

UAC

D 4061

TNO-rapport

**D4: INTERBEOORDELAARSBETROUWBAARHEID
WEBE-METHODIEK**

IBISSTAMBOEKNUMMER

7491

NIPG-publikatienummer
93.029

Juni 1993

Alle rechten voorbehouden.
Niets uit deze uitgave mag worden
vermenigvuldigd en/of openbaar gemaakt
door middel van druk, fotokopie, microfilm
of op welke andere wijze dan ook, zonder
voorafgaande toestemming van TNO.

Indien dit rapport in opdracht werd
uitgebracht, wordt voor de rechten en
verplichtingen van opdrachtgever en
opdrachtnemer verwezen naar de
'Algemene Voorwaarden voor Onderzoeks-
opdrachten aan TNO', dan wel de
betreffende terzake tussen partijen
gesloten overeenkomst.

Het ter inzage geven van het TNO-rapport
aan direct belanghebbenden is toegestaan.

© TNO

BIBLIOTHEEK NEDERLANDS INSTITUUT VOOR
PRAEVENTIEVE GEZONDHEIDSZORG TNO

08 JUL 1993

POSTBUS 124, 2300 AC LEIDEN

S. Dhondt

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Dhondt, S

D4: interbeoordelaarsbetrouwbaarheid WEBA-methodiek / S.

Dhondt. - Leiden: Nederlands Instituut voor Praeventieve
Gezondheidszorg TNO

NIPG-publikatienr. 93.029. - Onderzoek in het kader van
het DGA-TNO-Onderzoeksprogramma 'Kwaliteit van de arbeid
1991-1995'. - Met lit. opg.

ISBN 90-6743-253-9

Trefw.: arbeidsomstandigheden; kledingindustrie.

93.029.029
93.029.029
93.029.029
93.029.029

Deze uitgave is te bestellen door het overmaken van f 27,85 (incl. BTW) op postbankrekeningnr.
99.889 ten name van het NIPG-TNO te Leiden onder vermelding van bestelnummer 93.029.

INHOUD	pagina
SAMENVATTING	i
1. INLEIDING	1
1.1 De WEBA-methodiek	1
1.2 Betrouwbaarheidsonderzoek	2
1.3 Onderzoeksvragen	5
2. ONTWERP VAN HET BETROUWBAARHEIDSONDERZOEK	7
2.1 De WEBA-methodiek	7
2.2 Aantal beoordelaars	8
2.3 Functies	8
2.4 Design	9
2.5 Analyse	10
3. RESULTATEN	13
3.1 Ervaring beoordelaar	13
3.2 Functie	13
3.3 Variabele of conditie	14
3.4 Ervaring en variabele	16
3.5 Ervaring en functie	18
3.6 Functie en variabele	18
3.7 Gebruikte criteria	19
4. CONCLUSIE EN DISCUSSIE	21
LITERATUUR	25
BIJLAGEN	27

SAMENVATTING

Doel van het onderzoek was inzicht te krijgen in de interbeoordelaarsbetrouwbaarheid van de WEBA-methodiek, zoals toegepast voor de handleiding S-71. Daarbij werd onderzocht in welke mate de interbeoordelaarsbetrouwbaarheid afhankelijk is van de ervaring van de beoordelaar met de methodiek en van de complexiteit van de te beoordelen functie, en in hoeverre verschillen optreden per welzijnsconditie.

Door een zeer gestructureerd onderzoeksdesign te hanteren en de beoordelaars te vragen de overwegingen bij de beoordeling, dan wel de problemen hierbij nauwkeurig te beschrijven kon het onderzoek tevens informatie opleveren die behulpzaam zou kunnen zijn bij het specificeren van de beoordelingscriteria t.b.v. een betere interbeoordelaarsovereenstemming en bij het trainen van observatoren in het gebruik van de WEBA-methodiek.

Het onderzoek werd uitgevoerd in het kader van het DGA-TNO-Onderzoekprogramma 'Kwaliteit van de arbeid 1991-1995'.

Methode:

In het totaal hebben 22 beoordelaars aan het onderzoek deelgenomen. Alle beoordelaars waren bekend met de WEBA-methodiek volgens de principes van de handleiding S-71 en met haar theoretische achtergronden; zij hadden elk minimaal één WEBA-analyse uitgevoerd. Van de oorspronkelijke groep hadden 8 beoordelaars elf WEBA-beoordelingen of meer uitgevoerd. Zij werden derhalve als 'experts' aangeduid. De overigen werden, vanwege hun geringere WEBA-ervaring, als 'niet-experts' aangeduid.

Om de te beoordelen functies niet al teveel te laten verschillen op irrelevante kenmerken is ervoor gekozen om vier varianten van eenzelfde beroep, dat van naaister, te laten beoordelen. Het ging om twee functies die 'gunstiger' zouden moeten scoren op een aantal WEBA-criteria omdat dit functies in een taakgroep* waren; de overige twee functies zouden ongunstiger moeten scoren op de WEBA-criteria (de ene functie was een lijnfunctie, de andere een bewakingsfunctie aan een automaat).

*Een taakgroep vormt een cel of groep van werknemers die meer autonomie krijgen in de invulling van hun werk. Werknemers binnen een taakgroep kunnen over het algemeen meer functies uitoefenen, meer beslissingen nemen over de wijze en de volgorde waarop wordt gewerkt. Regelmogelijkheden van functies binnen een taakgroep zouden groter zijn dan van functies in klassieke lijnstructuren.

Vijf van de experts en zeven niet-experts hebben alle vier de functies beoordeeld. De overige beoordelaars hebben in ieder geval minimaal één functie beoordeeld en iedereen had in ieder geval éénzelfde functie, de eerste 'taakgroep' functie, beoordeeld.

Ten behoeve van standaardisering en opdat iedere beoordelaar op dezelfde wijze kennis zou nemen van het werk op een betreffende werkplek, zijn bij alle vier werkplekken video-opnamen gemaakt. Gesprekken met de uitvoerenden over hun functie zijn alle uitgetypt en tezamen met de video-opnamen aan de beoordelaars aangeboden.

De interbeoordelaarsbetrouwbaarheid is op twee wijzen gekwantificeerd:

- overeenstemmingspercentage met de 'norm'; één van de expert-beoordelaars was uitermate goed op de hoogte van het betreffende werk en heeft de video-opnamen geredigeerd. Deze beoordeling is als 'norm'beoordeling gehanteerd;
- Cohen's Kappa; een maat voor overeenstemming tussen beoordelaars, waarbij rekening wordt gehouden met overeenstemming als gevolg van kans. Bij een drie-antwoord keuze is de kans dat beoordelaars een score op basis van 'kans' invullen, 33%. Een Cohen's Kappa van 0 wil zeggen dat de overeenstemming niet van de kansverdeling verschilt, significant < 0 dat de kans geringer is dan de kansverdeling en significant > 0 dat er sprake is van een betere overeenstemming dan op basis van de kansverdeling. De Cohen's Kappa kan maximaal 1 worden, waarden $> .75$ worden als een 'uitstekende overeenstemming', beoordeeld, $.40 < \text{Kappa} \leq .75$ als redelijk en $\leq .40$ als matig tot slecht. De Cohen's Kappa is alleen per conditie berekend; het was niet zinvol deze maat over alle condities heen te berekenen omdat dan de vergelijkingen met te weinig eenheden zouden gebeuren.

De eerste maat laat toe om de afwijking van de norm te meten. Met de tweede maat wordt een correctie op deze eerste maat mogelijk gemaakt: indien iedereen zou afwijken van de norm en indien iedereen voor de rest onderling zou overeenkomen, dan zou de eerste maat een vertekend beeld geven. Cohen's Kappa laat toe om de algehele overeenstemming tussen beoordelaars, los van de norm, vast te stellen.

Resultaten:

Over alle beoordelingen is de gemiddelde overeenstemming tussen beoordelaars en normbeoordelaar, voor de functiematrix 57%, voor de regelproblemen 68% en voor de welzijnscondities (het

WEBA-profiel) eveneens 57%. Dit percentage is niet hoog, ongeveer een derde van alle beoordelingen is door de beoordelaars niet hetzelfde beoordeeld als door de 'norm-beoordelaar'.

Onderzocht is of de verschillen in overeenstemming te maken hebben met ervaring van de beoordelaar, complexiteit van de functie of de welzijnsconditie (elk van de 7 WEBA-criteria).

In tabel 1 staat het gemiddelde percentage overeenstemming en de Cohen's Kappa weergegeven voor het WEBA-profiel, uitgesplitst naar ervaring van de beoordelaar en naar te beoordelen conditie. De resultaten in deze tabel zijn in grote lijnen vergelijkbaar met de resultaten voor de functiematrix en de regelproblemen-inventarisatie. Uit tabel 1 (en ook uit de gegevens over de functiematrix en regelproblemen) blijkt dat ervaring er niet zoveel toe doet. Wel zijn er verschillen in interbeoordelaar overeenstemming tussen de welzijnscondities. Opgemerkt moet worden dat een enkele conditie, kort-cyclische taken, een hele grote overeenstemming laat zien in termen van het percentage maar een hele lage Kappa heeft. Dit heeft te maken met het feit dat wanneer er (bijna) volledige overeenkomst is tussen beoordelaars, er nauwelijks spreiding is. Als er nauwelijks spreiding in getallen is zal een associatie-maat, zoals de Cohen's Kappa, per definitie (bijna) nul zijn.

Wanneer de gegevens worden uitgesplitst naar type functie blijkt dat er geen systematisch verschil is in de mate van overeenstemming en de veronderstelde complexiteit van de functie. Wederom komen de eerder aangegeven verschillen in overeenstemming tussen welzijnscondities naar voren.

Conclusies:

Op basis van dit onderzoek moet worden geconcludeerd dat:

1. Circa 60% van alle beoordelingen (de drie analyseniveaus samengenomen) overeenstemt met de 'norm'. Deze overeenstemming is aan de lage kant.
2. Ervaring met de WEBA-methodiek niet leidt tot een betere overeenstemming.
3. De functie waarvan werd verondersteld dat hij het meest simpel was blijkt iets meer in overeenstemming met de norm beoordeeld te worden dan de complexere functies. De verschillen zijn echter zeer gering.
4. Er grote verschillen zijn in overeenstemming met de norm voor de verschillende welzijnscondities: de beste overeenstemming wordt behaald voor volledigheid, autonomie, contactmogelijkheden en organiserende taken. Wegens geringe spreiding in oordelen kon de consistentie in de beoordeling van kort-cyclische taken niet goed worden beoordeeld.

De voornaamste conclusie is dat de interbeoordelaarsbetrouwbaarheid van de WEBA-methodiek op grond van dit onderzoek laag moet worden genoemd. Dit is voornamelijk te wijten aan het

inconsistent toepassen van de verschillende criteria die zijn geformuleerd om een beoordeling te maken. De criteria laten teveel ruimte in de beoordeling toe. Dit betekent in de praktijk dat men bij het beoordelen van een functie aan de hand van WEBA een groot risico loopt ten onrechte welzijnsrisico's te identificeren waar deze er niet zijn of juist geen welzijnsrisico's te onderkennen waar ze er wel zijn.

Tabel 1 Overzicht van overeenstemmingsindicatoren tussen experts en niet-experts naar conditie (n=12)

Welzijnsconditie	% overeenstemming met norm			Cohen's Kappa
	gemiddeld	experts	niet-experts	gemiddeld
volledigheid	44%	36%	48%	.30**
organiserende taken	50%	63%	41%	.17**
niet-kortcyclisch	95%	92%	97%	.00
moeilijkheid	86%	77%	92%	.11
autonomie	29%	37%	25%	.28**
contactmogelijkheden	69%	78%	63%	.25**
informatievoorziening	24%	25%	22%	.01

** p < .001

1. INLEIDING

In 1990 is de nieuwe Arbeidsomstandighedenwet volledig van kracht geworden. Bij het organiseren van de arbeid, het inrichten van de arbeidsplaatsen en het bepalen van de produktie- en werkmethoden moet de werkgever dan, in het kader van zijn zorg voor het bevorderen van het welzijn bij de arbeid, het in artikel 3 gestelde in acht nemen, gelet onder meer de stand van de arbeids- en bedrijfskunde. Om werkplekken, werkplekverbeteringen en de organisatie van de arbeid volgens de criteria van de wet te kunnen beoordelen is door de Projectgroep WEBA (Welzijn Bij de Arbeid) de WEBA-methodiek ontwikkeld. In deze methodiek worden welzijnsrisico's die verbonden zijn aan arbeidssituaties in kaart gebracht en beoordeeld. Een andere doelstelling van de methodiek is aanwijzingen te geven hoe deze arbeidssituaties kunnen worden verbeterd.

Zowel vanuit het bedrijfsleven als vanuit de (internationale) onderzoekswereld wordt regelmatig gevraagd naar de methodologische onderbouwing van de WEBA-methodiek. Om deze vraag te kunnen beantwoorden is in het DGA-TNO-Onderzoekprogramma 'Kwaliteit van de arbeid 1991-1995' onderzoek voorzien in betrouwbaarheids- en validiteitsonderzoek van de methodiek. In dit project 'Interbeoordelaarsbetrouwbaarheid WEBA' wordt nagegaan in welke mate verschillende beoordelaars met de WEBA-methodiek een aanvaardbare interbeoordelaarsovereenstemming bereiken.

In deze inleiding wordt verder ingegaan (1.1) op de WEBA-methodiek, (1.2) op het doel en het belang van een betrouwbaarheidsonderzoek en (1.3) op de onderzoeksvragen voor dit onderzoek.

1.1 De WEBA-methodiek

Welzijn bij de arbeid wordt in verband gebracht met minimalisering van risico's voor psychische belasting of stress op de werkplek en met een maximalisering van de leermogelijkheden. *Psychische belasting* of *stress* ontstaat als werknemers in hun werk met problemen worden geconfronteerd terwijl de organisatie van het werk hen niet de mogelijkheid biedt deze problemen op te lossen. We noemen dit regelmogelijkheden. Belangrijke regelmogelijkheden zijn autonomie in het werk (zelf kunnen beslissen over werkplaats, methode, volgorde en timing), organiserende taken en contactmogelijkheden. Daarnaast is het ook van belang te kijken naar de cyclusduur van de belangrijkste taken en de informatievoorziening in het werk. Kortcyclische arbeid (bijv. lopende band-werk) gaat in het algemeen gepaard met extreem lage regelmogelijkheden. Een goede terugkoppeling van

informatie over hoe men presteert en over hoe het bedrijf werkt, is nodig om regelmogelijkheden te kunnen gebruiken. Werknemers kunnen voldoende *leren* in het werk als ze verschillende soorten taken alsook taken van verschillende moeilijkheidsgraden kunnen uitvoeren. Vakmanschap gaat verloren als men slechts eenvoudige taken moet uitvoeren. Een werknemer mag niet uitsluitend moeilijke taken uitvoeren, anders stijgen de kansen op psychische belasting. Voldoende regelmogelijkheden (bijv. autonomie om een eigen werkmethode te kiezen) vergroten de kansen op vakmatig leren. Vakmanschap is ook afhankelijk van het soort taken dat een werknemer moet uitvoeren: krijgt een werknemer alleen uitvoerend werk, dan leert hij of zij niet veel. De werknemer moet zijn eigen werk kunnen voorbereiden en, indien dat nodig is, ondersteuningstaken (bijv. reparaties uitvoeren) kunnen uitvoeren.

De WEBA-methodiek heeft tot doel om condities voor psychische belasting en leermogelijkheden te beschrijven en te beoordelen. In de beschrijvingsfase worden gegevens over de belangrijkste aspecten van de taakuitvoering verzameld. Deze gegevens worden samengevat in een functiematrix en een lijst van regelproblemen. Uitgaande van deze gegevens worden, al naargelang de situatie, samen met de taakuitvoerder zeven kwaliteitsvragen (zie onderstaand kader) beantwoord. Op elke vraag wordt geantwoord of de werkplek onvoldoende, beperkt voldoende ('voor verbetering vatbaar') of voldoende scoort. Vanuit deze beoordeling kunnen dan maatregelen worden afgeleid om de werkplek aan te passen.

Figuur 1 De zeven kwaliteitsvragen van WEBA

- is de functie volledig?
- bevat de functie voldoende organiserende taken?
- kan de functiehouder voldoende niet-kortcyclische taken uitvoeren?
- is de functie niet te moeilijk en niet te gemakkelijk?
- kan de functiehouder met voldoende autonomie zijn problemen oplossen?
- bezit de functiehouder voldoende contactmogelijkheden in zijn taakuitvoering?
- krijgt de functiehouder voldoende informatie over zijn taakuitvoering?

1.2 Betrouwbaarheidsonderzoek

In betrouwbaarheidsonderzoek worden twee aspecten van meetinstrumenten bekeken: in eerste instantie wordt gekeken of metingen stabiel blijven in de tijd ('wordt een zelfde uitkomst gemeten voor een fenomeen dat niet is veranderd?'); in tweede instantie kan worden gekeken naar de mate van overeenstemming tussen gebruikers van een instrument in acht genomen de werkelijke variatie in een variabele ('blijft de overeenstemming over de hele breedte van een variabele bewaard?').

in een variabele ('blijft de overeenstemming over de hele breedte van een variabele bewaard?'). Deze twee aspecten van betrouwbaarheid zijn niet hetzelfde. Men kan dus een hoge overeenstemming tussen beoordelaars bereiken, maar als de beoordelaars maar een klein deel van de werkelijke variantie van een gedrag (bijvoorbeeld) hebben beoordeeld dan is de echte overeenstemming nog niet bepaald (Goodwin, Sands, Kozleski, 1991). Het eerste aspect van betrouwbaarheid wordt gemeten aan de hand van test-hertestonderzoek. Het tweede aspect wordt gemeten aan de hand van interbeoordelaarsonderzoek.

Van de WEBA-methodiek is ons geen betrouwbaarheidsonderzoek bekend. Van de ASA-methodiek (Christis, Fortuin, 1989), waarvan de WEBA-methodiek is afgeleid, is er één betrouwbaarheids-onderzoek uitgevoerd (van Workum, 1990; Grosfeld, 1990). In een beperkt onderzoek met twee koppels beoordelaars die meerdere functies beoordeelden, werd voor vier van vijf ASA-variabelen een redelijke interbeoordelaarsbetrouwbaarheid gevonden. Een aantal aspecten van dit onderzoek laat niet toe de conclusies te veralgemenen: slechts vier beoordelaars deden aan het onderzoek mee wat de veralgemeenbaarheid inperkt; de onafhankelijkheid van de beoordelingen kon niet optimaal worden gerealiseerd (beoordelaars bekeken tegelijkertijd een werkplek); de overeenstemming werd in de hand gewerkt door een gezamenlijke training van de beoordelaars en door de beoordelaars een zeer gedetailleerde functiebeschrijving (fase 1 in WEBA) aan te bieden; de analyse beperkt zich tot het ASA-profiel van de beoordelaars zodat onduidelijk is hoe betrouwbaarheid speelt in de verschillende stappen van de ASA-methodiek; er is niet gekeken of de functies op alle ASA-condities voldoende spreiding vertoonden zodat niet duidelijk is of alleen stabiliteit dan wel interbeoordelaarsbetrouwbaarheid is gemeten.

In het huidige onderzoek wordt geprobeerd om interbeoordelaarsbetrouwbaarheid in de toepassing van alle stappen van de WEBA-methodiek te meten zodat een bredere onderzoeksopzet vereist is dan door Grosfeld is gehanteerd.

Het uitvoeren van een WEBA-analyse verloopt in verschillende fasen en stappen waarbij de eerdere fasen of stappen van invloed zijn voor de volgende fasen of stappen. In figuur 2 worden deze fasen en stappen, zoals ze in het handboek S71 voor de WEBA-methodiek zijn beschreven, grafisch voorgesteld.

Figuur 2 De stappen in een WEBA-beoordeling

Fase A. Beschrijven van de functie.**Stap 1: beschrijven van de kwaliteitskenmerken.**

Analyse van de taken waaruit de functie bestaat.

Stap 2: samenstellen van de functiematrix.

De gegevens van stap 1 worden in een matrix overzichtelijk bij elkaar gezet. Voor de kenmerken van de functie (soort taken, tijdsverdeling taken, cyclustijden, moeilijkheidsgraad, autonomieaspecten, contactmogelijkheden, informatievoorziening) worden totaalscores berekend.

Stap 3: inventariseren van regelproblemen.

De beoordeling van de functie hangt mede af van aantal en aard van voorkomende regelproblemen. Deze worden daarom geïnventariseerd. Regelproblemen situeren zich op de onderdelen die samen een arbeidstaak bepalen, namelijk de normen (productie-opdracht/werkinstructies), het materiaal, de middelen, de operaties, het arbeidsresultaat en de omgeving.

Fase B. Beoordelen van een functie.**Stap 4: beantwoorden van de kwaliteitsvragen.**

Uitgaande van de functiematrix en de regelproblemen worden de zeven kwaliteitsvragen beantwoord. Voor de zeven vragen, zie figuur 1.

Stap 5: samenstellen van het welzijnsprofiel.

De antwoorden uit stap 4 worden in een overzichtelijk profiel samengebracht.

Fase C. Verbeteren van de functie.**Stap 6: aangeven van welzijnsmaatregelen.**

Op basis van de beoordeling worden volgens een bepaalde typologie maatregelen aangegeven die van toepassing zijn.

De vraagstelling met betrekking tot betrouwbaarheid van de WEBA-methodiek luidt:

- komen verschillende beoordelaars in de fase van het beschrijven van een functie tot vergelijkbare resultaten voor de verschillende aspecten van de functie?,
- beoordelen de verschillende beoordelaars de zeven kwaliteitsvragen op een zelfde manier?,
- formuleren beoordelaars dezelfde soort maatregelen, uitgaande van dezelfde beoordeling van een werkplek?

In het betrouwbaarheidsonderzoek wordt nagegaan wanneer (in welke stap) en waarom bepaalde beoordelaars tot een verschillende beoordeling komen. In het onderhavige onderzoek wordt alleen de interbeoordelaarsbetrouwbaarheid van de eerste twee fasen (stappen 1 tot en met 5) onderzocht en wordt niet gekeken naar de laatste fase van de WEBA-methodiek, nl. het formuleren van maatregelen. In deze laatste fase van de WEBA-methodiek is geen sprake van een meetmethodiek en is interbeoordelaarsbetrouwbaarheid moeilijk te operationaliseren.

Het belang van een betrouwbaarheidsonderzoek is afhankelijk van de doelstelling van de WEBA-methodiek. In het Handboek WEBA (S71) wordt de methodiek gepresenteerd als een 'diagnose-' maar ook als een 'discussie-instrument'. In het laatste geval laat de methodiek toe om gegevens aan de oppervlakte te laten komen waarover met taakuitvoerders kan worden gediscussieerd. Het doel van de discussie is om, uitgaande van de beoordeling van de werkplek, te komen tot voorstellen om de werkplek te verbeteren. Indien WEBA alleen een instrument moet zijn om aan werkplaats-herontwerp te doen, dan is het nut van een betrouwbaarheidsonderzoek beperkt. Indien WEBA als diagnose-instrument wordt gebruikt, vooral wanneer deze diagnose directe consequenties heeft (bijv. een negatieve beoordeling door de arbeidsinspectie) dan is een aanvaardbare interbeoordelaarsbetrouwbaarheid een eis en is een betrouwbaarheidsonderzoek noodzakelijk. Met een onbetrouwbaar instrument loopt men het gevaar welzijnsrisico's ten onrechte te identificeren of juist ten onrechte over het hoofd te zien.

Afgezien van het doel waarvoor de WEBA wordt gebruikt, kan onderzoek waarin interbeoordelaarsovereenstemming -of gebrek hieraan- centraal staat belangrijke informatie opleveren over problemen die de gebruiker(s) hebben met het instrument.

1. Zo kunnen onduidelijkheden in de instructie, in de omschrijving van de belangrijke begrippen of onduidelijkheid van de criteria (onvoldoende - beperkt voldoende - voldoende) nauwkeurig worden aangegeven en kunnen oplossingen voor verduidelijking en standaardisering van begrippen en procedures worden gesuggereerd.
2. Ook kan dergelijk onderzoek belangrijke systematische verschillen in de beoordeling als gevolg van beoordelaarskenmerken (bijvoorbeeld ervaring of kennis van beoordelaar) of situatiekenmerken (complexiteit van de functie) aan het licht brengen of nuanceren. Dergelijke factoren waren in interbeoordelaarsbetrouwbaarheid bij andere meetinstrumenten ook van invloed (zie bijvoorbeeld DeNisi, Cornelius, Blencoe (1987) en Cornelius, DeNisi, Blencoe (1984) voor het PAQ-instrument).
3. Informatie uit betrouwbaarheidsonderzoek kan belangrijke informatie opleveren voor opleiden/trainen in het gebruik van de WEBA.

1.3 Onderzoeksvragen

De concrete onderzoeksvragen luiden als volgt:

- wat is de interbeoordelaarsbetrouwbaarheid bij gebruikers van het WEBA-instrument in de fase van het beschrijven van een functie?

- wat is de interbeoordelaarsbetrouwbaarheid bij gebruikers van het WEBA-instrument in de beoordelingsfase van een functie?
- zijn verschillen in overeenstemming terug te brengen tot:
 - verschillen in ervaring tussen beoordelaars: komen beoordelaars met veel ervaring met WEBA tot een grotere overeenstemming dan beoordelaars met een beperkte ervaring?
 - verschillen samenhangend met de complexiteit van functies?
 - bepaalde functieaspecten of WEBA-kwaliteitsvragen? (leveren bepaalde kwaliteitsvragen grotere variantie op dan andere vragen? zijn de criteria om een aspect te beoordelen onvoldoende duidelijk?)

2. ONTWERP VAN HET BETROUWBAARHEIDSONDERZOEK

In het onderzoeksontwerp zijn er beslissingen genomen over respectievelijk: (2.1) de WEBA-methodiek, (2.2) het aantal beoordelaars dat de WEBA-analyses moet uitvoeren, (2.3) het aantal functies dat in de beoordeling moet worden bekeken, (2.4) het onderzoeksdesign en (2.5) de analyseprocedure.

2.1 De WEBA-methodiek

In dit betrouwbaarheidsonderzoek wordt de WEBA-methodiek zoals ze is beschreven in het handboek WEBA -Functieverbetering en organisatie van de arbeid (S71)- onderzocht. De beoordelaars moeten de regels zoals ze in dit handboek zijn vastgelegd volgen. Beoordeling volgens ingeperkte protocollen, zoals bijvoorbeeld in het Conceptvoorlichtingsblad 22 voorzien, wordt niet toegelaten.

De interbeoordelaarsbetrouwbaarheid wordt gemeten op het niveau van de functiematrix (stap 2), de regelproblemen-inventarisatie (stap 3) en op het niveau van het welzijnsprofiel (stap 5). Bij de analyse op de beoordelingen wordt rekening gehouden met volgende waarden die de verschillende werkaspecten en welzijnscondities kunnen aannemen:

- functiematrix: in de functiematrix dienen de beoordelaars voor twaalf variabelen een score te geven variërend van '-' (=geen), 'o' (=beperkt) en '+' (ruim). De beoordelaar kan ook niets invullen als hij/zij van oordeel is dat de video-opname onvoldoende informatie bevat om tot een beoordeling te komen. In de analyse wordt de overeenstemming berekend op deze vier scores.
- regelproblemen-inventarisatie: regelproblemen kunnen zich volgens de WEBA-theorie situeren op zes gebieden. In het WEBA-handboek worden per gebied enkele voorbeelden gegeven van regelproblemen. Elke beoordelaar kan evenwel bijkomende problemen aanduiden. Bij de analyse van de regelproblemen is voornamelijk gekeken of een beoordelaar een bepaald gebied als problematisch aanduidt of niet. Een analyse op het niveau van de individuele problemen leidt te ver en levert een vertekend beeld op. In de praktijk betekent dit dat deze variabelen twee waarden kunnen aannemen: overeenstemming met norm of geen overeenstemming met norm.

- welzijnsprofiel: de zeven welzijnscondities kunnen vier waarden aannemen: onvoldoende - beperkt voldoende - voldoende - niet aangeduid. De overeenstemming wordt berekend op deze vier scores.

Tijdens het hele onderzoek moeten de beoordelaars nauwkeurig op formulieren bijhouden waarom zij bepaalde keuzes hebben gemaakt en of zij op moeilijkheden zijn gestoten. Zo kan achteraf worden vastgesteld waarom de beoordelaars tot verschillende uitspraken zijn gekomen.

2.2 Aantal beoordelaars

Om voldoende betrouwbare resultaten te verwerven moeten minstens 10 beoordelaars de WEBA-beoordelingen uitvoeren. In het onderzoek hebben in het totaal 22 beoordelaars meegedaan. De beoordelaars kunnen in vier groepen worden verdeeld: 6 NIPG-onderzoekers en/of stagiaires, 5 wetenschappelijke medewerkers van onderzoeksinstituten die ervaring hebben met WEBA, 5 medewerkers van het Ministerie van Sociale Zaken (Directoraat-Generaal van de Arbeid) of Arbeidsinspecteurs en, tenslotte, acht studenten (arbeidssociologie) die in hun werkcolleges met WEBA hebben gewerkt. Voor de verschillende beoordelaars geldt dat ze allen minstens één keer met de WEBA-methodiek hebben gewerkt en bekend zijn met de achtergronden van WEBA.

Tussen de beoordelaars bestaat een verschil in ervaring in het toepassen van de WEBA. Van de oorspronkelijke groep hadden acht beoordelaars elf of meer WEBA-beoordelingen uitgevoerd. Zij werden derhalve als 'experts' aangeduid. Dit bleken onderzoekers van het NIPG-TNO (evenwel geen stagiaires) en Arbeidsinspecteurs te zijn. De overige beoordelaars werden, vanwege hun geringe ervaring, als 'niet experts' aangeduid.

Vijf experts en zeven niet-experts hebben alle functies beoordeeld. De overige beoordelaars hebben één of meerdere beoordelingen uitgevoerd.

2.3 Functies

Om te kunnen vaststellen of de beoordelaars in hun uitspraken consistent zijn met een normbeoordeling zijn drie strategieën mogelijk: ofwel wordt het aantal beoordeelde functies per beoordelaar vergroot (zie Grosfeld, 1990), ofwel wordt het aantal beoordelaars per beoordeelde functie opgedreven, ofwel wordt gezocht naar een combinatie van beide strategieën. Er is hier gekozen voor een combinatie van beide strategieën:

- er zullen twee duidelijk van elkaar onderscheiden typen functies worden bekeken: deze types moeten duidelijk verschillen op basisvariabelen van WEBA. Indien daar niet op wordt gelet, kan, zoals reeds gezegd, alleen consistentie worden gemeten en niet inter-beoordelaarsbetrouwbaarheid. Er is gekozen voor een functie waarvan verwacht mag/moet worden dat deze functie op een aantal van de kwaliteitsvragen 'voldoende' zal scoren, en daarnaast voor een functie die op de kwaliteitsvragen 'onvoldoende' (of 'beperkt voldoende') scoort. Daarbij dienen de functies toch zo gelijk mogelijk te zijn opdat de verschillen niet door andere (niet relevante) kenmerken tot verschillen in beoordeling leiden. Zo kunnen we vaststellen of overeenstemming over verschillende functies daadwerkelijk aantoonbaar is. Daarvoor is gekeken naar functies uit de confectie-industrie: twee 'klassieke' naaifuncties, -een klassieke lijnfunctie en een naaister aan een naaiauto-maat- en twee naaifuncties werkend in taakgroepen. In recent onderzoek werd vastgesteld dat de 'klassieke' naaifuncties een groot aantal welzijnsrisico's bezitten en dat taakgroep-functies goed scoren op het welzijnsprofiel (zie Peeters, 1991; 1992).
- er moeten meerdere (minstens 2) functiehouders per type functie worden geobserveerd om toevallige verschillen of overeenkomsten uit te schakelen.

In het volledige design worden er dus (in principe) 4 functies door elke beoordelaar geanalyseerd.

2.4 Design

Bij het opzetten van een interbeoordelaarsbetrouwbaarheid-onderzoek wordt men geconfronteerd met drie problemen: ten eerste moet men ervoor zorgen dat de informatie die wordt beoordeeld, dezelfde is; ten tweede moet men ervoor zorgen dat beoordelaars onafhankelijk van elkaar een werkplek kunnen beoordelen; ten derde moet men vermijden dat de wijze waarop de beoordeelde functiehouder de taak uitvoert in de tijd verandert. Dergelijke eisen kunnen het best gegarandeerd worden door de beoordelaars videobeelden van vier functiehouders aan te bieden. De beelden werden opgesteld door een WEBA-expert die zeer goed bekend is met deze functies en van wie de WEBA-beoordeling als norm is gehanteerd. Bij de beoordeling van de werkplekken dienden de beoordelaars aan te geven of het videomateriaal onvoldoende was om de beoordeling uit te voeren. Dit bleek niet het geval te zijn.

2.5 Analyse

Er bestaan in de psychometrische literatuur verschillende maten om interbeoordelaarsbetrouwbaarheid voor nominale of ordinale data te bepalen. Twee belangrijke maten worden in dit onderzoek gebruikt: het percentage overeenstemming (a) en Cohen's Kappa (b).

ad a. De meest gebruikte maat voor nominale data is het *percentage overeenstemming* tussen beoordelaars (zie Grosfeld, 1990; Hartmann, 1977; Goodwin, Sands, Kozleski, 1991). In deze studie wordt het percentage overeenstemming berekend op de overeenstemming van de beoordelaars met een normbeoordelaar. Een perfecte overeenstemming bestaat er als 100% van de beoordelingen met de norm overeenkomen, een lage overeenstemming is er als minder dan 50% van de beoordelingen overeenstemmen met de norm (dan is er meer verschil van mening dan eensgezindheid). Een nadeel van deze maat is dat deze geen correctie voorziet voor overeenstemming die te wijten is aan het beperkt aantal score-categorieën van een variabele of overeenstemming door kans. Voor de WEBA betekent dit:

- functiematrix: er zijn in de beoordeling vier antwoorden mogelijk zodat de kans om toevallig hetzelfde antwoord als de norm te kiezen 25% bedraagt,
- regelproblemen: er zijn slechts twee categorieën mogelijk in de beoordeling zodat de kans om toevallig hetzelfde antwoord als de norm te kiezen 50% bedraagt,
- welzijnsprofiel: er zijn in de beoordeling vier antwoorden mogelijk zodat de kans om toevallig hetzelfde antwoord als de norm te kiezen 25% bedraagt.

ad b. Een maat die een correctie voor een dergelijke kans voorziet is *Cohen's Kappa* (Cohen, 1968; Fleiss, Cohen & Everitt, 1969). Deze maat werd speciaal ontwikkeld om interbeoordelaarsbetrouwbaarheid te meten en wordt als volgt berekend:

$$K = \frac{Pa - Po}{1 - Po} \quad (1)$$

P_a is het percentage geobserveerde overeenstemmingen in een populatie beoordelaars ** en P_o is het percentage van verwachte overeenstemmingen of overeenstemmingen dat men volgens de kansverdeling had kunnen verwachten. In de teller van de Kappa-formule, $(P_a - P_o)$, wordt het percentage geobserveerde overeenstemmingen expliciet gecorrigeerd voor het percentage van verwachte overeenstemmingen of overeenstemmingen dat men volgens de kansverdeling had kunnen verwachten. In de noemer van de Kappa-formule, $(1 - P_o)$, wordt op een zelfde wijze gecorrigeerd voor de kans-overeenstemmingen. Met andere woorden, Kappa geeft het percentage overeenstemming weer die er bestaat tussen beoordelaars, gecorrigeerd voor de kans-overeenstemmingen.

Kappa kan variëren tussen -1 en +1: '0' betekent dat er geen grotere overeenstemming tussen beoordelaars bestaat dan we zouden kunnen terugvinden door te kijken naar de overeenstemming die volgens kans mogelijk is. Negatieve waarden van Kappa betekenen een overeenstemming die lager is dan volgens kans te verwachten zou zijn. Positieve waarden houden een betere overeenstemming in dan volgens kans te verwachten is (Jenkins, Nadler, Lawler & Cammann, 1975). Landis en Koch (1977) hebben voor verschillende ranges van positieve waarden van de Kappa-maat een beoordeling gegeven van de overeenstemming tussen beoordelaars. Volgens hen betekent een Kappa van groter dan .75 een uitstekende overeenstemming in vergelijking met kans, waarden lager dan .40 stellen een slechte overeenstemming in vergelijking met kans voor, en waarden tussen deze twee grenzen betekenen een redelijke tot goede overeenstemming in vergelijking met kans.

Nog twee opmerkingen over het gebruik van de Kappa-maat zijn nodig. Allereerst worden in de hier gebruikte Kappa-maat alle verschillen tussen waarden even zwaar toegemeten. Soms is het aangewezen om een verschil in beoordeling evenwel niet even zwaar te laten doorwegen. Dan wordt aan afwijkingen van 1 (of meerdere) categorie(ën) toch een mate van overeenstemming toegerekend. De *gewogen Kappa-maat* laat een dergelijke 'geschaalde' afwijking toe. Omdat de onderzochte variabelen (op het niveau van de functiematrix, de regelproblemen-inventarisatie en het welzijnsprofiel) in de WEBA een zeer beperkt aantal categorieën bevat heeft een weging geen zin omdat de kans om af te wijken nihil wordt (Jenkins e.a., 1975; Gupta & Beehr, 1982; Goodwin, Sands, Kozleski, 1991). Ten

**Het verschil met de voorgaande maat 'percentage overeenstemming' is dat in de voorgaande maat de overeenstemming met de normbeoordelaar wordt berekend en dat in dit percentage overeenstemming een algemene overeenstemming wordt berekend. Met 'algemeen' wordt rekening gehouden met de verdeling van de antwoorden over alle categorieën van een variabele.

tweede wordt de Kappa-maat alleen gebruikt voor die vergelijkingen waarvoor een voldoende groot aantal beoordelingen voor handen is. In dit onderzoek wordt uitgegaan van minstens tien beoordelingen om met de Kappa-statistiek op een betrouwbare manier te kunnen werken. In de andere gevallen zal steeds met het percentage overeenstemming worden gewerkt.

3. RESULTATEN

De gemiddelde overeenkomst tussen alle beoordelaars en de normbeoordelaar is voor de functie-matrix 57%-gelijke beoordelingen, voor de regelproblemen-inventarisatie 68% en voor het welzijnsprofiel 57% overeenstemming. Dit percentage is in de drie gevallen niet hoog: ongeveer een derde van alle beoordelingen is door de beoordelaars niet hetzelfde beoordeeld als de 'norm'. Hieronder wordt nader ingegaan op de mogelijk verstorende invloed van ervaring van de beoordeelaar (3.1), de beoordeelde functie (3.2) en de beoordeelde variabele of conditie (3.3). Daarna wordt gekeken hoe deze verstorende invloeden in combinatie spelen: (3.4) ervaring en variabele; (3.5) ervaring en functie; (3.6) functie en variabele. Tenslotte wordt gekeken naar de criteria die werden geformuleerd om tot een beoordeling te komen (3.7).

3.1 Ervaring beoordelaar

Voor de beoordelaars die alle functies hebben beoordeeld vinden we volgende resultaten. Voor de *functiematrix* bereiken de vijf experts een overeenstemming van 55% met de normbeoordelaar, de niet-experts een overeenstemming van 59%. Voor de *regelproblemen-inventarisatie* bereiken de experts opnieuw geen hogere overeenstemming met de norm dan de niet-experts: de experts hebben slechts in 62% van de beoordelingen hetzelfde resultaat als de normbeoordelaar, de niet-experts bereiken in 71% van de beoordeling hetzelfde resultaat. Voor het *welzijnsprofiel* bereiken de experts een overeenstemming van 59% met de norm-beoordelaar en de niet-experts 55%. Ervaring hangt in dit experiment niet samen met een hogere overeenstemming met de norm.

3.2 Functie

In tabel 1 worden de percentages overeenstemming met de norm voor de vier beoordeelde functies gegeven. Van de vier functies wordt de 'klassieke' lijnfunctie het meest consistent met de norm beoordeeld met een maximale overeenstemming van 72%. De naaister aan de naaiautomaat wordt niet consistent met de norm beoordeeld dan de twee taakgroep-functies. De eerste taakgroep-functie werd als eerste functie beoordeeld omdat ze als eerste functie op de video was opgenomen: de overeenstemming met de norm is voor deze functie voor regelproblemen-inventarisatie en welzijnsprofiel het laagst.

Tabel 1 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar. Vergelijking tussen functies en niveau van beoordeling

Functie	Functiematrix	Regelproblemen	Welzijnsprofiel
Taakgroep 1 (n=19)	52%	63%	47%
Taakgroep 2 (n=12)	63%	72%	52%
Lijnfunctie (n=12)	68%	72%	72%
Naaiautomaat-bedieners (n=12)	46%	64%	55%

3.3 Variabele of conditie

De overeenstemming tussen de beoordelaars voor de variabelen en de condities laten grote verschillen zien. In tabel 2 worden de gegevens op twee manieren samengevat: enerzijds is er het percentage overeenstemming tussen beoordelaars en de normbeoordelaar, anderzijds is er de Kappa-maat om de absolute overeenstemming (los van de norm) tussen de beoordelaars, rekening houdend met kans, te bepalen. Het percentage overeenstemming varieert van zeer laag (29% overeenstemming) tot zeer hoog (97%). De Kappa-maat haalt voor geen enkele variabele een bevredigende hoogte: slechts voor 7 van de 25 variabelen wordt een Kappa van meer dan .20 bereikt, slechts één maal ('middelen') bereikt de Kappa-maat een redelijke overeenstemming tussen de beoordelaars in vergelijking met kans. Het feit dat hoge percentages overeenstemming met de norm gepaard gaan met zeer lage Kappa's wijst erop dat de beoordelaars voor de variabelen allen dezelfde waarde hebben gegeven: dit betekent dat er onvoldoende spreiding is binnen deze variabele om de Kappa een acceptabele hoogte te laten behalen. Wat betreft de *functiematrix* varieert de overeenstemming met de norm voor de 12 variabelen tussen de 31% en de 97% overeenstemming. Voor 4 variabelen is er minder dan 50% overeenstemming met de norm. De hoogste overeenstemming is er voor de variabelen 'cyclustijden' (97%), 'moeilijkheidsgraad' (72%) en 'autonomie over methode' (72%). De laagste overeenstemming is er voor 'functionele contactmogelijkheden' (31%) en 'autonomie werkplek' (37%). Wat de Kappa-maat betreft zien we dat geen enkele variabele een 'redelijke' overeenstemming bereikt.

De percentages overeenstemming op het niveau van de *regelproblemen-inventarisatie* variëren van 51% overeenstemming met de norm tot 84%. De hoogste overeenstemming in aanduiding van een regelprobleem is te vinden voor 'materiaal': de meeste beoordelaars vinden dat de naaiers met materiaalproblemen te maken hebben. De laagste overeenstemming is te vinden voor 'operaties'. Wat de algemene overeenstemming tussen de beoordelaars betreft, zien we dat alleen de variabele 'middelen' een redelijke Kappa tussen de beoordelaars oplevert.

De percentages overeenstemming op het niveau van het *welzijnsprofiel* variëren van 24% overeenstemming met de norm tot 95%. De hoogste overeenstemming is te vinden voor de condities 'niet

kortcyclische taken': bijna alle beoordelaars vinden dat de functies 'onvoldoende' scoren, met andere woorden dat de functies té kortcyclisch zijn. De laagste overeenstemming is te vinden voor de condities 'informatievoorziening'(24%) en 'autonomie' (29%). Voor 'autonomie' blijkt de afwijking met de normbeoordelaar systematisch te zijn: de Kappa-maat is hoger dan volgens het percentage overeenstemming is te verwachten.

Tabel 2 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar, algemene overeenstemming tussen beoordelaars met Kappa. Vergelijking tussen variabelen en niveau van beoordeling

	% overeenstemming	Kappa
Functiematrix	(n=55) ^{***}	(n=44) ^{****}
cyclustijden	97%	-0.02
moeilijkheidsgraad	72%	0.01
autonomie tempo	63%	0.01
autonomie methode	72%	0.34 **
autonomie volgorde	69%	0.29 **
autonomie werkplek	37%	0.05
autonomie omstandigheden	38%	-0.07
ondersteunende contacten	50%	0.06
functionele contacten	31%	0.00
sociale contacten	49%	0.05
taakinformatie	50%	-0.02
organisatie informatie	58%	-0.00
Regelproblemen-inventarisatie		
normen	55%	0.29 **
materiaal	84%	-0.07
middelen	79%	0.39 **
operaties	51%	0.06
resultaat	64%	-0.08
omgeving	72%	0.05
Welzijnsprofiel		
volledigheid	44%	0.30 **
organiserende taken	50%	0.17 **
niet-kortcyclische taken	95%	-0.00
moeilijkheid	86%	0.11

^{***}Dit aantal is de som van 19 beoordelaars voor de eerste taakgroepfunctie, 12 beoordelaars voor de overige drie functies.

^{****}Dit aantal is de som van de 11 beoordelaars voor de vier functies.

Vervolg tabel 2 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar, algemene overeenstemming tussen beoordelaars met Kappa. Vergelijking tussen variabelen en niveau van beoordeling

	% overeenstemming	Kappa
autonomie	29%	0.28 **
contactmogelijkheden	69%	0.25 **
informatievoorziening	24%	-0.01

** p < .001

3.4 Ervaring en variabele

In tabel 3 worden de percentages overeenstemming vergeleken per variabele tussen de expert en de niet-expert beoordelaars. Slechts voor 10 van de 25 meetpunten scoren de experts beter dan de niet-experts. Wat wel opvalt is dat de experts en de niet-experts naar de te beoordelen variabelen in dezelfde richting met de norm afwijken of overeenstemmen. Voor 21 van de 25 variabelen is de afwijking of overeenstemming in dezelfde richting. Dit wijst erop dat bepaalde criteria om variabelen of condities te beoordelen door experts en niet-experts op dezelfde manier worden toegepast als de norm, maar dat voor een reeks andere variabelen of condities de criteria tot verschillende toepassing leiden bij zowel experts als niet-experts. De enige variabelen waarvoor een groot verschil in overeenstemming met de norm tussen experts en niet-experts werd vastgesteld, waren: 'autonomie over tempo' en 'ondersteunende contacten' op het niveau van de functiematrix; 'organiserende taken' op het niveau van het welzijnsprofiel.

Tabel 3 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar. Vergelijking tussen experts - niet-experts voor de verschillende variabelen, naargelang niveau van beoordeling

	% overeenstemming experts	% overeenstemming niet-experts
Functiematrix	n=20 ^{****}	n=35 ^{*****}
cyclustijden	100%	96%
moeilijkheidsgraad	70%	71%
autonomie tempo	30%	82%
autonomie methode	80%	71%
autonomie volgorde	65%	73%
autonomie werkplek	30%	43%
autonomie omstandigheden	35%	39%
ondersteunende contacten	65%	43%
functionele contacten	45%	25%
sociale contacten	50%	48%
taakinformatie	45%	54%
organisatie informatie	50%	64%
Regelproblemen-inventarisatie		
normen	56%	55%
materiaal	77%	89%
middelen	73%	84%
operaties	46%	54%
resultaat	58%	70%
omgeving	64%	77%
Welzijnsprofiel		
volledigheid	36%	48%
organiserende taken	63%	41%
niet-kortcyclische taken	92%	97%
moeilijkheid	77%	92%
autonomie	37%	25%
contactmogelijkheden	78%	63%
informatievoorziening	25%	22%

^{****}Dit aantal is de som van 5 beoordelaars voor de vier functies.

^{*****}Dit aantal is de som van 14 beoordelaars voor de eerste taakgroepfunctie en 7 beoordelaars voor de andere drie functies.

3.5 Ervaring en functie

Uit tabel 4 valt af te leiden dat functie geen verklaring geeft voor het verschil in overeenstemming met de norm. Alleen voor de niet-experts is er een hogere overeenstemming vast te stellen voor de derde (meer simpele functie). Bij de experts verschilt de overeenstemming met de norm van functie tot functie. Ook uit deze tabel valt af te leiden dat expertise niet samengaat met een grotere overeenstemming met de norm.

Tabel 4 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar, vergelijking tussen experts - niet-experts voor de verschillende functies, naargelang het niveau van beoordeling

Functie	Functiematrix		Regelproblemen-inventarisatie		Welzijnsprofiel	
	experts	niet-experts	experts	niet-experts	experts	niet-experts
Taakgroep 1	63% (n=5)	48% (n=14)	58%	64%	49%	47%
Taakgroep 2	58% (n=5)	65% (n=7)	73%	71%	52%	52%
Lijnfunctie	60% (n=5)	74% (n=7)	64%	79%	79%	68%
Naaiautomaat-bedieners	40% (n=5)	50% (n=7)	53%	71%	54%	55%

3.6 Functie en variabele

In tabel 5 zijn grote verschillen op te merken zowel tussen variabelen als tussen functies. Voor sommige variabelen zijn er binnen de functies verschillen in overeenstemming van 80% à 90% vast te stellen. Tussen functies blijkt er ook geen stabiliteit in de beoordelingen te zijn: grote verschillen bestaan tussen variabelen en tussen functies. Deze verschillen worden teruggevonden op de drie analyse-niveaus. Beoordelaars passen per functie, naargelang de variabele, de criteria van S71 om tot een oordeel te komen op een verschillende manier toe.

Tabel 5 Percentage overeenstemming tussen beoordelaars en norm-beoordelaar. Vergelijking tussen variabelen en functietypes

	Taakgroep 1	Taakgroep 2	Lijnfunctie	Naaiautomaat-bewaker
Functiematrix	n=19	n=12	n=12	n=12
cyclustijden	89%	100%	100%	100%
moeilijkheidsgraad	79%	92%	100%	17%
autonomie tempo	84%	50%	75%	42%
autonomie methode	47%	67%	83%	92%
autonomie volgorde	42%	58%	92%	83%
autonomie werkplek	32%	42%	58%	17%
autonomie omstandigheden	42%	42%	33%	33%
ondersteunende contacten	32%	67%	50%	50%
functionele contacten	32%	25%	8%	58%
sociale contacten	53%	42%	67%	33%
taakinformatie	16%	92%	75%	17%
organisatie informatie	74%	75%	75%	8%
Regelproblemen-inventarisatie				
normen	70%	75%	8%	67%
materiaal	85%	83%	85%	83%
middelen	75%	75%	85%	83%
operaties	20%	92%	92%	0%
resultaat	80%	33%	69%	75%
omgeving	45%	75%	92%	75%
Welzijnsprofiel				
volledigheid	45%	36%	86%	8%
organiserende taken	55%	57%	43%	46%
niet-kortcyclische taken	86%	93%	100%	100%
moeilijkheid	59%	86%	100%	100%
autonomie	9%	14%	71%	23%
contactmogelijkheden	55%	71%	71%	77%
informatievoorziening	23%	7%	36%	31%

3.7 Gebruikte criteria

Er is tenslotte een inhoudsanalyse gemaakt van de criteria die gebruikt werden voor het opstellen van de *functiematrix* en het *welzijnsprofiel*. Voor de *regelproblemen-inventarisatie* werden door de beoordelaars geen criteria opgegeven. In bijlage 1 zijn de verschillende criteria weergegeven. In tabel 6 is aangeduid hoeveel verschillende criteria voor de verschillende variabelen zijn gehanteerd.

De verschillen zijn dikwijls minimaal, maar ze zijn er toch en kunnen verantwoordelijk zijn voor een verschillend eindoordeel. Er is geen verband vastgesteld tussen het aantal verschillende criteria dat is geformuleerd en het percentage overeenstemming. In een aantal gevallen zijn verschillende criteria geformuleerd en leidt dit tot een hoge overeenstemming, soms leidt 1 criterium tot een lage overeenstemming.

Tabel 6 Aantal criteria gebruikt voor de verschillende variabelen in de functiematrix en in het welzijnsprofiel

	aantal verschillende criteria
Functiematrix	
cyclus tijden	1
moeilijkheidsgraad	3
autonomie tempo	4
autonomie methode	2
autonomie volgorde	1
autonomie werkplek	2
autonomie omstandigheden	2
ondersteunende contacten	3
functionele contacten	4
sociale contacten	3
taakinformatie	1
organisatie informatie	1
Welzijnsprofiel	
volledigheid	4
organiserende taken	4
niet-kortcyclische taken	2
moeilijkheid	4
autonomie	4
contactmogelijkheden	2
informatievoorziening	1

4. CONCLUSIE EN DISCUSSIE

Doelstelling van het onderzoek was de interbeoordelaarsbetrouwbaarheid van de WEBA-methodiek vast te stellen. Het onderzoek heeft ten aanzien van de verschillende onderzoeksvragen volgende conclusies opgeleverd:

1. Ongeveer 60% van alle beoordelingen (de drie analyseniveaus samengenomen) is in overeenstemming met de normbeoordelaar. Deze overeenstemming is zowel in de informatieverzameling als in de beoordelingsfase aan de lage kant.
2. Ervaring met de WEBA-methodiek leidt niet tot een hogere overeenstemming tussen beoordelaars en de normbeoordelaar.
3. Er is geen groot verband tussen het percentage overeenstemming en de complexiteit binnen een te beoordelen functie: slechts één functie, waarvan werd verondersteld de meest simpele te zijn, wordt iets consistentter met de norm beoordeeld dan de veronderstelde complexere taakgroepfuncties. De andere (verondersteld) simpele functie (de naaiautomaat-bewaker) blijkt niet consistentter met de norm beoordeeld te worden dan de meer complexere taakgroepfuncties.
4. Wel van belang om verschillen tussen percentages overeenstemming te verklaren is de variabele of welzijnsconditie die moet worden beoordeeld: voor sommige variabelen of condities (bijvoorbeeld kortcyclische taken) is er een grote overeenstemming in de onderzoekspopulatie te vinden, voor vele andere variabelen of condities is er een lage overeenstemming aan te treffen. De overeenstemming op het niveau van de variabelen of condities is niet altijd even stabiel als het type functie in de analyse wordt betrokken: dan blijkt er per variabele of conditie ook een behoorlijke afwijking te bestaan.

De voornaamste conclusie is dat de interbeoordelaarsbetrouwbaarheid van de WEBA-methodiek op grond van dit onderzoek laag moet worden genoemd. Dit is voornamelijk te wijten aan het inconsistent toepassen van de verschillende criteria die zijn geformuleerd om een beoordeling te maken. De criteria laten teveel ruimte in de beoordeling toe. Dit betekent in de praktijk dat men bij het beoordelen van een functie aan de hand van WEBA een groot risico loopt ten onrechte welzijnsrisico's te identificeren waar deze er niet zijn of juist geen welzijnsrisico's te onderkennen waar ze er wel zijn. Geen enkele van de variabelen of condities haalt een redelijk niveau van interbeoordelaarsbetrouwbaarheid als gekeken wordt naar de Kappa-maat.

Dit resultaat is slechter dan werd vastgesteld door Grosfeld (1990) ten aanzien van de ASA-methodiek. De conclusies van dit ASA-onderzoek hebben evenwel een minder verregaande

reikwijdte dan de conclusies van dit onderzoek. Grosfeld formuleerde aan het eind van zijn onderzoek het advies om de objectiviteit en betrouwbaarheid van de methodiek te verbeteren door de instructies aan te scherpen en de training te standaardiseren en te uniformiseren. Daarom is het opvallend dat in het nieuwe onderzoek vastgesteld wordt dat ervaring (als indicatie voor meer training) geen belangrijke factor is voor een hogere overeenstemming. De achtergrond van de experts zou voor de lage overeenstemming verantwoordelijk kunnen zijn: deze experts hebben enerzijds een NIPG-achtergrond en anderzijds een Arbeidsinspectie-achtergrond. De Arbeidsinspectie heeft meer de gewoonte om met het Conceptvoorlichtingsblad 22 over de WEBA-methodiek te werken, de NIPG-experts meer met S-71. In CV 22 wordt een aantal criteria van S-71 versimpeld voorgesteld. Hoewel de Arbeidsinspecteurs gewerkt hebben met S-71, kan deze andere ervaringsachtergrond deels een verklaring vormen voor het gevonden verschil. De overeenstemming tussen NIPG-onderzoekers is evenwel ook niet hoog. Een andere verklaring kan zijn dat de verschillende experts allen op een andere wijze de WEBA-methodiek hebben aangeleerd (zelfstudie, on-the-job training) zodat dit leidt tot eigen interpretaties van de verschillende criteria. Meer gestandaardiseerde trainingen zouden een grotere overeenstemming in de hand kunnen werken. Een precisering van de criteria in de handleiding lijkt evenwel het meest aangewezen. Algemeen kan worden gezegd dat de meeste criteria in de basishandleiding S-71 onvoldoende zijn uitgewerkt opdat de beoordelaars tot een consistent oordeel zouden kunnen komen. De beperkte inhoudsanalyse heeft dit aan het licht gebracht.

Het onderzoek leidt tot volgende constatering.

- In eerste instantie zou het kunnen zijn dat het videomateriaal onvoldoende is om de beoordeling goed uit te voeren. Dit lijkt echter onwaarschijnlijk: de meeste beoordelaars gaven aan dat het materiaal voldoende was voor de beoordeling. Daarnaast is het zo dat het materiaal dermate gestandaardiseerd is dat de beoordelingen toch steeds een zelfde oordeel zouden moeten opleveren. Realiteitsmetingen zullen nooit de gestandaardiseerde informatie van video-beelden kunnen evenaren. De beelden werden zo samengesteld dat ze volgens een WEBA-expert een volledige beoordeling konden toelaten.
- Een ander probleem heeft te maken met het verschil tussen het hoge percentage overeenstemming voor bepaalde variabelen en een lage Kappa. Dit wijst op gebreken in het design dat is gehanteerd: er bestaat onvoldoende onderscheid op de verschillende variabelen opdat een uitspraak mogelijk zou zijn over interbeoordelaarsbetrouwbaarheid. Echter, dit heeft minder te maken met het design dan met het magere resultaat van functieverbetering. De functie van naaister werd geselecteerd omdat juist hierin een mooi voorbeeld van functieverbetering kon worden getoetst. Uit het onderzoek kan de ver-

betering niet in alle opzichten worden afgeleid: voor die variabelen waar een hoge overeenstemming met de norm werd gevonden, wees de overeenstemming in de zin van 'onvoldoende verbetering'. In vervolgonderzoek zal men aandacht moeten besteden aan de mate waarin functies op de verschillende variabelen of condities verbeterd zijn. Dit experiment toont aan dat er juist behoefte is aan degelijke testen op wat eigenlijk wordt verbeterd.

- Een laatste probleem is dat de criteria om de WEBA-methodiek te beoordelen misschien te streng zijn: in de meeste interbeoordelaarsonderzoeken (bijvoorbeeld ten aanzien van de JDS, zie Jenkins e.a., 1975) wordt gewerkt met gewogen maten, in dit onderzoek is dit niet gebeurd. Weging lijkt evenwel niet aangewezen omdat de meeste variabelen in de WEBA-methodiek slechts drie waarden kunnen aannemen: weging zou onmiddellijk leiden tot een hoog percentage overeenstemming zonder dat die overeenstemming er echt is.

LITERATUUR

CHRISTIS J, FORTUIN R. ASA-Handleiding. Amsterdam: NIA, 1989.

COHEN J. Weighted Kappa: nominal scale agreement with provision of scaled disagreement or partial credit. Psychol Bull 1968;70:213-20.

CORNELIUS ET, DENISI AS, BLENCOE AG. Expert and naive raters using the PAQ: Does it matter? Personnel Psychol 1984;37(3):453-64.

DENISI AS, CORNELIUS ET, BLENCOE AG. Further investigation of common knowledge effects on job analysis ratings. J Appl Psychol 1987;72(2):262-8.

DHONDT S, HOUTMAN I. NIPG-Onderzoeksvragenlijst Arbeidsinhoud: constructie en eerste toets op betrouwbaarheid en validiteit. Leiden: NIPG-TNO, 1992. Publ.nr. 92.088.

FLEISS J, COHEN J, EVERITT B. Large sample standard error for Kappa and weighted Kappa. Psychol Bull 1969;72:323-7.

FUNCTIE-INHOUD ANALYSEREN EN BEOORDELEN: DE WEBB-METHODE. Den Haag: Directoraat-Generaal van de Arbeid, 1991. (CV 22)

GOODWIN LD, SANDS DJ, KOZLESKI EB. Estimating interinterviewer reliability for interview schedules used in special education research. J Special Educ 1991;25(1):73-89.

GROSFELD JAM et al. De Arbeidssituatie-Analyse: eerste indrukken over objectiviteit en betrouwbaarheid van een experimentele methode ter vaststelling van welzijnsrisico's. Rapport B. Nijmegen: Katholieke Universiteit, Psychologisch Laboratorium, augustus 1990.

GUPTA N, BEEHR TA. A test of correspondence between self-reports and alternative data sources about work organizations. J Vocational Behav 1982;20:1-13.

HARTMANN DP. Considerations in the choice of interobserver reliability estimates. J Appl Behav Anal 1977;10:103-16.

HOUTMAN I. Onderzoeksnotitie: kwantitatief onderzoek welzijnsrisico's, gezondheid en welbevinden. Leiden: NIPG-TNO, 1992.

JENKINS GD jr, NADLER DA, LAWLER EE III, CAMMANN C. Standardized observations: an approach to measuring the nature of jobs. *J Appl Psychol* 1975;60:171-81.

LANDIS JR, KOCH GG. The measurement of observer agreement for categorical data. in: *Biometrics* 1977;22:154-74.

PEETERS MHH, POT FD, DELLEMAN N. Nieuwe produktiemethoden voor de kledingindustrie: een verkennend onderzoek naar de mogelijkheden van integratie van het aspect kwaliteit van de arbeid in een integraal (her)ontwerp. Den Haag: Directoraat-Generaal van de Arbeid, 1991. S 117.

PEETERS MHH. Groepswerk in de confectie: beschrijving en analyse van vier bedrijven met een integraal aangepaste productie- en arbeidsorganisatie. Den Haag: Directoraat-Generaal van de Arbeid, 1992. S 156.

PROJECTGROEP WEBA. Functieverbetering en organisatie van de arbeid. Den Haag: Directoraat-Generaal van de Arbeid, 1989. S 71.

SIEGEL S, CASTELLAN J jr. *Nonparametric statistics for the behavioral sciences*. 2nd ed. New York: McGraw-Hill, 1988.

WORKUM R van. De Arbeidssituatie-Analyse van het Nederlands Instituut voor Arbeidsomstandigheden te Amsterdam: meetinstrument of discussiestuk?. Rapport A. Nijmegen: Katholieke Universiteit, Psychologisch Laboratorium, 1990.

BIJLAGEN

pagina

BIJLAGE 1	Welke criteria werden gebruikt om de functiematrix en het welzijnsprofiel op te stellen?	29
BIJLAGE 2	Resultaat van de beoordelingen op het niveau van de functiematrix, regelproblemen-inventarisatie en het welzijnsprofiel.	37

BIJLAGE 1

**Welke criteria werden gebruikt om de functiematrix en het
welzijnsprofiel op te stellen?**

Welke criteria werden gebruikt om de functiematrix en het welzijnsprofiel op te stellen?

A. Functiematrix

1. Cyclustijd

Geen verschil in criteria.

2. Moeilijkheidsgraad

De meeste beoordelaars gebruiken gelijksoortige criteria, alleen de mate van precisie verschilt danig. Volgende criteria werden geformuleerd:

- *uitgaan van wat door taakuitoefenaar wordt gezegd;*
- *uitgaan van 'indruk dat uitvoerende taken snel te leren zijn';*
- *is er sprake van 'denken vooraf (=+)', 'aandacht tijdens werk (=0) en 'routinewerk (= -)'?*

3. Autonomie tempo

De criteria verschillen tussen beoordelaars naar de mate van precisie. Om de autonomie over tempo te bepalen worden volgende criteria gebruikt:

- *blijkt uit de beelden of er tempo-vrijheid is in de uitvoerende taken?*
- *het aantal stuks dat moet worden geproduceerd is het criterium;*
- *kan men binnen dag- of weekplanning zelf variëren?*
- *zijn er buffers aanwezig of niet?*

4. Autonomie methode

De meeste beoordelaars gebruiken dezelfde criteria, maar toch zijn hier ook verschillen vast te stellen:

- *kan men zelf de methode uitzoeken?*
- *worden alle handelingen aan één produkt gedaan: dit houdt volgens een beoordelaar in dat er geen methodevrijheid is.*

5. Autonomie volgorde

Geen verschillen opgemerkt.

6. Autonomie werkplek

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *heeft men verschillende werkplekken of niet?*
- *is men gebonden aan machines of niet?*

7. Autonomie arbeidsomstandigheden

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *kan men arbeidsomstandigheden beïnvloeden?*
- *kan men zelf wijzigingen ter verbetering aandragen?*

8. Ondersteuningsmogelijkheden

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *kan men hulp inroepen als het nodig is?*
- *gebruikt men hulp van anderen of niet?*
- *kan men de productie ongestoord laten verlopen of niet?*

9. Functionele contacten

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *men moet eigen interpretatie van beelden volgen;*
- *komen de functionele contacten voor en zo ja met wie?*
- *kan men hulp inroepen met betrekking tot tempo etc of niet?*
- *heeft men contacten met chef, collega's, andere afdelingen?*

10. Sociale contacten

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *men moet eigen interpretatie van beelden volgen;*
- *wordt er in groepsverband gewerkt met beperkte voorwaarden?*
- *heeft men de mogelijkheid om tijdens het werk contacten te leggen?*

11. Informatie taakniveau

De meeste gebruiken dezelfde criteria. Volgens sommige beoordelaars zijn er voor deze variabele in S71 geen criteria gegeven. Het meest gebruikte criterium was:

- *hoe zit het met de werkinstructie, afdelings- bedrijfsinformatie?*

12. Informatie organisatieniveau

De beelden waren voor deze variabele onvoldoende volgens de meeste beoordelaars.

B. Welzijnsprofiel

1. Volledigheid

De meeste gebruiken dezelfde criteria. Toch zien we een vermenging van criteria die eigenlijk voor andere variabelen (moeilijkheid) dienen. Volgende verschillende criteria werden gebruikt:

- *wat is de omvang van de uitvoerende taken? Zijn er voorbereidende en ondersteunende taken aanwezig; dan zijn de leermogelijkheden positief;*
- *als technische taken (onderhoud etc) ontbreken dan is het oordeel negatief;*
- *is sprake van voldoende afwisseling, een voldoende moeilijkheidsgraad, dan is het oordeel positief;*
- *is het werk een samenhangend geheel?*

2. Organiserende taken

De meeste gebruiken hetzelfde criterium: is er werkoverleg aanwezig? Volgende verschillende criteria werden gebruikt:

- *is er sprake van groepswerk: dan betekent dit dat er veel onderling geregeld dient te worden wat het stressaspect bepaalt. Éénmaal per week overleg is een indicatie van organiserende taken;*
- *worden veel problemen opgelost dan wordt stress gereduceerd; leren wordt bepaald door de aanwezigheid van mogelijkheden tot overleg;*
- *worden regelproblemen extern opgelost?*
- *regelproblemen moeten worden gerelateerd aan oplossingsmogelijkheden.*

3. Niet-kortcyclische taken

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *meer dan 50% van de tijd besteed aan cycli van minder dan 90 seconden;*
- *meer dan 40% van de taken minder dan 90 seconden.*

4. Moeilijkheid

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *kijk naar de inleertijd;*
- *is de inleertijd maximum een week, dan is het niet moeilijk;*
- *de meeste taken zijn zonder nadenken uit te voeren, het langdurig gemakkelijke taken uitvoeren waardoor men niet naar eigen behoefte kan wisselen leidt tot een negatief oordeel;*
- *moet men 'voor' - 'tijdens' - 'niet' denken bij de taakuitvoering; evenwichtige verdeling voor deze drie criteria is nodig om van leren te kunnen spreken.*

5. Autonomie

De meeste criteria komen met elkaar overeen, maar men verwijst toch naar verschillende dingen.

- *is er keuzevrijheid; worden veel problemen zelf opgelost: deze twee criteria bepalen het stressaspect;*

- *stress wordt gereduceerd als men veel zelf moet uitzoeken;*
- *autonomie over methode is van belang voor leren;*
- *men moet letten op het oplossend vermogen in de functie, wat belangrijk is om stress te reduceren; autonomie ten aanzien van tempo en werkplek zijn ook van belang;*
- *als er weeklijsten zijn, dan leggen deze tempo en volgorde vast; men dient ook na te gaan of de taak de werknemer aan de werkplek bindt;*
- *laat de weekplanning toe om keuzen te maken, dan is er voldoende autonomie.*

6. Contactmogelijkheden

De meeste gebruiken dezelfde criteria. Volgende verschillende criteria werden gebruikt:

- *groepswork, ondanks de belemmeringen door herrie en produktiedraaien, laat toe, als het moet, met elkaar te spreken;*
- *de ervaring van de respondent over contactmogelijkheden is van belang.*

7. Informatievoorziening

De meeste beoordelaars gebruikten volgend criterium:

- *heeft men problemen met beschikbare informatie.*

BIJLAGE 2

**Resultaat van de beoordelingen op het niveau van de functiematrix,
regelproblemen-inventarisatie en het welzijnsprofiel**

Resultaat van de beoordelingen op het niveau van de functiematrix, regelproblemen-inventarisatie en het welzijnsprofiel.

In onderstaande tabel is de verdeling weergegeven van de beoordelaars naar de verschillende variabelen of condities en naar de verschillende functies. Het gaat hier om die beoordelaars die alle functies hebben beoordeeld.

a) *Functiematrix (n=12)*

1. CYCLUSDUUR				
	?	-	o	+
Functie 1	0	11	1	0
Functie 2	0	12	0	0
Functie 3	0	12	0	0
Functie 4	0	12	0	0
2. MOEILIKHEID				
	?	-	o	+
Functie 1	0	9	3	0
Functie 2	0	11	1	0
Functie 3	0	12	0	0
Functie 4	0	9	3	0
3. AUTONOMIE TEMPO				
	?	-	o	+
Functie 1	0	0	9	3
Functie 2	0	2	6	4
Functie 3	0	2	9	1
Functie 4	0	5	5	2
4. AUTONOMIE METHODE				
	?	-	o	+
Functie 1	0	3	9	0
Functie 2	0	3	9	0
Functie 3	0	10	2	0
Functie 4	0	11	1	0

5. AUTONOMIE VOLGORDE				
	?	-	o	+
Functie 1	1	1	3	7
Functie 2	0	2	3	7
Functie 3	0	11	0	1
Functie 4	1	10	0	1
6. AUTONOMIE WERKPLEK				
	?	-	o	+
Functie 1	2	1	6	3
Functie 2	1	1	6	4
Functie 3	1	7	2	2
Functie 4	0	7	3	2
7. AUTONOMIE ARBEIDSONDERSTANDIGHEDEN				
	?	-	o	+
Functie 1	4	6	2	0
Functie 2	3	5	4	0
Functie 3	4	5	3	0
Functie 4	4	6	2	0
8. CONTACTMOGELIJKHEDEN ONDERSTEUNING				
	?	-	o	+
Functie 1	1	1	3	7
Functie 2	0	0	3	9
Functie 3	0	1	7	4
Functie 4	0	3	7	2
9. CONTACTMOGELIJKHEDEN FUNCTIONEEL				
	?	-	o	+
Functie 1	0	1	3	8
Functie 2	0	1	4	7
Functie 3	0	2	7	3
Functie 4	0	0	7	5
10. CONTACTMOGELIJKHEDEN SOCIAAL				
	?	-	o	+
Functie 1	0	0	8	4
Functie 2	0	0	6	6
Functie 3	0	2	8	2
Functie 4	2	5	4	1

11. INFORMATIEVOORZIENING TAAK				
	?	-	o	+
Functie 1	1	0	3	8
Functie 2	1	0	0	11
Functie 3	1	1	1	9
Functie 4	0	0	3	9
12. INFORMATIEVOORZIENING ORGANISATIE				
	?	-	o	+
Functie 1	10	0	0	2
Functie 2	10	0	0	2
Functie 3	9	1	1	1
Functie 4	6	0	2	4

b) Regelproblemen (n=12)

1. NORMEN		
	GEEN	WEL
Functie 1	2	10
Functie 2	2	10
Functie 3	10	2
Functie 4	8	4
2. MATERIAAL		
	GEEN	WEL
Functie 1	1	11
Functie 2	1	11
Functie 3	2	10
Functie 4	2	10
3. MIDDELEN		
	GEEN	WEL
Functie 1	2	10
Functie 2	10	2
Functie 3	10	2
Functie 4	2	10
4. OPERATIES		
	GEEN	WEL
Functie 1	8	4
Functie 2	12	0
Functie 3	11	1
Functie 4	11	1
5. RESULTAAT		
	GEEN	WEL
Functie 1	9	3
Functie 2	8	4
Functie 3	8	4
Functie 4	9	3
6. OMSTANDIGHEDEN		
	GEEN	WEL
Functie 1	6	6
Functie 2	10	2
Functie 3	11	1
Functie 4	9	3

c) Welzijnsprofiel (n=13)

1. VOLLEDIGHEID				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	1	6	6
Functie 2	0	2	6	5
Functie 3	0	12	1	0
Functie 4	0	11	1	1
2. ORGANISERENDE TAKEN				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	0	4	9
Functie 2	0	0	5	8
Functie 3	0	7	5	1
Functie 4	0	7	5	1
3. KORTCYCLISCHE TAKEN				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	11	2	0
Functie 2	0	12	1	0
Functie 3	0	13	0	0
Functie 4	0	13	0	0
4. MOEILIKHEID				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	9	4	0
Functie 2	0	12	1	0
Functie 3	0	13	0	0
Functie 4	0	13	0	0
5. AUTONOMIE				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	0	10	3
Functie 2	0	2	8	3
Functie 3	0	11	1	1
Functie 4	0	9	3	1

6. CONTACTMOGELIJKHEDEN				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	0	0	3	10
Functie 2	0	0	3	10
Functie 3	0	1	10	2
Functie 4	0	1	10	2
7. INFORMATIEVOORZIENING				
	?	Onvoldoende	Beperkt	Voldoende
Functie 1	2	0	2	9
Functie 2	2	0	1	10
Functie 3	1	2	4	6
Functie 4	1	0	5	7

Reprografie: NIPG-TNO
Projectnummer: 5741