

Kengetallen of grenswaarden voor psychische vermoeidheid in stressonderzoek. Een vergelijkende studie met de data van de 'Vlaamse werkbaarheidsmonitor'

Ongeveer een jaar geleden pakte STV innovatie & arbeid uit met de 'Vlaamse Werkbaarheidsmonitor' (Bourdeaud'hui, Janssens & Vanderhaeghe, 2004). De 'Vlaamse Werkbaarheidsmonitor' (WBM) is ondertussen uitgegroeid tot een onmiskenbare referentie inzake 'stress op het werk' in Vlaanderen. Op basis van een 'receiver operating characteristics' curve (ROC) werd een verkeerslichtenmodel voor werkstress uitgewerkt. Aan de hand van een representatieve steekproef van de Vlaamse beroepsbevolking levert dit een schatting op van de omvang van de stressproblematiek. Aan de hand van de schaal 'psychische vermoeidheid' werd geïllustreerd dat 10,5% van de werknemers eigenlijk voor het 'rood licht' staan, terwijl 18,5% op hun rem dienen te gaan staan, omdat het licht voor hen op oranje staat. Ongeveer 71% kan rustig door het groen licht rijden. In deze bijdrage willen we de conclusies van de WBM valideren door de resultaten over psychische vermoeidheid te vergelijken met de resultaten die via een andere methode werden bekomen: de latente clusteranalyse. Een latente clusteranalyse (verder: LCA) bepaalt de werkbaarheidsgrenzen op basis van de antwoordpatronen van de respondenten zelf, en niet op basis van een extern criterium zoals het geval is bij de WBM.¹

Inleiding

De onderzoekers van de werkbaarheidsmonitor stellen dat het valideren van grenswaarden ten dele een beleidsmatige kwestie is. Bij het bepalen van grenswaarden moet steeds een beleidsmatige afweging gemaakt worden tussen een voldoende ruime signalering van risicogevalen enerzijds en het vermijden van een cijfermatige overschatting van de problematiek anderzijds. Door te werken met een 'verkeerslichtenmodel' met oranje knipperlicht (ruime signalering van risicogevalen) en rood licht (strikt conservatieve inschatting van de probleemgroep) kan volgens de onderzoekers dit dilemma ondervangen worden. De databank wordt dan ook ingedeeld in twee groepen: een problematische en een niet-problematische groep. Deze indeling kwam tot stand op basis van een door de on-

derzoekers zelf bepaalde drempel: het vertonen van een bepaald aantal symptomen.

In de hier gerapporteerde analyse gaan we anders te werk. We analyseren via een latente clusteranalyse in hoeverre we de respondenten op basis van hun antwoordpatronen in homogene groepen kunnen indelen (zie: Eid, Diener & Langheine, 2003). We laten de indeling in groepen niet afhangen van een extern criterium, maar van de vraag: hoeveel groepen ('latente klassen') hebben we nodig om de samenhang tussen de verschillende uitspraken in de vragenlijst zo goed mogelijk te beschrijven (Vermeunt, 2004)?

Onderzoekopzet

Er zijn natuurlijk meer methoden om grenswaarden voor psychische vermoeidheid te bepalen dan de twee methoden (de ROC-analyse en de LCA-analyse) die hier met elkaar vergeleken worden. De zinvolheid van de categorisering en de grenswaarden, kan bepaald worden door de construct- en predictieve validiteit van beide methoden te vergelijken. De constructvaliditeit wordt nagegaan door de samenhang te analyseren tussen beide manieren om grenswaarden te bepalen en de oorspronkelijke schaal 'psychische vermoeidheid'. De predictieve validiteit heeft betrekking op de mate waarin beide methoden externe criteria kunnen voorspellen. Een voor de hand liggend extern criterium is de registratie van afwezigheden door bijvoorbeeld het bedrijf of arbeidsgeneeskundige diensten. Over zulke 'harde' criteria beschikken we echter niet. Daarom gebruiken we de vragen van de WBM als 'extern' criterium: zelfgerapporteerde afwezigheid wegens ziekte, psychosomatische klachten en lichamelijke klachten.

In deze bijdrage wordt gebruik gemaakt van de SERV-databank 'Vlaamse Werkbaarheidsmonitor'. Het betreft een representatieve steekproef (met een netto respons van 60,6%) van de Vlaamse beroepsbevolking, die na kwaliteitscontrole 11 099 respondenten telt (Bourdeaud'hui et al., 2004). Naar achtergrondkenmerken (zoals geslacht, sector en functieniveau) wordt de populatie van de loontrekkenden in het Vlaamse gewest (exclusief (werk)studenten) dus goed weerspiegeld. Enkel de min-25-jarigen zijn lichtjes ondervertegenwoordigd in de steekproef.

Receiver Operating Characteristics (ROC)

Voor de WBM werd gekozen om met 'gelede kengetallen' te werken. Analoot met de kleurensymboliek van een verkeerslicht, wordt daarbij een problematische groep afgebakend ("oranje knipperlicht"). Daarbinnen wordt tevens een subgroep bepaald die acute problemen vertoont ("rood alarmsignaal"). De grenswaarden voor psychische vermoeidheid (op basis van de VBBA-schaal 'herstelbehoefte') worden vastgelegd via een ROC-analyse. Dat is een simulatietechniek om een cutoff score te bepalen voor een test (bijvoorbeeld een HIV-test) in functie van een maximale discriminatie tussen twee groepen (bijvoorbeeld besmet versus niet besmet of geen overspanningverschijnselen versus wel overspanningverschijnselen).

De ROC-curve laat een trade-off zien tussen de sensitiviteit (de percentages effectieve probleemgevallen die een test vaststelt) en de specificiteit (de percentages niet-probleemgevallen die een test correct indeelt).² Een optimale cutoff is deze waar zo min mogelijk probleemgevallen door de mazen van het testnet glippen, terwijl tevens zo weinig mogelijk niet-probleemgevallen als probleemgeval worden beschouwd. Bij de ROC-analyse wordt uitgegaan van een extern criterium voor het onderscheiden van de 'geen probleem/probleem'-groepen. Die indeling wordt in de WBM gemaakt op basis van andere gegevens uit de vragenlijst, omdat externe informatie over de respondenten ontbreekt. Voor problematische psychische vermoeidheid wordt geopteerd voor een samengesteld criterium op basis van de rapportering van slaapstoornissen, psychosomatische klachten en langdurig ziekteverzuim, gecombineerd met een gevalideerde maat voor de psychische gezondheidstoestand (de COOP-WONCA-chart 'emotionele toestand'; Köning-Zahn, 1993). Zo komen de STV-onderzoekers tot een probleemgroep van 111 werkenden. De ROC-test voor psychische vermoeidheid levert bij vergelijking van deze probleemgroep met de rest van de onderzoekspopulatie, rekening houdend met de trade-off tussen sensitiviteit en specificiteit, volgende kengetallen (of cutoffs) op: minstens instemmen met zeven van de elf vragen is 'problematisch' en minstens instemmen met tien van de elf vragen is 'acuut problematisch'.

Latente Clusteranalyse

Met behulp van latente clusteranalyse onderscheiden we eveneens groepen werknemers in functie van de mate waarin ze psychisch vermoeid zijn (Notelaers, Vermunt, De Witte & van Veldhoven, 2003). Bij de bepaling van deze groepen wordt echter geen beroep gedaan op een extern criterium. Op basis van de probabiliteiten van mogelijke antwoordpatronen wordt nagegaan hoeveel het minimum aantal groepen is dat we nodig hebben om de verschillende soorten antwoorden te beschrijven (zie: Vermunt, 2004).

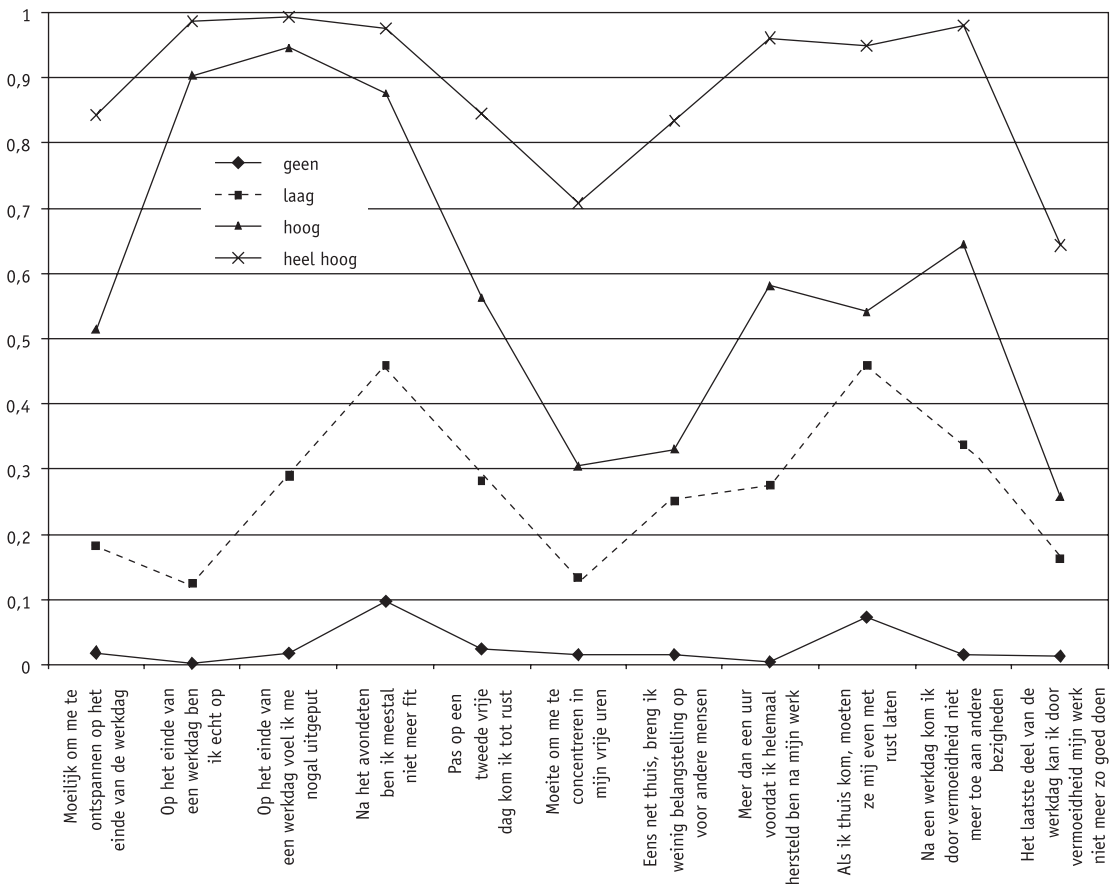
Met latente clusteranalyse in Latent Gold (Vermunt & Magidson, 2003) bekomen we een model dat

de 211 mogelijke combinaties kan herleiden tot slechts 4 clusters.³ De betekenis van de vier clusters kan grafisch worden weergegeven door de proportie 'ja'-antwoorden op de psychische vermoeidheidsvragen af te beelden. In figuur 1 worden de vier clusters weergegeven. Op de x-as staan de vragen over psychische vermoeidheid en op de y-as staat de kans dat men 'ja' antwoordt op een vraag over vermoeidheid.

Twee clusters komen heel duidelijk naar voren: het cluster respondenten dat 'geen vermoeidheid' rapporteert en de cluster respondenten dat 'heel hoge vermoeidheid' rapporteert. De eerste cluster wordt gekenmerkt door de erg lage kans dat de respondenten 'ja' antwoorden. Die kans is voor alle vra-

Figuur 1.

Vier clusters inzake psychische vermoeidheid.



Bron: SERV-databank 'Vlaamse Werkbaarheidsmonitor'

gen immers bijna nul. Deze groep omvat 28,75% van de respondenten uit de werkbaarheidsmonitor. De andere cluster wordt gekenmerkt door een erg hoge kans dat de respondenten 'ja' antwoorden: deze is voor alle vragen bijna 90%. Deze cluster telt 13,47% van de respondenten. Ook de twee clusters tussen de clusters 'geen vermoeidheid' en 'heel hoge vermoeidheid' zijn goed van mekaar te onderscheiden. Hoewel de kansen om 'ja' te antwoorden erg op mekaar gelijken vanaf 'pas op een vrije dag kom ik tot rust', onderscheiden de werknemers 'met hoge psychische vermoeidheid' zich vooral van de werknemers 'met lage psychische vermoeidheid' door hun antwoorden op de eerste vier vragen in figuur 1 ('moeilijk onspannen op het einde van de werkdag', 'eind van de werkdag ben ik echt op', 'eind van de werkdag nogal uitgeput' en 'na het avondeten niet meer fit'). Op deze vragen is de kans dat de werknemers uit de 'hoge vermoeidheidsgroep' 'ja' antwoorden immers opvallend hoger. Ongeveer 28,72% van de werknemers behoort tot het 'laag vermoeidheidscluster', en 28,72% scoort hoog op vermoeidheid. Naar analogie met de problematische groep van de WBM kunnen we stellen dat werknemers die een 'hoge psychische vermoeidheid' rapporteren 'erg moe' zijn na hun werk. De werknemers die een 'heel hoge psychische vermoeid' vertonen zijn net zoals de 'acuut problematische groep' uit de WBM niet alleen erg moe, maar hebben daarenboven ook erg veel nood aan herstel.

ROC en LCA: construct- en predictieve validiteit vergeleken

Als we de constructvaliditeit onderzoeken aan de hand van de samenhang tussen beide methoden (ROC versus LCA) en de oorspronkelijke VBBA-schaal 'herstelbehoefte', dan stellen we vast dat de correlatie (Kendal tau b) tussen de LCA-resultaten en psychische vermoeidheid ($r = 0,86$) groter is dan de correlatie van de ROC-methode met de psychische vermoeidheidschaal ($r = 0,71$).

De predictieve validiteit kan onderzocht worden door te analyseren hoeveel werknemers juist geklasseerd werden door beide methoden. We moeten dus nagaan hoeveel werknemers die geen problemen rapporteren als niet-problematisch geklasseerd worden en hoeveel werknemers die problemen rapporteren als problematisch geklasseerd worden.

Een eerste variabele die bij de berekening van de kengetallen in de WBM gebruikt wordt, betreft het aantal keren dat men afwezig was wegens ziekte. In tabel 1 worden hiervan de steekproefpercentages weergegeven. Per methode is de som van deze percentages 100. Om te vergelijken hoeveel respondenten juist geklasseerd worden, vergelijken we de resultaten van beide methoden met deze in de laatste kolom, die het totale rijpercentage weergeeft.

Tabel 1.

Ziekteperiodes gekruist met de kengetallen van de WBM en de clusters van de LCA.

Ziekteperiodes de afgelopen 12 maanden	LCA-model blootstelling aan psychische vermoeidheid				ROC-WBM			Totaal rij - percentage
	niet problematisch		proble- matisch	acuut proble- matisch	niet proble- matisch	proble- matisch	acuut	
	geen	laag	hoog	heel hoog	groen	oranje	rood	
geen enkele keer	16,3	12,3	12	4,69	34,7	7	3,73	45,4
1 keer	10,5	8,59	9,49	4,42	23,4	6,3	3,27	33
2 keer	3,57	3,7	4,52	2,67	9,21	3,3	1,98	14,4
3 of 4 keer	0,93	1,27	2,15	1,46	2,95	1,7	1,19	5,82
5 keer of meer	0,22	0,29	0,4	0,46	0,63	0,3	0,4	1,38

Bron: SERV-databank 'Vlaamse Werkbaarheidsmonitor'

De LCA klasseert 63% ($= ((16,3 + 12,3) / 45,4) * 100$) van de werknemers die geen enkele keer afwezig zijn wegens ziekte in de niet-problematische clusters. Bij de ROC-analyse is dat 76% ($= (34,7 / 45,4) * 100$). Van de respondenten die één keer afwezig zijn, wordt 58% door de LCA geklasseerd als niet-problematisch, terwijl 71% als niet problematisch gezien wordt door de WBM. Van de werknemers die twee keer afwezig zijn, wordt 50% als problematisch beschouwd door de LCA-methode. Bij de ROC-methode is dat echter minder: 36,6%. Van de werknemers die drie of vier keer afwezig zijn wegens ziekte, heeft 62% volgens de LCA-methode een problematisch niveau van psychische vermoeidheid. Bij de WBM is dat bijna 50%. Van de respondenten die vijf keer of meer afwezig zijn, wordt 62,3% door de LCA-methode geklasseerd in de problematische clusters. Bij de WBM is dat 50,7%. We stellen dus vast dat de LCA-methode sterker is in het identificeren van de probleemgevallen, ook al werd bij het ontwerp van de WBM rekening gehouden met dit externe criterium.

Een tweede reeks variabelen die in aanmerking komt voor een soortgelijke analyse zijn vragen (in de WBM-vragenlijst) naar klachten, zoals rugpijn, pijn in maag en darmen, of een gestoorde slaap. De antwoordmogelijkheden zijn 'ja' of 'neen'. In tabel 2 worden enkel de 'ja'-antwoorden weergegeven. Ditmaal gaan we na in welke mate beide methodes de 'ja'-antwoorders als psychisch vermoeid klasseren. In tabel 2 wordt na elke methode een kolom voorzien die weergeeft hoeveel procent van de 'ja'-antwoorders als problematisch geklasseerd werd.

Uit de vergelijking van de kolommen '% problematisch' blijkt dat de LCA-methode meer respondenten die geestelijke of lichamelijke klachten melden onder de problematische clusters klasseert dan de ROC-methode. We besluiten dan ook dat de LCA-methode meer 'probleemgevallen' juist klasseert dan de ROC-methode. Dat de ROC-methode meer respondenten met problemen in de niet-problematische groep klasseert, is verrassend. Deze vragen werden bij het ontwerp van de kengetallen immers als extern criterium in de analyse opgenomen.

Conclusie

Een eerste vaststelling is dat de latente clusteranalyse, die twee niet-problematische clusters en twee problematische clusters onderscheidt, meer probleemgevallen rapporteert dan de WBM. Op het eerste zicht lijkt het alsof de LCA-methode het aantal probleemgevallen overschat. Men zou daaruit kunnen concluderen dat deze methode minder geschikt is om het aantal psychisch vermoeiden te bepalen in een representatieve steekproef. Verdere analyses spreken deze suggestie echter tegen. De overeenstemming tussen de oorspronkelijke psychische vermoeidheidsschaal en de LCA-methode is immers groter dan de overeenstemming tussen de ROC-methode en deze schaal. Verder blijkt de LCA-methode aanzienlijk beter probleemgevallen te identificeren dan het verkeerslichtenmodel dat voorgesteld werd door de WBM. De door ons gerapporteerde analyse blijkt dus een grotere constructvaliditeit én predictieve validiteit te vertonen dan het verkeerslichtenmodel dat gerapporteerd werd in de werkbaarheidsmonitor.

Verder onderzoek over de LCA-methode is echter nodig, voor we definitieve conclusies kunnen trekken. Vooreerst is het nog onduidelijk in hoeverre de LCA-analyse te streng is. Uit Nederlands onderzoek blijkt immers dat slechts één derde van alle verzuimgevallen veroorzaakt wordt door stress of vermoeidheid (Schaufeli, Bakker & De Jonge, 2003). De percentages die door de LCA worden gesuggereerd liggen echter hoger. Dit laatste geldt overigens ook voor de kengetallen van de WBM. In toekomstig onderzoek hebben we 'hardere' externe criteria nodig dan zelfgerapporteerd ziekteverzuim. Het zou dan ook goed zijn om betrouwbare registraties van ziekteverzuim te koppelen aan de gegevens van de WBM.

Verder gelden onze conclusies enkel voor psychische vermoeidheid. Deze variabele is weliswaar een cruciale parameter voor de kwaliteit van de arbeid. Toch moet ook onderzocht worden of we tot analoge conclusies komen, wanneer we andere indicatoren van de kwaliteit van de arbeid onderzoeken, zoals bijvoorbeeld plezier in het werk of leer-mogelijkheden.

Naar de toekomst toe stelt zich de vraag of de afkappunten die in de WBM gehanteerd worden, ge-

Tabel 2.
% ja-antwoorden op klachten verdeeld over classificatiemethoden

Had u de laatste twee weken last van...	blootstelling niveau aan psychische vermoeidheid met LCA				ROC – WBM kengetallen				
	niet problematisch		problematisch		problematisch				
	geen	laag	hoog	heel hoog	niet problematisch	oranje	rood		
Rugpijn?	9,41	11,2	15,1	8,76	53,6	10,74	6,64	39	44,5
Maag en darmklachten?	3,85	5,87	8,51	6,28	60,3	6,447	5,07	47	24,5
Hoofdpijn?	7,85	10,1	14,5	8,56	56,2	10,16	6,84	41,5	41
Borst of hartstreek?	1,55	2,41	4,67	4,42	69,9	3,99	3,75	59,5	13
Nek- of schouderpijnen?	9,57	12,3	16,5	9,37	54,2	11,38	7,2	39	47,7
Moelijk in slaap kunnen komen?	5,87	8,85	11,9	8,52	58,2	9,425	6,89	46,5	35,1
Spierpijnen in de ledematen?	5,66	7,25	11	7,26	58,5	8,137	5,59	44	31,2
Aanslepende vermoeidheid?	2,05	5,99	13,7	10,5	75,2	11,23	8,37	61	32,2
Tintelingen/verdoofd gevoel in de ledematen?	2,35	3,3	5,7	4	63	4,512	3,3	50,7	15,4
Onrustige of gestoorde slaap?	6,91	10,5	14,7	9,81	58,4	11,32	7,79	45,5	42

Bron: SERV-databank 'Vlaamse Werkbaarheidsmonitor'

voelig genoeg zijn om veranderingen te meten, wanneer men een trendanalyse wil uitvoeren. Er dienen immers vrij veel respondenten te verschuiven van bijvoorbeeld tien naar elf antwoorden, voordat een verandering kan worden vastgesteld. Dit probleem stelt zich niet met latente clusteranalyse. Daarin kan met latente Markovmodellen (Vermunt, 1999; Paas, Bijmolt & Vermunt, 2004) immers een statistisch meetmodel ontwikkeld worden, waarbij getoetst wordt of clustertransities verantwoordelijk zijn voor veranderingen in bijvoorbeeld zelfgerapporteerde afwezigheid wegens ziekte.

Guy Notelaers

Directie Onderzoek voor de Verbetering van de Arbeidsomstandigheden.

Onderzoeksgroep Stress, Gezondheid en Welzijn,

Departement Psychologie

K.U. Leuven

Hans De Witte

Onderzoeksgroep Stress, Gezondheid en Welzijn,

Departement Psychologie

K.U. Leuven

Marc van Veldhoven

Faculteit der Sociale Wetenschappen.

Departement Personeelwetenschappen

Universiteit van Tilburg

Noten

1. Voor dit onderzoek is gebruik gemaakt van de SERV-databank 'Vlaamse Werkbaarheidsmonitor'. De inhoud van de tekst bindt alleen de auteurs.
2. In ROC-termen betekent dit dat bij de cutoff 'oranje knipperlicht' geopteerd werd voor het meettechnische optimum 'maximale sensitiviteit + specificiteit'. De cutoff 'rood alarmsignaal' werd bepaald in functie van een sterke specificiteit van +90% (potentiële classeringsfout 10%).
3. Het LCA-model met een L^2 van 1895,3278 en 1979 vrijheidsgraden is significant en vertoont de laagste BIC (-16243,7139) (zie passingsmaten in Magidson & Vermunt, 2004).

Literatuur

- Bourdeaud'hui, R., Janssens, F. & Vanderhaeghe, S. (2004). 'Nulmeting Vlaamse Werkbaarheidsmonitor'. *Indicatoren voor de kwaliteit van de arbeid op de Vlaamse arbeidsmarkt*. Brussel: STV innovatie & arbeid. Sociaal Economische Raad Vlaanderen.
- Eid, M., Langeheine, R. & Diener, E. (2003). Comparing Typological Structures across Cultures by Multigroup Latent Class Analysis. *Journal of Cross-Cultural Psychology*, 34(2), pp. 195-210.
- Magidson, J. & Vermunt, J.K. (2004). Latent Class Models. In D. Kaplan (Red.), *The Sage Handbook for Quantitative Methodology* (pp. 175-198). Thousand Oakes: Sage Publications.
- Notelaers, G., Vermunt, J., De Witte, H. & van Veldhoven, M. (2003). Estimating exposure rates to psychosocial hazards. The use of Latent Class Analysis to conduct risk assessment with standardized questionnaires. In Flexman, e.a. (Reds.), *Proceedings of the 5th Conference of the European Academy of Occupational Health Psychology*. Berlin, 20-21 November 2003: Flexibility, Quality of Working Life and Health.
- Notelaers, G., De Witte, H. & Vermunt, J.K. (2005). The Usefulness of Latent Class Analysis for Work and Organisational Psychology: An Illustration Regarding the Job Demand Control Support Model. *Paper presented at the European Congress Work and Organizational Psychology: book of abstracts*, Istanbul, Turkey.
- Paas, L.J., Bijmolt, T.H. & Vermunt, J.K. (2004). *Extending dynamic Segmentation with Lead Generation: A Latent Class Markov Approach*. [Manuscript submitted for publication].
- Schaufeli, W., Bakker, A. & De Jonge, J. (2003). *De psychologie van arbeid en gezondheid*. Houten/Mechelen: Bohn Stafleu Van Loghum.
- van Veldhoven, M. & Meijman, T.F. (1994). *Het meten van psychosociale arbeidsbelasting met een vragenlijst: de vragenlijst beleving en beoordeling van de arbeid (VBBA)*. Amsterdam: NIA.
- Vermunt, J.K., Langeheine, R. & Böckenholt, U. (1999). Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates. *Journal of Educational and Behavioral Statistics*, 24, pp. 178-205.
- Vermunt, J.K. & Magidson, J. (2002). Latent Class Cluster Analysis. In J. Hagenaars & McCutcheon (Reds.), *Applied Latent Class Analysis* (pp. 89-106). Cambridge: Cambridge University Press.
- Vermunt, J. & Magidson, J. (2003). *Latent Gold. User's Guide*. Belmont: Statistical Innovations.
- Vermunt, J.K. (2004a). Latent Markov models. In M.S. Lewis-Beck, A. Bryman & T.F. Liao (Reds.), *The Sage Encyclopedia of Social Sciences Research Methods* (pp. 553-554). Thousand Oakes, CA: Sage Publications.
- Vermunt, J.K. (2004b). Toepassingen van latente klasseanalyse in sociaal wetenschappelijk onderzoek. *Sociale Wetenschappen*, 47, pp. 2-14.