbeoordelingsmethoden worden toegepast.

De voornaamste conclusie uit hoofdstuk 2 met betrekking tot het vervolg van het proefschrift is dan ook, dat het werken met zelfbeoordelingsmethoden niet theoriegebonden is. Vragenlijstinformatie kan nuttig zijn, zowel in een risicobenadering als in een interactionistische benadering. Er wordt daarom niet voor gekozen om de te ontwikkelen vragenlijst a priori vanuit één specifieke theorie te ontwerpen. Het tweede deel van dit proefschrift beschrijft een andere aanpak.

Hoofdstuk 2 sluit af met een bespreking van drie fundamentele methodologische verschillen tussen de besproken theorieën: het gewenste analyse-niveau, de aard van de veronderstelde verbanden tussen psychosociale arbeidsbelasting en werkstress, en de veronderstelde oorzakelijkheidsketen. Deze drie onderwerpen komen in het derde deel van dit proefschrift aan bod.

Hoofdstuk 3 plaatst het onderwerp psychosociale arbeidsbelasting en werkstress tegen de achtergrond van de algemene psychologische stresstheorie. Hiertoe wordt een beknopte historische schets gegeven van de ontwikkeling in het denken over het ontstaan van psychologische stress. Een zestal benaderingen komt achtereenvolgens aan bod: stress als respons, stress als stimulus, de individuele verschillenbenadering, het transactioneel model, het zelfregulatie-model, en het conservation of resources-model.

Wat zegt de ontwikkeling binnen de algemene psychologische stresstheorie over het terrein psychosociale arbeidsbelasting en werkstress? In de eerste plaats dat er een zwaar accent ligt op de situatie, wat gezien het toepassingsgebied te verwachten is. In de tweede plaats dat het transactionele model, uitermate dominant op de achtergrond aanwezig is bij de meeste theorieën op het gebied van psychosociale arbeidsbelasting en werkstress. De meeste theorieën verschillen niet zozeer qua onderliggende stresstheorie, maar in de mate waarin het begrip balans of homeostase expliciet in het model is opgenomen. Het vitamine-model is het enige interactionistische model dat op deze materie een ander licht werpt: hierin is geen sprake van een simpele balans tussen persoon en situatie, waarbij elke vorm van overbelasting leidt tot dezelfde werkstressgevolgen. In het vitamine-model is sprake van specifieke stressreacties ten gevolge van specifieke oorzakelijke factoren.

Zo gezien is een groot deel van de theorievorming op het gebied van psychosociale arbeidsbelasting en werkstress vrij smal, en wordt sinds de jaren zeventig gedomineerd door een transactionele visie. Er zijn twee stromingen aan te geven in de stress-literatuur die de transactionele visie intact laten, maar ieder een eigen

kleur hieraan geven. De eerste hiervan is de verzameling van theorieën die men collectief sociale vergelijkingstheorie noemt. Een aantal hiervan wordt kort besproken. Bij de billijkheidstheorie wordt iets langer stil gestaan. De toepassing hiervan op het terrein psychosociale arbeidsbelasting en werkstress wordt behandeld. Vervolgens komt de symptoomperceptie-theorie aan bod. In dit kader is het begrip negatieve affectiviteit ontwikkeld. Dit begrip heeft fundamentele gevolgen voor vragenlijstonderzoek naar psychosociale arbeidsbelasting en werkstress. Op de billijkheidstheorie en negatieve affectiviteit wordt in het derde deel van het proefschrift teruggekomen.

In **hoofdstuk 4** wordt de literatuur behandeld die ingaat op het meten van psychosociale arbeidsbelasting en werkstress in het algemeen, en op de vragenlijst-methode in het bijzonder. Voor- en nadelen van de vragenlijstmethode en alternatieven hiervoor komen aan bod. De conclusie van deze bespreking is dat een zwart-wit discussie voor of tegen het gebruik van zelfbeoordelingsmethoden niet zinvol is. Het is verstandiger te bekijken welke aspecten beter kunnen worden gemeten met een zelfbeoordelingsmethode, en welke beter met alternatieve methoden.

Vervolgens wordt de aandacht gevestigd op een belangrijke en terechte kritiek ten aanzien van het werken met een zelfbeoordelingsmethode: de zogenaamde triviality trap. Dit houdt in dat psychosociale arbeidsbelasting en werkstress op vergelijkbare manieren en in vergelijkbare bewoordingen worden bevraagd, zodat er schijn-correlaties ontstaan tussen deze twee. Met dit probleem wordt in het tweede deel van dit proefschrift bij de schaalconstructie rekening gehouden.

Het laatste deel van hoofdstuk 4 slaat een brug naar de schaalconstructie. Aan de hand van het integraal model van Beehr en Newman (1978) wordt bekeken, op welke gebieden van een zelfbeoordelingsmethode een zinvolle bijdrage mag worden verwacht en op welke gebieden niet. Het hoofdstuk sluit af met een bespreking van de praktischeisen die aan een zelfbeoordelingsmethode op het gebied van psychosociale arbeidsbelasting en werkstress kunnen worden gesteld.

Hoofdstuk 5 behandelt de achtergrond en het ontwerp van de zelfbeoordelings-schalen die in dit proefschrift worden ontwikkeld. Aansluitend bij de conclusies uit eerdere hoofdstukken, wordt als methode voor het komen tot te meten begrippen gebruik gemaakt van een facetanalyse. Uit 50 bestaande instrumenten voor het meten van psychosociale arbeidsbelasting en werkstress worden de gemeten begrippen verzameld. Dit levert als opmerkelijk resultaat dat er op de terreinen van persoonskenmerken en coping geen begrippen worden gevonden, waarover qua wenselijkheid en wijze van meten consensus bestaat. De grootste gemene deler van het huidige instrumentarium lijkt te bestaan uit begrippen die betrekking hebben op de arbeidstaak, de context waarin die arbeidstaak wordt uitgeoefend, en de mogelijke gevolgen van de arbeid voor de werknemer.

Het resultaat van de facetanalyse wordt als startmateriaal gebruikt voor het construeren van een itempool bestaande uit ruim 400 vragen, verdeeld over 27 itemverzamelingen.

Het laatste deel van hoofdstuk 5 gaat in op de keuze van te meten begrippen in de itempool, het ordenen van items in de itemverzamelingen, en de keuze van het type antwoordmogelijkheden per itemverzameling. Hierbij spelen inhoudelijke maar ook pragmatische afwegingen een rol.

Hoofdstuk 6 bespreekt het schaalconstructie-onderzoek. Het hoofdstuk opent met een bespreking van de techniek die voor de schaalconstructie wordt gebruikt: Mokken-analyse. Het schaalconstructie-onderzoek werd uitgevoerd bij twee gezondheidszorginstellingen en een productiebedrijf. Voor de schaalconstructie werd een ontwikkelgroep en een testgroep gemaakt. De schaalconstructie verliep volgens een van te voren opgesteld stappenplan, en vond in eerste instantie plaats bij de ontwikkelgroep. Na vijf analysestappen was een verzameling van schalen geconstrueerd met goede tot acceptabele meettechnische eigenschappen (schaalbaarheid, betrouwbaarheid, schaalengte, spreiding van de itemgemiddelden over de potentiële antwoordrange, acceptabele onderlinge redundantie, hoeveelheid schendingen van de assumptie van dubbele monotonie).

Vervolgens werd bij de testgroep onderzocht welke schaal kwaliteiten de ontwikkelde schalen hadden. Op grond hiervan viel een aantal schalen af. De uiteindelijke vragenlijst bevat 26 Mokken-schalen met redelijke tot goede schaal kwaliteiten en drie blokken met losse restitems. De vragenlijst kreeg de naam "Vragenlijst Beleving en Beoordeling van de Arbeid" (VBBA) mee.

Op een onderzoekspopulatie afkomstig uit elf verschillende organisaties, met een zeer heterogene samenstelling, werden de schaal kwaliteiten van 15 schalen opnieuw getest. De resultaten weken niet veel af van die uit het eerdere schaalconstructie-onderzoek. Alle schalen bleken opnieuw redelijke tot goede Mokken-schalen te zijn.

Het laatste deel van hoofdstuk 6 slaat een brug naar het derde deel van het proefschrift. Met behulp van een exploratieve factoranalyse wordt een eerste verkenning uitgevoerd van de structuur van psychosociale arbeidsbelasting en werkstress. Psychosociale arbeidsbelasting blijkt te ordenen te zijn in drie factoren: psychologische taakeisen, sociaal-organisatorische aspecten en regelmogelijkheden. Werkstress valt uiteen in twee factoren: spanning en welbevinden.

In **hoofdstuk 7** wordt dieper ingegaan op de vraag naar de structuur van psychosociale arbeidsbelasting en werkstress. Hiertoe worden twee analyses uitgevoerd: een tweede-orde analyse van de interne consistentie en een confirmatorische factoranalyse. Deze analyses worden uitgevoerd op een bestand van één gezondheidszorginstelling en drie productiebedrijven.

Uit de tweede-orde analyse van de interne consistentie blijkt dat de gevonden structuur uit de exploratieve factoranalyse in hoofdstuk 6 redelijk voldoet, maar zeker niet zonder problemen is.

Met de confirmatorische factoranalyse wordt gezocht naar een verbeterde structuur. De confirmatorische factoranalyse is bovendien gericht op een aantal additionele vragen: Is het hele gebied psychosociale arbeidsbelasting en werkstress te herleiden tot één onderliggende dimensie van algemene onvrede of heeft het zin om afzonderlijke factoren te onderscheiden? Hoeveel afzonderlijke factoren zijn in dit geval nodig?

Een laatste vraag heeft betrekking op het demand-control model. Hierin wordt het begrip decision latitude opgevoerd, dat is opgebouwd uit de onderdelen skill variety en task authority. In de confirmatorische factoranalyse wordt de vraag gesteld hoe het begrip decision latitude moet worden geoperationaliseerd, gegeven de beschikbare verzameling van Mokken-schalen. Ten behoeve van deze analyse is opnieuw een ontwikkelgroep en een testgroep geconstrueerd, middels het trekken van random steekproeven uit het beschikbare bestand.

De confirmatorische factoranalyse werd uitgevoerd middels een analyseplan van vier stappen. De analyse leverde heldere antwoorden op de gestelde vragen. De Mokken-schalen zijn niet goed te herleiden tot één onderliggende dimensie van algemene onvrede. Het heeft derhalve zin om meerdere dimensies te onderscheiden in psychosociale arbeidsbelasting en werkstress. De structuur die het beste bleek te voldoen was de volgende. Psychosociale arbeidsbelasting heeft vier dimensies: psychologische taakeisen, veelzijdigheid, regelmogelijkheden en sociale ondersteuning. De schalen die deze laatste dimensie afdekken blijken onderling nogal redundant te zijn. Werkstress heeft twee dimensies: spanning en welbevinden. De gevonden structuur van zes dimensies blijft de beste, ongeacht of deze structuur wordt beschreven door het volledige aantal van negentien beschikbare schalen, dan wel door een minimum van twaalf schalen (twee per dimensie). Wat betreft het begrip decision latitude levert de confirmatorische factoranalyse als resultaat, dat het beter is om de dimensies veelzijdigheid en regelmogelijkheden als afzonderlijk te behandelen.

De gevonden dimensiestructuur is onderzocht op generaliseerbaarheid. Afzonderlijke passing van de structuur op een groep werknemers in de gezondheidszorg versus een groep werknemers in productiebedrijven leverde nagenoeg dezelfde passingsmaten op.

Een laatste stap die in hoofdstuk 7 wordt gezet is, om in de gevonden dimensiestructuur relaties aan te brengen tussen de psychosociale arbeidsbelasting-dimensies en de werkstress-dimensies. Dit levert een redelijk goed passend model op. Psychologische taakeisen en sociale ondersteuning hangen hierin samen met spanning. Alle psychosociale arbeidsbelasting-dimensies hangen samen met welbevinden.

In **hoofdstuk 8** worden de verbanden tussen psychosociale arbeidsbelasting en werkstress nader onderzocht. Is er sprake van een algemeen effect van psychosociale arbeidsbelasting op werkstress of is er sprake van differentiatie in meerdere werkstressmechanismen, zoals wordt gepostuleerd in het vitamine-model?

Een aantal inhoudelijk relevante onderwerpen wordt tegelijkertijd aangesneden: de verstorende rol van negatieve affectiviteit, het gewenste analyse-niveau, en de rol van ervaren onbillijkheid.

Na te hebben besproken hoe de begrippen negatieve affectiviteit en ervaren onbillijkheid in het onderzoek zijn geoperationaliseerd, volgt een uiteenzetting over de gebruikte analysetechniek: multi-level analyse. De gegevens voor de analyse zijn afkomstig van een gezondheidszorginstelling. Van deze instelling is precies beschreven welke individuen werken binnen welke afdelingen binnen welke organisatie-onderdelen. De analyse probeert de individuele beleving van spanning en welbevinden te verklaren vanuit persoonsgegevens, negatieve affectiviteit, ervaren onbillijkheid en psychosociale arbeidsbelasting. Het effect van psychosociale arbeidsbelasting is hierbij onderzocht op drie analyse-niveaus tegelijkertijd: de individuele werknemer, de afdeling en het organisatie-onderdeel.

Er blijkt inderdaad sprake te zijn van een verstorend effect van negatieve affectiviteit. Dit is groter bij spanning dan bij welbevinden. Toch levert psychosociale arbeidsbelasting een significante eigen bijdrage aan de verklaring van zowel spanning als welbevinden, bovenop de verklaarde variantie door persoonskenmerken en negatieve affectiviteit. De verklaring van spanning blijkt voor een groter deel te moeten worden gezocht op het individuele verklaringsniveau dan bij welbevinden. Bij welbevinden speelt het afdelings- en organisatieniveau een belangrijkere rol. De multi-level analyse laat zien hoe de begrippen spanning en welbevinden, beide gemeten met een zelfbeoordelingsmethode, systematisch verschillen in de mate waarin ze samenhangen met de verschillende verklaringsniveaus (individu, afdeling, organisatie) en predictoren. Dit werpt een nieuw licht op de discussie over het al dan niet subjectief zijn van vragenlijstgegevens. Een laatste opmerkelijke vondst op dit gebied is, dat met het onderzochte model en het beperkte aantal variabelen dat hierin is opgenomen, nagenoeg alle variantie kon worden verklaard op het boven-individuele verklaringsniveau (afdelingen en organisatie-onderdelen).

Wat betreft de aard van de verbanden tussen psychosociale arbeidsbelasting en werkstress werden een aantal aanwijzingen gevonden die er op duiden dat er eerder sprake is van differentiatie in meerdere werkstressmechanismen, dan van één algemeen werkstressmechanisme. Dit bevestigt de uitgangspunten van het vitamine-model. Binnen dit model spelen non-lineaire verbanden tussen psychosociale arbeidsbelasting en werkstress een belangrijke rol. In het laatste deel van hoofdstuk 8 zijn deze verbanden nader onderzocht. Geheel volgens de predicties van het vitamine-model werden non-lineaire verbanden gevonden tussen psychologische taakeisen en spanning, tussen regelmogelijkheden en spanning, en tussen veelzijdigheid en welbevinden. Er werd echter tevens een non-lineair verband gevonden tussen psychologische taakeisen en welbevinden. Dit verband wordt niet voorspeld door het vitamine-model.

De voorspellende waarde van psychosociale arbeidsbelasting en werkstress voor ziekteverzuim staat centraal in **hoofdstuk 9**. Van de gevonden dimensiestructuur, negatieve affectiviteit en ervaren onbillijkheid wordt onderzocht, in hoeverre ziekteverzuim (gedurende een periode van een half jaar na het vragenlijstonderzoek) hiermee kan worden voorspeld. Dit onderzoek vindt opnieuw plaats bij een gezondheidszorginstelling.

Voor zowel de verzuimfrequentie als de verzuimduur blijkt het LISREL-model het beste te voldoen, waarin ziekteverzuim wordt voorspeld door psychosociale arbeidsbelasting, spanning, welbevinden, en ervaren onbillijkheid. Spanning en welbevinden worden in dit model voorspeld door psychosociale arbeidsbelasting en negatieve affectiviteit. Ervaren onbillijkheid blijkt een goede voorspeller van de verzuimduur, maar niet van de verzuimfrequentie. Spanning blijkt een goede voorspeller van zowel de verzuimfrequentie als de verzuimduur te zijn, evenals veelzijdigheid van het werk. Opmerkelijk is dat welbevinden een significant effect heeft op zowel de verzuimfrequentie als de verzuimduur, totdat een model wordt onderzocht dat een rechtstreeks effect van veelzijdigheid van het werk op verzuim toelaat. Hierdoor wordt het effect van welbevinden op verzuim kennelijk teniet gedaan.

In **hoofdstuk 10** worden de belangrijkste resultaten uit het derde deel van het proefschrift overzichtelijk besproken en voorzien van commentaar. Achtereenvolgens komen aan bod: het tekort schieten van een verklaring in termen van algemene onvrede, het tekort schieten van het koepelbegrip decision latitude, de verstorende rol van negatieve affectiviteit, de meerwaarde van onderzoek op meerdere verklaringsniveaus tegelijkertijd, de differentiatie in meerdere werkstresmechanismen, het bestaan van non-lineaire effecten, en de voorspelling van ziekteverzuim. Deze resultaten worden geïntegreerd tot een multi-level model voor de beleving en beoordeling van psychosociale arbeidsbelasting en werkstress. Vervolgens worden de beperkingen van het onderzoek besproken en de betekenis van de onderzoeksresultaten voor de praktijk.

Summary

M. van Veldhoven. *Psychosocial job demands and job stress*.

This thesis is about psychosocial job demands and job stress. This topic is approached with a self-report method that is developed in this thesis.

In the first part (chapter 2-4) three different perspectives are presented that are important for the following parts of the thesis: theories about psychosocial job demands, theories about the development of (job) stress, and the strengths and weaknesses of the self-report method for measuring psychosocial job demands and job stress.

In the second part (chapter 5-6) a collection of self-report scales is constructed for measuring psychosocial job demands and job stress.

The third part (chapter 7-9) uses the constructed scales for answering fundamental questions about psychosocial job demands and job stress: the structure of and relationships between these topics, and their value in predicting absenteeism.

Chapter 1 is the introduction to the thesis. Goal and importance of the topic are presented, followed by an introduction of the methodological and theoretical questions that are the main focus of the thesis. An elementary framework of concepts is defined.

In **chapter 2** eight theories about psychosocial job demands are discussed. Successively these are: the sociotechnical approach, the action-theory approach, the job characteristics model, the Michigan-model, the demand-control model, the vitamin model, the work demands-work capacity approach, and the integrated model by Beehr and Newman. For every model two questions are raised that are important for the construction of self-report scales.

The first question is about the idea that follows from the theory about the way job stress can be prevented or treated. In answer to this question it is concluded that there are basically two viewpoints. First there are theories that have a purely situational view. Eliminating the risks in the situation is their primary strategy for prevention and intervention in the area of job stress. Second there are theories that focus on the interaction between person and situation. Their strategy for prevention and intervention is directed toward maintaining or restoring the balance between the capacities of the person and the demands from the situation.

The second question for each theory was about the preference for a specific measurement-method for psychosocial job demands. Although the specific concepts to be measured differ from theory to theory, it appears that there is certainly a big amount of overlap in concepts. There are a few concepts that appear in most theories, be it in different terms. It is striking that a theoretical preference for a situational or risk approach doesn't coincide with the rejection of the self-report method. In the demand-control model for instance, questionnaires are used very frequently.

It is also striking that a theoretical preference for an interactionistic approach doesn't coincide with the exclusive use of self-report methods. For instance: in the action-theory approach expert ratings and observations are often used together with self-reports.

The most important conclusion of chapter 2 for the rest of the thesis is, that working with self-report methods is not bound to a specific theory of psychosocial job demands. Questionnaire information can be useful in both a situational risk approach and an interactionistic approach. Therefore the choice is made not to develop a new questionnaire from one theoretical perspective only. In the second part of the thesis another approach is used to determine the content of the questionnaire.

Chapter 2 closes with a discussion of three fundamental methodological differences between the eight theories: the required level of analysis, the supposed relationship between psychosocial job demands and job stress, and the supposed chain of causality. These three topics are covered in the third part of the thesis.

In **chapter 3** the topic of psychosocial job demands and job stress is placed against the background of general psychological stress theory. A short historical presentation is given about the development of a psychological view of stress. Six approaches are discussed successively: stress as response, stress as stimulus, individual differences-approaches, the transactional model, the self-regulation model, and the conservation of resources model.

What has the history of the development of a psychological stress theory to offer to the area of psychosocial job demands and job stress? First it is clear that in this particular field, there is a particularly heavy emphasis on situational factors, as could be expected. The second point is that the transactional approach is very dominant in most theories in the area of psychosocial job demands and job stress. Most theories agree on the transactional viewpoint, but there are considerable differences in how explicitly theories incorporate the notion of balance or homeostasis in the model. The vitamin model is the only interactionistic model, that has a different perspective. According to this model there is not a simple balance between person and situation, where every form of overload leads to the same reactions of job stress. In the vitamin model specific stress reactions are caused by specific causal factors.

Looking at the area of psychosocial job demands and job stress from the angle of general psychological stress theory, the perspective is actually very narrow and has been dominated by a transactional model since the 1970's. There are however two fields of scientific work that leave the transactional viewpoint intact, but do make new contributions to the field of psychosocial job demands and job stress. First of these are the theories that are collectively called social comparison theory. A selection of these are discussed, especially equity theory. The application of equity theory to psychosocial job demands and job stress is illustrated. Next is a discussion of symptom perception theory, the second field that makes new contributions to the psychosocial job demands and job stress area. Especially the

notion of negative affectivity is important. This personality characteristic has fundamental consequences for questionnaire based research in the field of psychosocial job demands and job stress. Equity theory and negative affectivity form part of the research that is presented in part three of this thesis.

Chapter 4 deals with the literature about measuring psychosocial job demands and job stress in general, and the questionnaire method in specific. Strengths and weaknesses of the questionnaire method and its alternatives are discussed. The conclusion from this chapter is, that a general acceptance or rejection of the questionnaire method is not adequate. Rather it makes more sense to consider which factors can better be measured with a self-report method, and which factors can better be measured with alternative methods.

Next attention is focused on an important criticism of research using self-report measures: the triviality trap. This means that psychosocial job demands on the one hand and job stress on the other hand, are measured in a similar way and in similar wording, resulting in spurious correlations. This problem is considered and dealt with in the second part of the thesis, during scale construction.

The last section of chapter 4 is the bridge to the scale construction. Using the integrated model by Beehr and Newman (1978), it is considered for which factors a self-report method is useful and for which factors it is not. The chapter closes with a discussion of the practical demands, that have to be met by self-report methods in the research field of psychosocial job demands and job stress.

Chapter 5 is about the background and design of the self-report scales that are developed in this thesis. Linking up with the conclusions from preceding chapters, a facet analysis is chosen as the method for determining the concepts to be measured. 50 existing instruments for measuring psychosocial job demands and job stress are used to collect concepts. A striking result from the facet analysis is, that no concepts were found in the area of personality and coping, that were shared by a sufficient amount of existing instruments. The overlap in existing instruments can be found in three major areas: the work task, the context of the work, and the possible consequences of the work.

The result of the facet analysis is used as the basic material for the construction of an itempool consisting of more than 400 questions, divided in 27 possible scales. The last part of chapter 5 deals with the choice of concepts in the itempool, the arrangement of the items in the possible scales, and the choice of answering categories for the possible scales. Theoretical and practical considerations are used to make decisions at this point.

Chapter 6 deals with the scale construction research. The chapter opens with a review of the analysis technique that is used during scale construction: Mokken-analysis. The scale construction research was conducted on a population from two health care organisations and a production firm. Before starting the analysis, a developmentgroup and a testgroup were created.

The scale construction proceeded according to a plan of analysis, and was first conducted on the developmentgroup. After five steps a collection of scales was established with reasonable or good measurement characteristics (scalability, reliability, scale length, spreading of the item means across the possible range of answering categories, redundancy, and amount of violations of the assumption of double monotonicity).

After this the measurement characteristics of the hitherto constructed scales were calculated for the testgroup. Because of the results of this analysis a few scales were dropped. The definitive questionnaire contains 26 Mokkenscales with reasonable or good scale characteristics. Furthermore the questionnaire contains three blocks of separate items that don't form scales. The questionnaire was named "Vragenlijst Beleving en Beoordeling van de Arbeid (VBBA)".

On a research population from eleven different companies and institutions, with a very heterogeneous composition in terms of age, sex and job, the scale characteristics were again tested for 15 of the constructed scales. The results were not very different from those that were found during scale construction. All scales again appeared to be reasonable or good Mokkenscales.

The last part of chapter 6 is a bridge toward the third part of the thesis. Using factoranalysis a first exploration is done of the factor structure of psychosocial job demands and job stress. Psychosocial job demands appear to be clustered in three groups: psychological demands, social-organisational aspects, and job control. Job stress has two factors: strain and well-being.

In **chapter 7** the matter of the structure of psychosocial job demands and job stress is further investigated. Two types of analyses are performed: a second-order analysis of internal consistency and a confirmatory factoranalysis. These analyses are done on data from one health care organisation and three industrial companies. The second-order analyses of internal consistency shows that the structure that was found in the explorative factoranalysis of chapter 6 is reasonably adequate, but certainly not without problems. The confirmatory factoranalysis is used to search for an improved structure. Furthermore the confirmatory factoranalysis is used to adress some additional questions: can the entire field of psychosocial job demands and job stress be reduced to a single dimension of general dissatisfaction, or are there good reasons for distinguishing several separate factors? In the latter case: how many factors are necessary?

A last additional question is about the demand-control model. The concept of decision latitude is introduced in this model as a single concept derived from two other concepts: skill variety and task authority. The confirmatory factoranalysis addresses the question how decision latitude can best be operationalised, given the set of constructed Mokkenscales. For this analysis a new developmentgroup and a new testgroup were formed, drawing random samples from the available data. The confirmatory factoranalysis followed a plan of analysis consisting of four steps. Clear answers to the research questions were found.

The Mokkenscales cannot satisfactorily be reduced to a single underlying dimension of general dissatisfaction. It is better to discriminate several factors within the field of psychosocial job demands and job stress. The following structure appeared to be the most adequate of those investigated. The area of psychosocial job demands contains four separate dimensions: psychological demands, many-sidedness, control and social support. The questionnaire scales that cover the social support dimension, appear to be rather redundant. Job stress has two dimensions: strain and well-being. This dimensional structure remains the best, whether the confirmatory factoranalysis is performed on the maximum amount of nineteen available scales or on a minimum of twelve (two scales belonging to every dimension). As to the concept of decision latitude, the confirmatory factoranalysis shows that it is better to consider many-sidedness and control as separate dimensions.

Next it was investigated whether the dimensional structure can be generalized across different types of organisations. Separate analyses of model fit in a health care population and an industrial population were performed, and showed comparable fit measures.

So far the dimensional structure that was investigated with confirmatory factoranalysis treated all dimensions as uncorrelated. The last step in chapter 7 is to calculate the model fit for a model where the dimensions of psychosocial job demands have relationships with the dimensions of job stress. This results in a model with reasonably good fit measures. In this model psychological demands and social support appear to be correlated with strain, and all dimensions of psychosocial job demands appear to be correlated with well-being.

Chapter 8 more explicitly addresses the question of relationships between psychosocial job demands and job stress. Is there a general effect of psychosocial job demands on job stress, or is there a differentiation in several job stress mechanisms, as is postulated by the vitamin model? A few relevant other questions are investigated at the same time: the amount of confounding caused by negative affectivity, the required level of analysis, and the possible role of experiences of inequity as a confounder.

After a discussion of the operationalisation of the concepts of negative affectivity and experiences of inequity, the type of data analysis in this chapter is treated: multi-level analysis. The data in this chapter come from a health care organisation. The multi-level analysis uses a complete and precise description of which individual employees work in which departments, and which departments form which organisational units. The analysis tries to explain individual strain and well-being of employees using sociodemographical and job characteristics, negative affectivity, experiences of inequity, and psychosocial job demands. The effect of psychosocial job demands is investigated on three levels of analysis at the same time: the individual worker, the department, and the organisational unit.

As expected, negative affectivity appeared to have a confounding effect. This effect is larger for strain than for well-being.

Nevertheless psychosocial job demands add significantly to the explained variance in strain and well-being, over and above the variance explained by sociodemographical characteristics and negative affectivity. The explanation of variance is bound more to the individual level of explanation for strain than for well-being. In the case of well-being, the supra-individual level of explanation (department and organisational unit) is relatively more important. The multi-level analysis shows that strain and well-being, both concepts measured by a self-report method, systematically differ in the way they correlate with levels of explanation and predictors. This puts the old discussion of the subjectiveness of self-report methods in a new perspective. A last remarkable finding on this point is, that with the research model and the relatively few variables it contains, nearly all variance in strain and well-being that exists at the supra-individual level (departments and organisational units) can be explained.

As for the relationships between psychosocial job demands and job stress: a few arguments were found in this chapter that point towards a differentiation of several job stress mechanisms, as opposed to the existence of one general job stress mechanism. This confirms one of the basic premises of the vitamin model. Another important part of this model is the existence of non-linear relationships between psychosocial job demands and job stress. In the last part of chapter 8 these non-linear relationships were investigated. Exactly as predicted by the vitamin model, non-linear relationships were found between psychological demands and strain, between control and strain, as well as between many-sidedness and well-being. However a non-linear relationship was also found between psychological demands and well-being. This relationship is not predicted by the vitamin model.

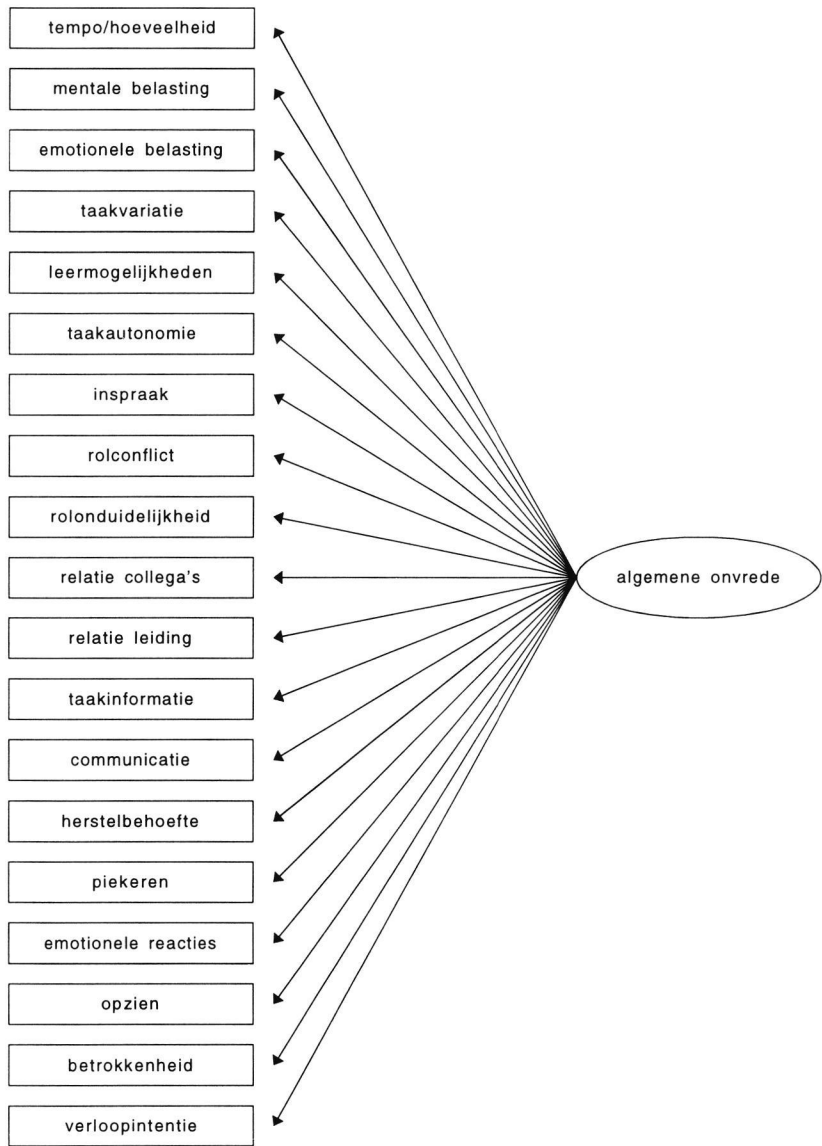
The value of psychosocial job demands and job stress in predicting absenteeism is the central topic in **chapter 9**. It is investigated to what extent absenteeism (measured during the period of one half year after the questionnaire research) can be predicted, using the dimensional structure found in chapter 7, negative affectivity, and experiences of inequity as predictors. Just like in chapter 8, this analysis is restricted to data from one health care organisation.

For both frequency- and duration-type measures of absenteeism the best fitting model is a model that predicts absenteeism using psychosocial job demands, strain, well-being, and experiences of inequity. In this model strain and well-being are predicted by psychosocial job demands and negative affectivity. Experiences of inequity appear to be a good predictor of duration, but not of frequency. Strain is found to be a good predictor of both duration and frequency, just like many-sidedness. It is remarkable that well-being predicts both duration and frequency, until a direct relationship between many-sidedness and absenteeism is incorporated in the model. The significant effect of well-being on absenteeism appears to be annihilated by this.

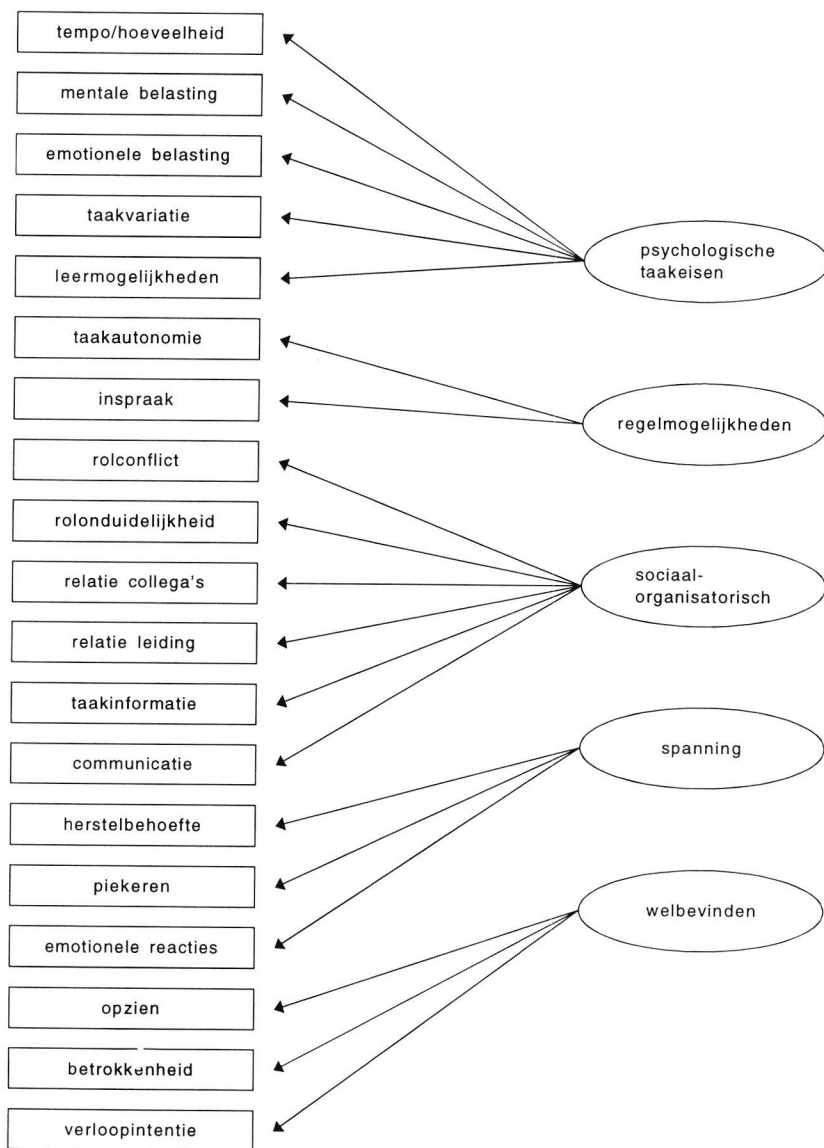
In **chapter 10** the most important results from the third part of the thesis are reviewed and discussed. One after another, the following topics are treated: the limitations of an explanation of stress in terms of general dissatisfaction, the limitations of the concept of decision latitude, the confounding effect of negative affectivity, the surplus-value of research using several levels of explanation at the same time, the existence of several job stress mechanisms, the existence of non-linear effects, and the prediction of absenteeism. These results are integrated and summarised in

a multi-level model for the experience and judgement of psychosocial job demands and job stress. Concluding this thesis, the limitations and the practical relevance of the research are discussed.

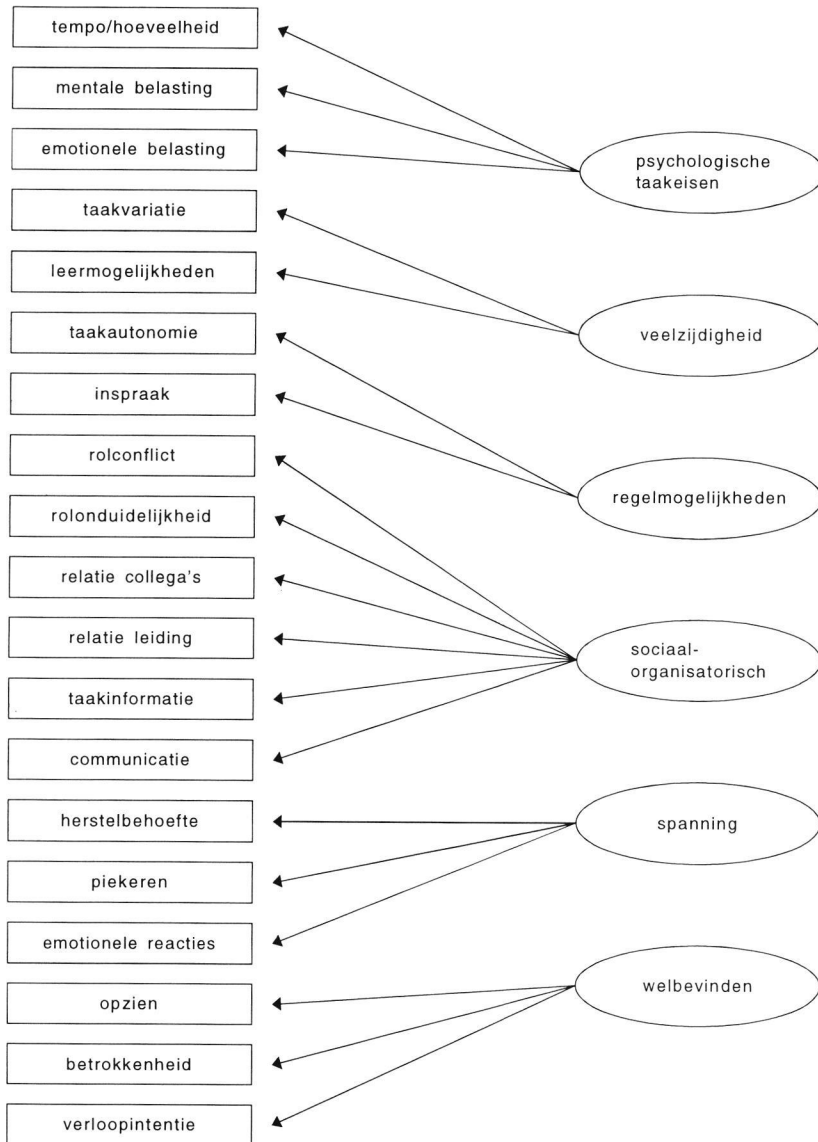
Bijlage 1: LISREL-modellen hoofdstuk 7



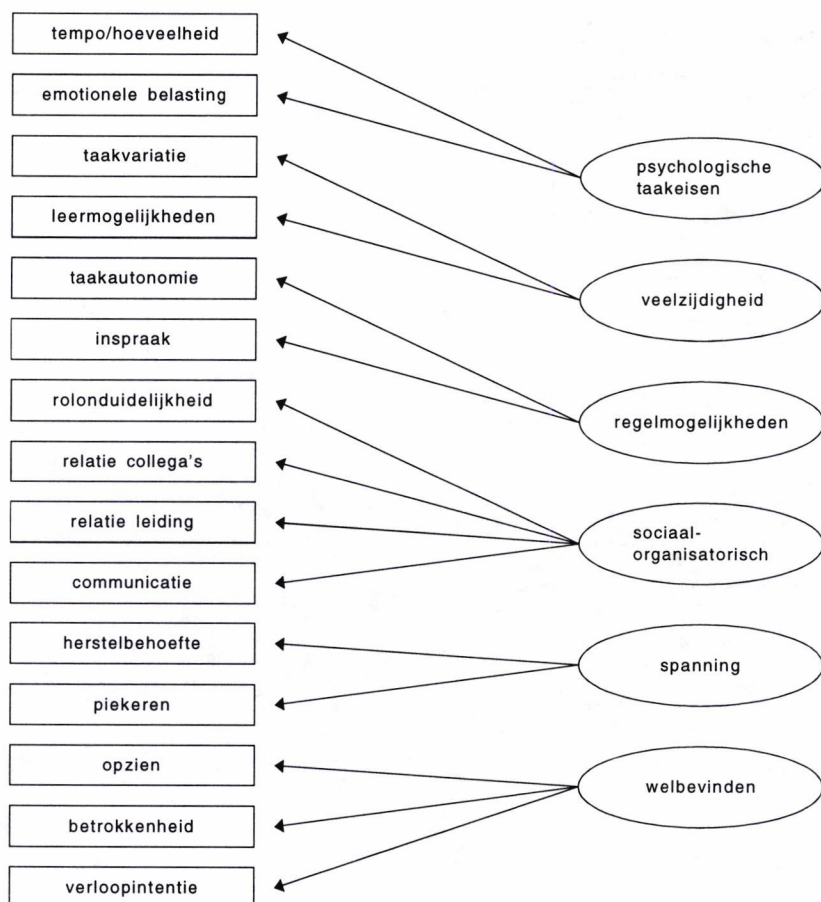
Bijlage 1.1: Confirmatorische factoranalyse - Nulmodel



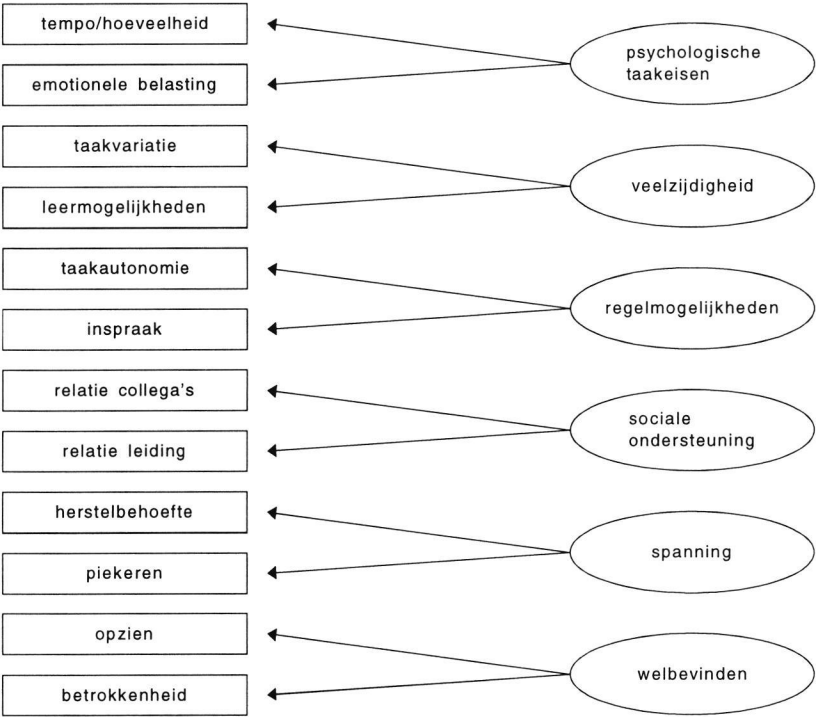
Bijlage 1.2: Confirmatorische factoranalyse - Model 1



Bijlage 1.3: Confirmatorische factoranalyse - Model 2

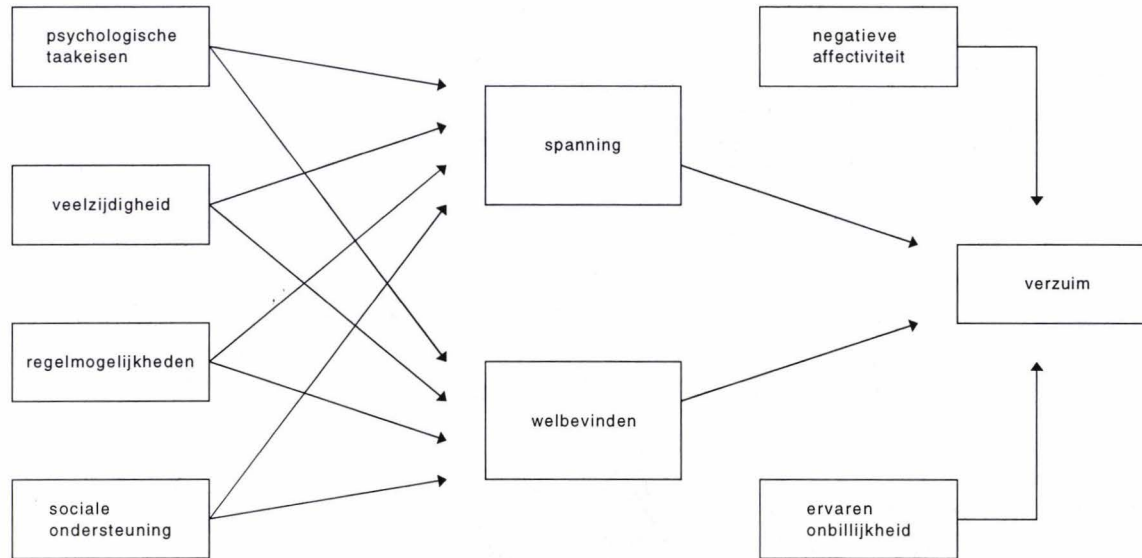


Bijlage 1.4: Confirmatorische factoranalyse - Model 5

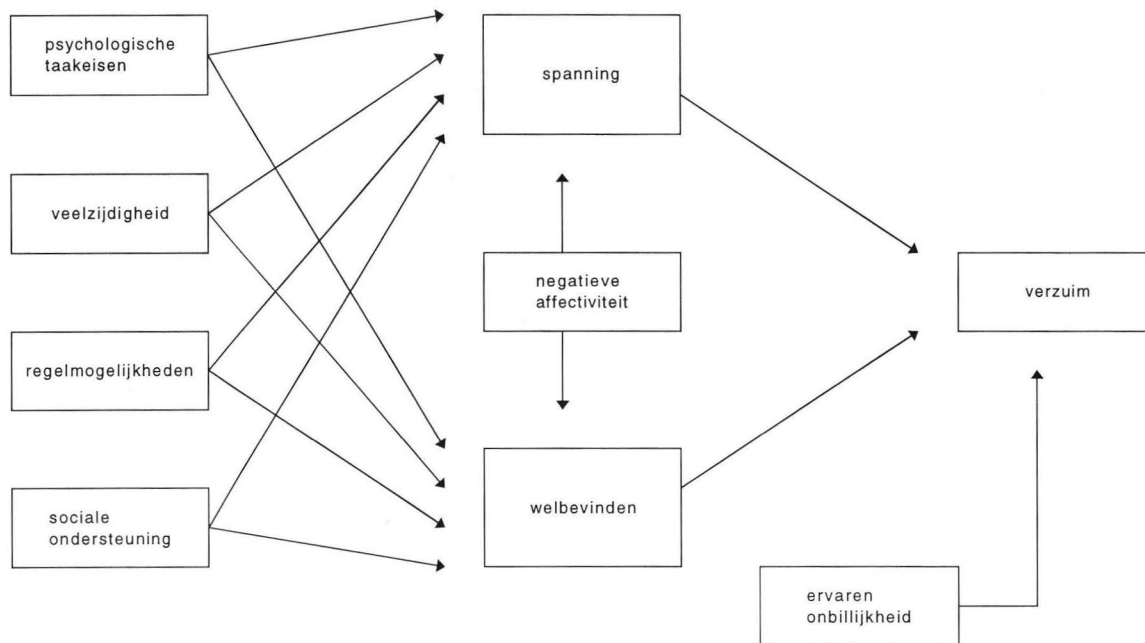


Bijlage 1.5: Confirmatorische factoranalyse - Model 6

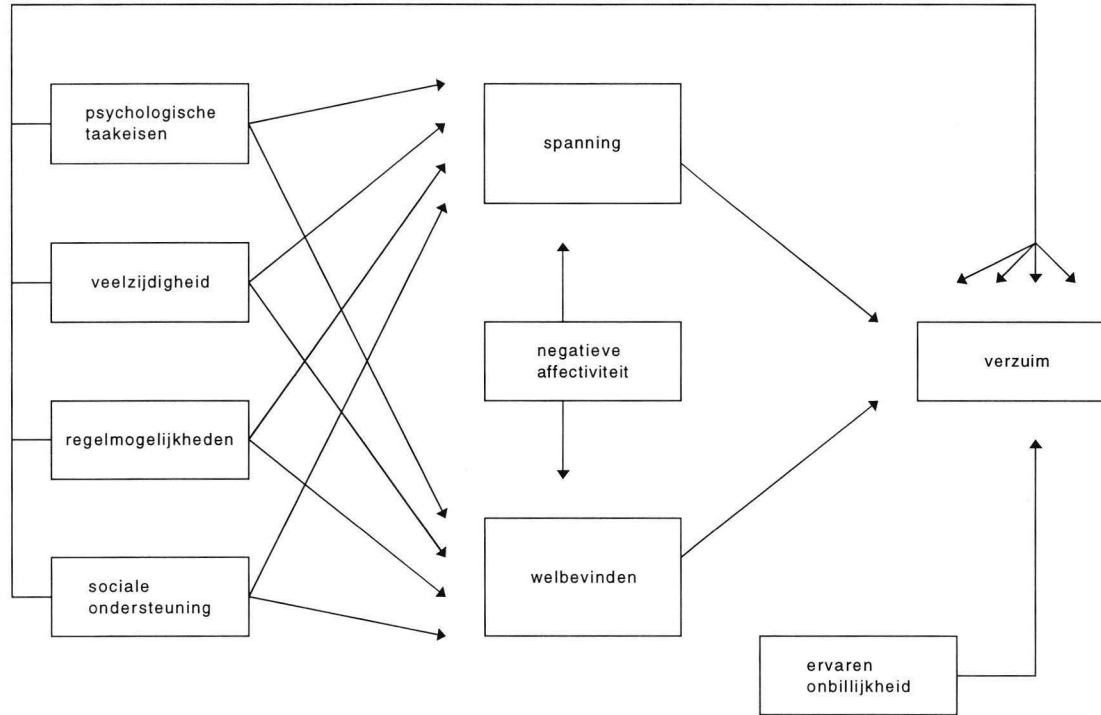
Bijlage 2: LISREL-modellen hoofdstuk 9



Bijlage 2.1: PSA, werkstress en ziekteverzuim - Model 1



Bijlage 2.2: PSA, werkstress en ziekteverzuim - Model 2



Bijlage 2.3: PSA, werkstress en ziekteverzuim - Model 3