

TNO Kwaliteit van Leven

Preventie en Zorg
Wassenaarseweg 56
Postbus 2215
2301 CE Leiden

www.tno.nl

T 071 518 18 18
F 071 518 19 10
info-zorg@tno.nl

TNO-rapport

KvL/B&G 2007.138

Productiviteit Meten met een Vragenlijst

Stand van zaken november 2007

Datum	November 2007
Auteur(s)	Claire Bernaards Ingrid Hendriksen
Oprachtgever	Ministerie VWS
Projectnummer	031.11479/01.04
Aantal pagina's	32 (incl. bijlagen)
Aantal bijlagen	1

Alle rechten voorbehouden. Niets uit dit rapport mag worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook, zonder voorafgaande schriftelijke 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 onderzoekopdrachten aan TNO, dan wel de betreffende terzake tussen de partijen gesloten overeenkomst.

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

© 2007 TNO

Samenvatting

Inleiding

De meeste vragenlijsten die productiviteit meten richten zich op verminderde productiviteit t.g.v. ziekte. TNO Kwaliteit van Leven heeft echter juist behoefte aan een vragenlijst die productiviteit meet in een algemene werknemerspopulatie, bestaande uit voornamelijk gezonde werknemers. Het uiteindelijke doel van TNO is om met deze vragenlijst de relatie tussen arbeidsproductiviteit en leefstijlfactoren te kwantificeren. Productiviteitsverlies kan optreden als men verzuimt, maar ook door een verminderde werkprestatie. TNO richt zich in dit onderzoek alleen op het meten van verminderde werkprestatie. In dit rapport wordt een samenvatting gegeven van een stageverslag waarin de volgende vraagstellingen zijn onderzocht: 1. Welke vragenlijsten zijn bruikbaar om werkprestatie te meten bij gezonde werknemers? 2. Hoe hangen de drie meest kansrijke vragenlijsten met elkaar samen? 3. Hoe komen werknemers tot een antwoord op deze vragenlijsten? 4. Is de spreiding bij vragenlijsten met abstracte vragen lager dan bij vragenlijsten met concrete vragen? Vervolgens worden de belangrijkste bevindingen en criteria voor een nieuw te ontwikkelen vragenlijst op een rij gezet. Daarnaast wordt een model gepresenteerd waarin de relatie tussen bewegen en productiviteit in kaart is gebracht. Tot slot worden de geplande vervolgtacties van TNO op dit gebied op een rij gezet.

Methode

Bij aanvang van de stage werd op basis van de literatuur (Pubmed en Psychlit) een overzicht gemaakt van alle vragenlijsten die aangaven productiviteit te meten. Vervolgens werden de drie meest kansrijke vragenlijsten voor het meten van werkprestatie bij gezonde werknemers geselecteerd en uitgezet bij een groep werknemers (N=94). Zij kregen een samengestelde vragenlijst die bestond uit de uitgekozen drie vragenlijsten, waarbij de Latin Square methode werd gebruikt om het effect van de vragenlijstvolgorde tegen te gaan. Met behulp van hardop-denken protocollen werd bij een subgroep van de deelnemers (N=6) nagegaan hoe zij tot een antwoord kwamen op de vragenlijsten. Om de samenhang tussen de vragenlijsten te bepalen werd gebruik gemaakt van Pearson correlatie coëfficiënten, gecorrigeerd voor de onbetrouwbaarheid van de vragenlijsten. Om de spreiding op abstracte vragenlijsten te vergelijken met de spreiding op concrete vragenlijsten werden de varianties van de vragenlijsten met elkaar vergeleken voor ieder item en voor de hele vragenlijst. De resultaten van beide methoden werden met elkaar vergeleken om na te gaan hoe robuust de resultaten waren.

Resultaten

De Health and Performance Questionnaire (HPQ), Endicott Work and Productivity Scale (EWPS) en Work Limitations Questionnaire (WLQ) werden geselecteerd als de drie meest kansrijke instrumenten voor het meten van werkprestatie bij gezonde werknemers. De correlatie tussen de drie vragenlijsten was matig tot groot voor de hele onderzoekspopulatie. De spreiding op de vragenlijst met voornamelijk abstracte vragen (HPQ) was lager dan die van de vragenlijsten met concrete vragen (EWPS/WLQ), zowel op itemniveau als op het niveau van de hele vragenlijst. De resultaten uit de hardop-denken protocollen ondersteunden deze conclusie. De drie vragenlijsten hebben ieder hun eigen beperkingen. Het nadeel van de HPQ is dat deze vragenlijst abstracte vragen bevat die gevoelig lijken te zijn voor het geven van sociaal wenselijke antwoorden. Het nadeel van de EWPS is dat deze vragenlijst samengestelde vragen

bevat. De WLQ heeft als nadeel dat deze vragenlijst alleen prestatievermindering ten gevolge van gezondheidsklachten meet en hiermee zonder aanpassingen niet inzetbaar is in een populatie van (voornamelijk) gezonde werknemers. Daarnaast bevatten zowel de EWPS als de WLQ concrete vragen die niet gemakkelijk te vertalen zijn naar kosten ten gevolge van productiviteitsverlies.

Conclusie

Geen van de drie onderzochte vragenlijst is in de huidige vorm inzetbaar bij projecten waar de werkprestatie gemeten moet worden van gezonde werknemers. Het ligt op basis van de bevindingen uit deze studie voor de hand om een nieuwe vragenlijst te ontwikkelen. Deze nieuw te ontwikkelen vragenlijst zou bij voorkeur zowel concrete als abstracte vragen moeten bevatten, onder meer omdat concrete vragenlijsten moeilijk te vertalen zijn naar productiviteitsverliezen. Ook dient rekening te worden gehouden met de gevonden nadelen en belemmeringen van de drie bestaande vragenlijsten.

Summary

Introduction

Several questionnaires have been developed to assess productivity but many of these questionnaires focussed on reduced productivity as a result of health problems. However, the goal of TNO Quality of Life is to assess productivity in a relatively healthy working population in order to quantify the relationship between productivity and lifestyle factors. The focus of this project was on work performance which is, besides absenteeism, one of the two main components of productivity. The goal of this study was to answer the following questions: 1. Which questionnaires are available to assess work performance in relatively healthy workers? 2. What is the correlation between three selected questionnaires? 3. How do workers come to a certain answer on the questionnaires? 4. Do abstract questionnaires give rise to smaller variances in work performance than concrete questionnaires? This report presents the main results of this study and the most important findings and criteria for the development of a new instrument to assess work performance in a healthy working population. Furthermore, a model is presented that describes the relationship between physical activity and productivity. Finally, future plans of TNO with regard to this topic will be presented.

Methods

In order to answer the first question, the literature (Pubmed and Psychlit) was searched and three questionnaires that were most suited to be used in our future studies were selected. For the second question, 94 workers filled out all three selected questionnaires. The Latin Square method was used in order to study the order effect. To investigate the correlation between questionnaires, Pearson correlation coefficients were used with correction for the reliability of questionnaires. To answer the third question, the Think Aloud protocol was used; six workers were asked to think aloud while filling out the three questionnaires. To answer the fourth question variances on single items and complete questionnaires were compared between questionnaires with mostly abstract questions and questionnaires with concrete questions.

Results

The Health and Performance Questionnaire (HPQ), Endicott Work and Productivity Scale (EWPS) and the Work Limitations Questionnaire (WLQ) were selected for this study. Correlations between pairs of questionnaires were moderate to large for the complete study population. The abstract questionnaire gave rise to lower variances (HPQ) than the concrete questionnaires (EWPS and WLQ), both on item level as well as on the level of total scores. Results of the think aloud protocol supported these findings.

All three questionnaires have their limitations. The disadvantage of the HPQ is the high chance of yielding social desirable answers due to the abstract questions. The main disadvantage of the EWPS is its use of compound questions. The use of the WLQ is limited because it is explicitly focussed on reduced work performance resulting from health problems. Without changes this questionnaire cannot be used in a population of relatively healthy workers. In addition, the results of both abstract questionnaires (the EWPS and the WLQ) are difficult to translate into costs due to productivity losses.

Conclusions

In their original form, none of the three selected questionnaires can be used to quantify the relationship between work performance and lifestyle factors in healthy workers. Therefore, it seems inevitable to develop a new questionnaire. This new questionnaire should, preferably, consist of both abstract and concrete questions because concrete questions are difficult to translate into productivity losses. Also all disadvantages of the three existing questionnaires should be taken into account.

Inhoudsopgave

	Samenvatting	2
	Summary	4
1	Inleiding	7
2	Samenvatting stageverslag	9
2.1	Inleiding.....	9
2.2	Methode.....	10
2.3	Resultaten	12
2.4	Discussie.....	18
2.5	Conclusie	19
3	Belangrijkste bevindingen en vervolgstappen	20
3.1	Belangrijkste bevindingen	20
3.2	Criteria voor een nieuwe vragenlijst.....	22
3.3	Model bewegen en productiviteit.....	23
3.4	Stappenplan voor de komende jaren.....	28
4	Referenties	29
	Bijlage(n)	
	A Hardop-denken protocollen	

1 Inleiding

Om de kosten en baten van bedrijfsbewegingsprogramma's in kaart te brengen, zal naast verzuim ook productiviteit gemeten moeten worden. Immers, indien een werknemer niet verzuimt, kan verminderde werkprestatie toch tot productiviteitsverliezen leiden. Onder productiviteitsverliezen verstaan we in dit geval de verliezen ten gevolge van verminderde werkprestatie van medewerkers op het moment dat zij niet verzuimen. Dit kan het gevolg zijn van een minder goede gezondheid maar ook van factoren zoals verminderd concentratievermogen, gebrek aan inspiratie/motivatie of factoren in de bedrijfsorganisatie. Een bedrijfsbewegingsprogramma kan naast een vermindering in verzuim mogelijk ook bijdragen aan een verhoging van de werkprestatie en daarmee minder productiviteitsverlies opleveren.

Voor het meten van arbeidsproductiviteit zijn er in het afgelopen decennium diverse vragenlijsten ontwikkeld (o.a. de Health and performance Questionnaire, Work Limitations Questionnaire en Prodisq). De meeste van deze vragenlijsten hebben als primair doel om productiviteitsverliezen in kaart te brengen van werknemers met gezondheidsklachten of gezondheidsrisico's. Het verschijnsel van verminderde werkprestatie bij werknemers met gezondheidsklachten die niet verzuimen wordt ook wel presentisme genoemd.

Bij TNO hebben we echter juist behoefte aan een vragenlijst die productiviteit meet in een algemene werknemerspopulatie die voornamelijk bestaat uit gezonde werknemers en slechts gedeeltelijk uit werknemers met gezondheidsklachten. Het uiteindelijke doel van TNO is om met deze vragenlijst de relatie tussen arbeidsproductiviteit en leefstijlfactoren te kwantificeren. Voor dit doel vallen alle vragenlijsten af die alleen specifiek vragen naar verminderde werkprestatie ten gevolge van gezondheidsklachten. Daarnaast heeft TNO nog een aantal criteria opgesteld waaraan de vragenlijst zou moeten voldoen (o.a. wat betreft validiteit, betrouwbaarheid, gevoeligheid voor verandering). Een eerste inventarisatie van bestaande vragenlijsten om productiviteit te meten leverde de conclusie op dat geen van de bestaande vragenlijsten volledig voldeed aan de TNO criteria. De meeste vragenlijsten waren niet gevalideerd in een gezonde werknemerspopulatie en zeker niet bij kantoorpersoneel ('white collar workers'). Er zijn wel drie vragenlijsten geselecteerd die als kansrijk werden beschouwd, namelijk de Health and Performance Questionnaire (HPQ) (Kessler et al 2003), de Endicott Work and Productivity Scale (EWPS) (Endicott & Nee 1997) en de Health Risk Appraisal (AHM-HRA) (Musich et al 2006). Deze zijn voorgelegd aan een aantal experts (binnen en buiten TNO). Geen van de vragenlijsten werd door de experts als 'de beste' gezien. Een aantal experts had twijfel over de mate waarin de EWPS arbeidsproductiviteit meet. Deze vragenlijst zou alleen indirect arbeidsproductiviteit meten. De HPQ werd vaak als 'aardig' beschouwd maar sommige experts vonden deze vragenlijst te abstract of te gevoelig voor sociaal wenselijke antwoorden. Eén expert verwachtte plafondeffecten bij deze vragenlijst. De categorieën die gebruikt worden bij de AHM-HRA spraken over het algemeen wel tot de verbeelding. Een aantal experts vroegen zich echter af of deze vragenlijst wel het hele spectrum van productiviteit meet en of hij gevoelig genoeg is om kleine effecten te meten.

Recentelijk heeft een stagiair opnieuw geïnventariseerd welke vragenlijsten er voor het meten van productiviteit beschikbaar zijn. Op basis van de literatuur werden drie vragenlijsten geselecteerd die het meest kansrijk leken om te gebruiken in onderzoek naar het effect van bewegen/gezonde leefstijl op productiviteitsverandering bij gezonde werknemers. Dit waren de HPQ, de EWPS en de Work Limitations Questionnaire (WLQ) (Lerner et al 1998, 2001). De AHM-HRA is in het stageonderzoek vervangen door de WLQ omdat de spreiding op de AHM-HRA naar verwachting laag zou zijn. Ondanks dat de WLQ specifiek vraagt naar productiviteitsverliezen ten gevolge van lichamelijke klachten, en hij hiermee dus eigenlijk niet geschikt is voor ons

onderzoeksdoel, hebben we hem wel meegenomen in het onderzoek vanwege twee redenen. Ten eerste denken we dat deze vragenlijst na aanpassing heel geschikt kan zijn voor het meten van werkprestatie bij gezonde werknemers. Ten tweede wilden we nagaan hoe gezonde werknemers omgaan met specifieke vragen naar productiviteitsverliezen ten gevolge van gezondheidsklachten. Door één abstracte vragenlijst (HPQ) mee te nemen en twee concrete vragenlijsten (EWPS/WLQ) konden we bovendien het effect van concrete en abstracte vragenlijsten bestuderen op de spreiding van de vragenlijsten. De drie vragenlijsten zijn uitgezet onder een groep kantoormedewerkers om inzicht te krijgen in wat deze vragenlijsten nu eigenlijk meten en hoe ze met elkaar samenhangen.

In dit rapport wordt een samenvatting gegeven van het bovengenoemde stageverslag (hoofdstuk 2) en worden vervolgens de belangrijkste bevindingen en criteria voor een nieuw te ontwikkelen vragenlijst op een rij gezet (hoofdstuk 3). Omdat het uiteindelijke doel van TNO is om de relatie tussen bewegen en productiviteit te kwantificeren, is een model opgesteld waarin deze relatie in kaart is gebracht (hoofdstuk 3). Dit model wordt eveneens in hoofdstuk 3 weergegeven en nader toegelicht. Bij de geplande vervolgacties van TNO op dit gebied zal dit model als basis dienen.

2 Samenvatting stageverslag

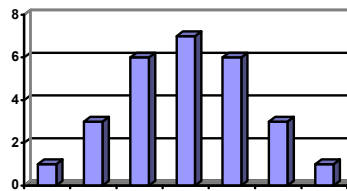
2.1 Inleiding

De stage bestond uit een literatuurstudie, een kwantitatief deel en een kwalitatief deel.

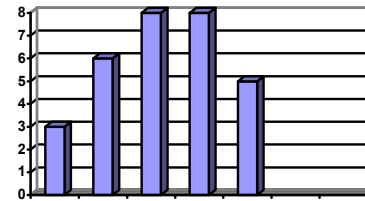
De onderzoeksvragen binnen de literatuurstudie waren als volgt:

- a) Welke instrumenten zijn er om productiviteit te meten?
- b) In hoeverre meten deze instrumenten werkprestatie?
- c) Hoe wordt werkprestatie door de makers van deze instrumenten gedefinieerd?
- d) Wat is de beoogde doelgroep? (bijv. specifieke beroepsgroep, werknemers met specifieke gezondheidsproblemen, algemene werknemerspopulatie).
- e) In hoeverre zijn deze instrumenten ingezet binnen studies? Wat was binnen deze studies de spreiding in werkprestatie? In het geval van interventiestudies, was er sprake van verandering in werkprestatie?

In de kwantitatieve studie zijn drie vragenlijsten uitgezet bij werknemers uit verschillende bedrijven. Het eerste doel van de kwantitatieve studie was om de samenhang tussen de vragenlijsten te bestuderen. Er is tevens gekeken naar de mate van overeenstemming in de berekende percentages productiviteitsverlies. Het tweede doel van het kwantitatieve onderzoek was om na te gaan of concrete vragenlijsten die productiviteit beogen te meten meer last hebben van “impression management” dan abstracte vragenlijsten. Een van de drie vragenlijsten was de Health and Performance Questionnaire (HPQ). Dit is een vragenlijst met een abstracte vraag als hoofdvraag. Uit eerder onderzoek bleek dat de spreiding op deze abstracte vraag heel laag was (Rijnaard en Jans 2006; Bernaards et al. 2007). Mogelijk kan dit het gevolg zijn van impression management. Impression management is een doelgerichte bewuste of onbewuste poging om de perceptie van andere mensen ten aanzien van een persoon, voorwerp of gebeurtenis te beïnvloeden door het reguleren en controleren van informatie tijdens sociale interactie. Het vertonen van sociaal wenselijk gedrag is hier een voorbeeld van. Omdat men bij het invullen van de vragenlijst door invloed van impression management vaker het sociaal wenselijke antwoord geeft, in dit geval dat men bijvoorbeeld zeer productief is, verschuift het gemiddelde van de scores in de richting van het meer sociaal wenselijke antwoord en neemt de variantie van de test scores af (Podsakoff & Organ, 1986). De ware scores zijn dan bijvoorbeeld verdeeld zoals in figuur 1, de test scores daarentegen zijn dan door de invloed van impression management verdeeld zoals in figuur 2.



Figuur 1. Ware score.



Figuur 2. Test scores

Wanneer er minder variantie is én deze variantie ook nog op een systematische manier is verstoord, is het lastiger een verband aan te tonen tussen de variabelen.

Tenslotte heeft een kwalitatieve studie plaatsgevonden waarin met behulp van de hardop-denken methode is nagegaan hoe werknemers tot een antwoord komen op de drie vragenlijsten en welke problemen ze hierbij ondervonden.

2.2 Methode

2.2.1 Literatuurstudie

Voor de literatuurstudie is in Psychlit en Pubmed gezocht naar meetinstrumenten die productiviteit of werkprestatie beogen te meten. De volgende kenmerken van de meetinstrumenten zijn in kaart gebracht en in een tabel geplaatst: het aantal items, de antwoordschaal (aantal antwoordmogelijkheden per item), de aanwezigheid van een neutraal antwoord zoals “niet van toepassing”, de terugvraagperiode, de aanwezigheid van concrete vragen (hierbij wordt gevraagd naar concrete taken), de aanwezigheid van abstracte vragen (hierbij wordt gevraagd naar abstracte termen zoals “uw algehele werkprestatie”), benodigde invultijd, beschikbaarheid in het Nederlands, de inzetbaarheid bij gezonde werknemers, de gevoeligheid voor geheugenproblemen, de mogelijkheid om productiviteitsverliezen te meten, de validiteit, het aantal studies waarin de vragenlijst gebruikt is, de link van items met productiviteit, eenduidigheid en de gevoeligheid voor bias.

2.2.2 Kwantitatieve studie

Deelnemers

Voor de uitvoering van de kwantitatieve studie zijn 120 medewerkers gevraagd deel te nemen aan de studie. Voorwaarde voor participatie in dit onderzoek was dat de respondenten drie volle dagen per week werkzaam waren. De respondenten zijn geworven door “White collar workers” uit de (privé) omgeving van de stagiair te vragen te participeren in dit onderzoek. Vervolgens is aan de respondenten gevraagd om collega’s en kennissen te vragen om te participeren in dit onderzoek. Dit is ook wel bekend als de ‘sneeuwbal methode’. Doordat er gebruik is gemaakt van de ‘sneeuwbal methode’ is het response percentage hoog.

Procedure

Het design van het kwantitatieve deel is een cross-sectioneel onderzoek met drie condities. Alle respondenten kregen een samengestelde vragenlijst die bestond uit drie vragenlijsten. De onderzochte vragenlijsten zijn de Endicot Work Productivity Scale (EWPS) (Endicott & Nee, 1997), Work Limitations Questionnaire (WLQ) (Lerner et al.

1998, 2001) en het presenteïsme deel van de Health and work Performance Questionnaire (HPQ) (Kessler et al. 2003, 2004). De volgorde van de vragenlijsten is via de Latin Square methode geroteerd om de effecten van de verschillende opvolgende vragenlijsten tegen te gaan. Een compleet Latijns Vierkant (Latin Square) voor drie vragenlijsten geeft drie versies met verschillende volgordes (123, 231, 312). De verschillende versies werden a-select toegewezen. De vragenlijsten werden bij de mensen thuis of op het werk ingevuld. De deelnemers kregen een enveloppe om de privacy van de medewerkers te garandeerd.

Analyses

Om te bepalen in hoeverre de verschillende vragenlijsten overeenkomen werd steeds de correlatie tussen scores van de twee vragenlijsten berekend. Bij het bepalen van de overeenstemming moest er gecorrigeerd worden voor de onbetrouwbaarheid van de vragenlijsten. De HPQ heeft een lage betrouwbaarheid omdat de hoofdvraag maar uit 1 item bestaat. Een onbetrouwbare vragenlijst zal per definitie een lage overeenstemming vertonen met andere vragenlijsten. Er werd gecorrigeerd voor de onbetrouwbaarheid van de vragenlijsten door een correctie voor “attenuatie” toe te passen. De formule die hiervoor gebruikt is, staat beschreven in het stageverslag (Van den Berg 2007, rapportage 1). In de resultatensectie worden zowel de ongecorrigeerde correlaties als de voor onbetrouwbaarheid gecorrigeerde correlaties gepresenteerd. De ongecorrigeerde correlatie geeft de geobserveerde overeenstemming tussen de vragenlijsten weer. De gecorrigeerde correlatie geeft de overeenstemming weer tussen de vragenlijsten indien de betrouwbaarheid van de vragenlijsten optimaal zou zijn.

Om te analyseren of de abstracte methode (HPQ) meer last heeft van impression management dan de concrete methode (EWPS en WLQ) werden de varianties van de vragenlijsten met elkaar vergeleken. Dit gebeurde op twee manieren: per item en voor de hele test. De resultaten van beide methoden werden met elkaar vergeleken om na te gaan hoe robuust de resultaten zijn.

Voor het vergelijken van de variantie per item werd elk afzonderlijk item uit de EWPS en de WLQ vergeleken met het abstracte item uit de HPQ. Hierbij is rekening gehouden met het feit dat de vragenlijsten gebruik maken van verschillende antwoordschalen. De HPQ gebruikt namelijk een 11 puntsschaal terwijl de EWPS en de WLQ een 5 puntsschaal hanteren. Hierbij is aangenomen dat een 5 puntsschaal hetzelfde gebruikt wordt als een 11 puntsschaal. Wanneer de schaalgrootte afneemt, i.e. er minder antwoordcategorieën zijn, neemt ook de variantie af. Bij de transformatie van een 11 puntsschaal naar een 5 puntsschaal werd de variantie 6.25 keer zo klein (zie voor verdere uitleg het stageverslag).

Bij het vergelijken van de variantie van de gehele concrete vragenlijsten (EWPS en WLQ) met de abstracte vragenlijst (HPQ) moest rekening worden gehouden met verschillen in het aantal items per vragenlijst en met verschillen in de antwoordschalen. De EWPS bevatte 25 items op een 5 puntsschaal, de WLQ bevatte 19 items op een 5 puntsschaal en de HPQ bevatte slechts één item dat voor deze test relevant was en deze heeft een 11 puntsschaal. Het probleem met de verschillen in antwoordschalen werd opnieuw opgelost zoals beschreven staat bij de vergelijking tussen de afzonderlijke items. Het probleem met de verschillen in aantallen items werd opgelost door van alle items uit één test (EWPS en WLQ) een “synthetisch (gemiddeld) item” te creëren. Het probleem hierbij is echter dat een vragenlijst bestaande uit één item (HPQ) een lagere betrouwbaarheid heeft (en daarmee een hogere variantie) dan een vragenlijst bestaande uit meerdere items. Het synthetische item, gebaseerd op een vragenlijst met meerdere items, heeft dus een hogere betrouwbaarheid (en lagere variantie) dan het enkele item uit de HPQ. Aangezien de betrouwbaarheid van het item uit de HPQ onbekend is (geen

interne consistentie) wordt een ‘truc’ uitgehaald om de betrouwbaarheid te schatten. Er wordt gedaan alsof het item onderdeel uitmaakt van de WLQ of EWPS. Hierbij wordt de betrouwbaarheid van één item uit de EWPS of WLQ gebruikt om de betrouwbaarheid van het item uit de HPQ te schatten. Op basis van deze geschatte betrouwbaarheden zijn de varianties geschat.

Bij alle analyses werd gekeken of er verschillen waren tussen mannen en vrouwen.

2.2.3 *Kwalitatieve studie*

Een kleine selectie van de deelnemers (n=6) werd gevraagd een ‘hardop-denken sessie’ te ondergaan. Dit hield in dat de deelnemers bij het invullen van de vragenlijsten hun gedachten hardop moesten uitspreken. De hardop-denken sessies werden op dictafoon opgenomen. Dit gebeurde bij de respondenten thuis. Voordat begonnen werd met de hardop-denken methode kregen de respondenten een instructieblad en drie oefenopgaven. Welke oefenopgaven men kreeg hing af van de vragenlijst waarmee begonnen werd. Men kreeg oefenopgaven die soortgelijk waren aan de vragen uit de eerste vragenlijst van de betreffende versie (zie voor verdere uitleg het stageverslag).

De observator had een passieve rol in het geheel en hield zich dus geheel afzijdig. De vragenlijsten werden ook hier volgens de Latin Square methode geroteerd. Ook hier ontstonden door het Latijnse vierkant drie versies en deze drie versies zijn allen door twee personen ingevuld. Het gebruik van de verschillende versies werd ook hier a-select gedaan. Na afloop van de hardop-denken sessie werden tevens de volgende vragen gesteld:

1. “Hoeveel productieverlies koppelt u aan uw cijfer op productiviteit in de afgelopen twee weken?” (verwijzend naar vraag 3 uit de verkorte HPQ).
2. “Waren er nog onduidelijkheden bij het invullen van de vragenlijst en zo ja welke?”

Ook werd er gevraagd naar zaken die de observator niet duidelijk vond tijdens de hardop-denken sessie. Een voorbeeld hiervan is: “Ik zag dat u tijdens het invullen van de vragenlijsten af en toe moeite had met het kiezen tussen ‘soms’ en ‘zelden’. Kunt u me vertellen hoe dat komt?”

2.3 **Resultaten**

Kenmerken deelnemers

Van de 120 verstuurd vragenlijsten zijn er 94 ingevuld teruggekomen. Dit komt neer op een responsepercentage van 78%. De 94 respondenten waren werknemers met een leeftijd tussen de 22 en 64 jaar met een gemiddelde leeftijd van 37,5 jaar ($SD=11,3$). Dit waren 41 vrouwen tussen de 22 en 61 jaar met een gemiddelde leeftijd van 35,5 jaar ($SD=10,1$) en 53 mannen tussen de 22 en 64 jaar met een gemiddelde leeftijd 36,3 jaar ($SD=11,9$). Zeven personen hebben hun leeftijd niet ingevuld (2 vrouwen en 5 mannen). Van de respondenten hadden er 16 personen een leidinggevende functie, 61 personen een functie op HBO of WO niveau, 16 een functie op MBO niveau en van één vrouw is het functieniveau niet bekend (zie tabel 1).

Tabel 1. Respondenten kwantitatief & kwalitatief onderzoek ingedeeld naar sekse en functieniveau.

	Kwantitatief onderzoek			Kwalitatief onderzoek		
	Man	Vrouw	Totaal	Man	Vrouw	Totaal
Leidinggevende	10	6	16	2	0	2
HBO/WO niveau	34	27	61	1	1	2
MBO niveau	9	7	16	2	0	2
Totaal	53	40	93	5	1	6

Voor het kwalitatieve deel zijn zes “white collar workers” benaderd om de vragenlijst met behulp van de hardop-denken methode in te vullen. Hun leeftijd was tussen de 22 jaar en 57 jaar met een gemiddelde leeftijd van 39,7 ($SD=17,9$) jaar meegedaan. Dit waren 5 mannen tussen de 24 en 57 jaar met een gemiddelde leeftijd van 43,2 jaar ($SD=17,6$) en een vrouw van 22 jaar. Van deze respondenten hadden er twee een leidinggevende functie, twee een functie op HBO niveau en twee een functie op MBO niveau (zie tabel 1).

Betrouwbaarheid vragenlijsten

De betrouwbaarheid van de EWPS (Cronbach's α) was 0.89. De betrouwbaarheidscoëfficiënten (Cronbach's α) van de subschalen van de WLQ (tijd management, mentale en inter-persoonlijke vaardigheden en output vereisten) waren respectievelijk 0.80, 0.91 en 0.88. Dit komt overeen met de betrouwbaarheden zoals vermeld in de literatuur.

2.3.1 Literatuuronderzoek

Kenmerken van instrumenten om productiviteit te meten

In het document “Vragenlijst uitwerking” (Van den Berg 2007, rapportage 2) staan alle instrumenten die productiviteit beogen te meten beschreven. In tabel 2 staat weergegeven in hoeverre de link met productiviteit duidelijk is. Met andere woorden: “In hoeverre meten deze instrumenten werkprestatie?” De link met productiviteit varieert tussen matig en hoog. Voor de drie vragenlijsten die zijn opgenomen in deze studie was de link met productiviteit matig, voor de EWPS, groot voor de HPQ en redelijk voor de WLQ. Opvallend is dat geen van de makers een duidelijk definitie geeft van werkprestatie of productiviteit. In het document “Vragenlijst uitwerking” staan bij iedere vragenlijst referenties vermeld. Het aantal studies waarin de drie vragenlijsten zijn ingezet, is meer dan zes voor de EWPS en meer dan tien voor de HPQ en de WLQ. De overige vragenlijsten in tabel 2 zijn allen minder vaak ingezet of deze informatie is niet bekend. In het document wordt niet ingegaan op de spreiding in werkprestatie in de verschillende studies en de vraag of er sprake was van een verandering in werkprestatie (bij interventiestudies). Zes van de veertien vragenlijsten zijn ook bruikbaar bij gezonde werknemers (o.a. de EWPS en HPQ). Zeven van de veertien vragenlijsten zijn niet bruikbaar bij gezonde werknemers, waaronder de WLQ. De WLQ is ontwikkeld om prestatievermindering te meten bij werknemers met gezondheidsklachten. Dat wil zeggen dat de vragenlijst ook specifiek vraagt naar problemen ten gevolge van gezondheidsklachten. Gezonde werknemers zullen de vragen uit deze vragenlijsten waarschijnlijk al snel beschouwen als niet van toepassing. Van één vragenlijst is de inzetbaarheid bij gezonde werknemers niet bekend.

Tabel 2. Kenmerken van de vragenlijsten die productiviteit beogen te meten

	EWPS	HPQ	AHM-HRA	Prodisq	Prodisq uitbr.	WLQ	HWQ	VPM	APA	VOZW/HLQ	WPAI	SAHAPS	MHIHQ	WPSI
Aantal Items	25	4	5	2	4	15	13	59	6	8	1	32	17	2
Schaal	5	1, 2, 5 & 7	4	11	11	5	11	7	5	4	11	5, 3 en continu	6, 2 & 11	geen
Neutraal / Nvt	Nee	Nee	Ja	Nee	Nee	Ja	Nee	Ja	?	Nee	Nee	Nee	Ja	Nee
Periode	7 dagen	4 weken	4 weken	1 dag	1 week	2 weken	1 week	Geen	2 weken	2 weken	7 dagen	1 Maand	30 dagen	1 jaar
Concreet	Ja	Nee	Ja	Nee	Ja	Ja	Ja	Ja	Ja	Ja	Nee	Ja	Nee	Nee
Abstract	Nee	Ja	Nee	Ja	Ja	Nee	Ja	Nee	Nee	Nee	Ja	Ja	Ja	Ja
Relatief	Nee	Ja	Nee	Ja	Nee	Nee	Ja	Nee	Nee	Nee	Nee	Nee	Ja	Nee
Lengte tijd	5 min	5 min	2 min	2 min	3 min	5 min	7 min	15 min	3 min	5 min	2 min	15 min	10 min	3 min
Beschikbaar NL	Ja	Ja	Ja	Ja	Ja	Ja	Ja	Ja	Nee	Ja	Nee	Nee	Nee	Nee
Gezonde mensen	Ja	Ja	Ja	Ja	Nee	Nee	Ja	Ja	?	Nee	Nee	Nee	Nee	Nee
Geheugen Problemen	Nee	Ja	Nee	Nee	Nee	Nee	Nee	Nee	Nee	Nee	Nee	Ja	Ja	Ja
Aanwezig	Ja	Ja	Ja	Ja	Ja	Ja	Ja	Ja	Nee	Ja	Ja	Ja	Ja	Ja
Verlies Berekenbaar	Nee	Ja	Nee	Ja	Ja	Nee	?	Nee	Nee	Ja	Ja	Ja	Ja	Ja
Validiteit	+	+	?	?	?	++	+	+	?	+	+	?	?	?
Aantal Studies	>6	>10	1	>1	>1	>10	>5	1	>3	?	?	?	?	?
Item link productiviteit	+/-	++	+	++	++	+	+	+/-	+	+	++	+/-	+/-	++
Eenduidig	-	+	+	+	+	+	+	++	?	+	+	+	+/-	+
Gevoelig voor Bias	Nee	Ja	Nee	Ja	Nee	Nee	?	Nee	Nee	Nee	Nee	Nee	Nee	Nee
Aantal Items	Aantal Items dat de vragenlijst bevat om productiviteit te meten.						Gezonde mensen							
Schaal	Aantal antwoord mogelijkheden per item.						Geheugen Problemen							
Neutraal / Nvt	Is er een neutrale antwoord mogelijkheid bv. "niet van toepassing"						Aanwezig							
Periode	De terugvraag periode van de test.						Verlies Berekenbaar							
Concreet	Zijn er items in de test die concreet zijn ja/nee.						Validiteit							
Abstract	Zijn de items in de test die van een abstract niveau zijn ja/nee.						Aantal Studies							
Relatief	Bavet de test vergelijkende elementen ja/nee.						Item link productiviteit							
Lengte tijd	De tijd die nodig is om de test in te vullen (geschat).						Eenduidig							
Beschikbaar NL	Is de test beschikbaar in de Nederlandse taal ja/nee.						Gevoelig voor Bias							

2.3.2 *Kwantitatieve studie*

De drie versies waarin de volgorde van de drie vragenlijsten verschilden (123, 231, 312) zijn respectievelijk 34, 33 en 33 keer ingevuld.

Productiviteitsverliezen

De drie testen geven verschillende manieren om de scores op productiviteitsverlies te berekenen. De EWPS geeft een score van 0 tot 100, de HPQ en de WLQ geven een procent score. In tabel 3 zijn de score overzichtelijk weergegeven. De WLQ geeft een lagere score op productieverlies dan de HPQ.

Tabel 3. Score op productiviteitsverlies gemeten met de verschillende vragenlijsten.

	totale steekproef (N = 94)			mannen (N = 53)			vrouwen (N = 41)		
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
EWPS ⁺	21.6	10.8	0 - 51	22.9	11.2	7 - 51	20.0	10.2	0 - 40
HPQ	28.2	11.8	10 - 70	28.1	13.5	10 - 70	28.3	9.5	10 - 50
WLQ	5.6	3.1	0 - 13	5.8	3.1	0 - 13	5.3	3.2	0 - 12

⁺ geen procent score.

Overeenstemming tussen vragenlijsten

In tabellen 4 en 5 staan respectievelijk de ruwe en voor betrouwbaarheid gecorrigeerde correlaties tussen de vragenlijsten weergegeven. Zowel de resultaten voor de hele groep als per geslacht worden gepresenteerd.

EWPS versus HPQ. De ongecorrigeerde correlatie tussen de EWPS en de HPQ is 0.49 en de voor betrouwbaarheid gecorrigeerde correlatie is 1. Voor mannen is de gecorrigeerde correlatie 2.1. Deze correlatie is groter dan 1 door de correctie voor onbetrouwbaarheid. Dit kan gebeuren wanneer de items van een vragenlijst onderling minder samenhangen dan dat deze samenhangen met de andere test. Na correctie voor betrouwbaarheid is de correlatie tussen de EWPS en de HPQ hoog voor de gehele populatie en voor de mannen en laag voor de vrouwen.

WLQ versus HPQ. De ongecorrigeerde correlatie tussen de WLQ en de HPQ is 0.34 en de voor betrouwbaarheid gecorrigeerde correlatie 0.54. Na correctie voor betrouwbaarheid is de correlatie tussen de WLQ en de HPQ matig voor de hele populatie, redelijk voor mannen en laag voor vrouwen.

WLQ versus EWPS. De ongecorrigeerde correlatie tussen de WLQ en de EWPS is 0.66 en de voor betrouwbaarheid gecorrigeerde correlatie 0.71. Na correctie voor betrouwbaarheid is de correlatie voor de hele populatie en voor de mannen redelijk tot hoog. Voor de vrouwen is de correlatie redelijk ($r=0.63$).

Tabel 4. Ruwe correlaties tussen de vragenlijsten.

	Totale steekproef (N = 94)			Mannen (N = 53)			Vrouwen (N = 41)		
	EWPS	HPQ	WLQ	EWPS	HPQ	WLQ	EWPS	HPQ	WLQ
EWPS									
HPQ	0.49*			0.63*			0.24		
WLQ	0.66*	0.34*		0.71*	0.44*		0.59*	0.19	

* significant met $\alpha = 0,05$

Tabel 5. Voor betrouwbaarheid gecorrigeerde correlaties tussen de vragenlijsten.

	Totale steekproef (N = 94)			Mannen (N = 53)			Vrouwen (N = 41)		
	EWPS	HPQ	WLQ	EWPS	HPQ	WLQ	EWPS	HPQ	WLQ
EWPS									
HPQ	1*			2.1*+			0.49		
WLQ	0.71*	0.54*		0.77*	0.70*		0.63*	0.26	

* significant met $\alpha = 0,05$, + correlatie groter dan 1 als gevolg van correctie.

Invloed van impression management

Om te onderzoeken of de concrete vragenlijsten minder last hebben van impression management dan de abstracte HPQ vraag vergelijken we de variantie op de scores van beide vragenlijsten met de HPQ. In tabel 6 staan de varianties van de vragenlijsten als geheel weergegeven. Er bestaan twee geschatte varianties voor de HPQ. Deze twee varianties verschillen van elkaar omdat de ene variantie geschat is via de betrouwbaarheid van de EWPS en de andere variantie via de betrouwbaarheid van de WLQ.

EWPS versus HPQ. Van de scores op de 25 concrete vragen uit de EWPS hebben er 24 een significant ($\alpha \leq 0,05$) grotere variantie dan de score op de abstracte vraag uit de HPQ, na correctie voor de schaalgrootte. Er was geen enkele score op een concrete vraag die een significant lagere variantie had dan de variantie op de score van de abstracte vraag uit de HPQ. De variantie van de totaalscores op de EWPS is 2.76 keer zo groot ($0.160/0.058$) als de variantie van de scores op de abstracte vraag uit de HPQ (Tabel 6). Er waren geen significante verschillen tussen mannen en vrouwen.

WLQ versus HPQ. Alle scores op de 19 concrete items uit de WLQ vertonen een significant grotere variantie ($\alpha = 0,05$) dan de score op het abstracte item uit de HPQ na correctie voor de schaalgrootte. Ook hier is er geen enkele score op een concrete vraag die een significant lagere variantie heeft dan de score op de abstracte vraag uit de HPQ. De variantie van de totaalscores op de WLQ is 2.68 keer zo groot ($0.281/0.105$) als de variantie van de scores op de abstracte vraag uit de HPQ (zie tabel 6). Ook hier werden geen significante verschillen tussen mannen en vrouwen gevonden.

Tabel 6. Geschatte ware score varianties van de EWPS, WLQ en HPQ (hele vragenlijsten).

	totale steekproef (N = 94)	mannen (N = 53)	vrouwen (N = 41)
	Variantie	Variantie	Variantie
EWPS synthetisch item ¹	0.160	0,181	0,146
HPQ op 5 puntsschaal via betrouwbaarheid EWPS	0.058	0,073	0,037
HPQ op 5 puntsschaal via betrouwbaarheid WLQ	0.105	0.123	0.075
WLQ synthetisch item ¹	0.281	0.251	0.290

¹ Synthetisch item= Een 'gemiddeld item' dat gecreeërd is uit alle items uit 1 test

Bovenstaande resultaten ondersteunen onze hypothese dat abstracte vragen meer last hebben van impression management dan concrete vragen. De scores op concrete vragen hebben per item en over een hele test een grotere variantie dan de scores op de abstracte vraag.

2.3.3 *Kwalitatieve studie*

Uit de hardop-denken protocollen bleek dat de concrete vragen kort en snel aan de hand van concrete informatie beantwoord werden. Men las de vraag en antwoordde in hoeverre dit specifieke gedrag voorkwam. Er waren echter wel wat problemen met onduidelijkheden in de vragen. Vooral in de EWPS, waarin een groot aantal samengestelde vragen zitten, was enige onduidelijkheid onder de respondenten.

De abstracte vragen uit de HPQ werden ook kort en snel beantwoord, maar bij navraag hoe men tot een antwoord kwam antwoordde men dat men dit op gevoel deed of aan de hand van een beoordelingsgesprek. Dit ondersteunt de hypothese dat abstracte vragen gevoeliger zijn voor impression management dan concrete vragen. Het antwoorden op gevoel laat meer ruimte voor impression management dan antwoorden uit het geheugen aan de hand van concrete informatie. In de Appendix wordt uitgebreider inzicht gegeven in de belangrijkste resultaten van de hardop-denken protocollen per vragenlijst.

Interpretatie van de schaal

Bij de schaal 5 puntsschaal van de EWPS en de WLQ bestaande uit 'nooit', 'zelden', 'soms', 'vaak' en 'bijna altijd' bleken een aantal mensen het verschil tussen zelden en soms moeilijk te kunnen maken. Een respondent antwoordde bij navraag over dit fenomeen "Soms is als ik het zeker weet dat het in de afgelopen twee weken is voorgekomen en zelden is als het wel eens voorkomt, maar ik niet zeker weet of het in de afgelopen twee weken is voorgekomen". Een andere respondent zei toen er naar dit fenomeen gevraagd werd; "Ik heb het idee dat er aan zelden een groot tijdsaspect zit. Ik vind het een beetje vreemd om zelden te gebruiken wanneer het over een periode van 2 weken gaat". Bij navraag onder drie deelnemers hoe men de 11 puntsschaal van de HPQ interpreteerde, kwam het volgende naar voren. De respondent die zichzelf een 8 gaf voor zijn prestatie in de afgelopen weken vond dat dit redelijk overeen kwam met 20% productieverlies die de HPQ daar aan koppelt. De respondent die zichzelf een 7 gaf vond dit niet, deze vond 20% tot 25% productieverlies een redelijke schatting. De respondent die zichzelf een 6 gaf vond dat dit overkomt met ongeveer 30% productieverlies. Wat opviel tijdens de hardop-denken sessie was dat een respondent zei "Mijn collega's zijn wel een 7" bij een vraag uit de HPQ. Dit zou men kunnen interpreteren als een beoordeling waarbij een soort rapportcijfer gegeven werd. Het is niet helemaal duidelijk wat ermee bedoeld werd, maar het is goed voor te stellen dat een 10 puntsschaal als een rapportcijfer gebruikt wordt.

2.4 Discussie

Tijdens de uitvoering van dit onderzoek kwamen er een aantal factoren aan het licht die het meten van werkprestatie en productiviteitsverliezen met behulp van een vragenlijst bemoeilijken.

Dit is ten eerste impression management, het proces waarbij mensen het beeld dat anderen van hen hebben proberen te beïnvloeden. Met name bij het gebruik van abstracte vragen dient rekening te worden gehouden met impression management en de verminderde spreiding die dit tot gevolg kan hebben. Door alleen gebruik te maken van concrete vragen zou dit probleem grotendeels kunnen worden omzeild. Echter, de vertaling van concrete vragen naar productiviteitsverliezen (liefst uitgedrukt in geld) is moeilijker dan de vertaling van abstracte vragen naar productiviteitsverliezen.

Ten tweede kunnen er problemen ontstaan bij de interpretatie van de schaal. Wat opviel tijdens de hardop-denken sessie was dat men de schaal zou kunnen interpreteren als een beoordeling waarbij een soort rapportcijfer gegeven werd. Dit hoeft niet nadelig te werken, maar het probleem zou kunnen zijn dat men de getallen onder de zes als 'onvoldoende' beschouwt. Men is door deze grens minder bereid dit deel van de schaal te gebruiken, waardoor wellicht 'restriction of range' optreedt. Wanneer men het gebruik van de schaal door de huidige respondenten bekijkt, blijkt het ook zo te zijn dat men de getallen onder de zes nauwelijks gebruikt. Deze scheve respons is echter ook te verwachten, want de steekproef bestaat uit gezonde mensen. Het zou daarom ook vreemd zijn als er veel mensen laag (i.e. "onvoldoende") scores op productiviteit.

Een derde factor is de terugvraagperiode. Over het algemeen geldt dat hoe langer de periode is waarnaar teruggevraagd wordt, des te groter de invloed van geheugenfouten is. Wanneer teruggevraagd wordt naar een periode langer dan twee weken blijkt dat informatie van de laatste week de meeste invloed heeft op het antwoord en dat een korte terugvraagperiode betrouwbaarder is (Lerner et al. 2001). Een verschijnsel dat samenhangt met de terugvraagperiode is "mood-state-dependent retrieval". Dit houdt in dat de stemming of de 'state' van een persoon invloed heeft op de toegankelijkheid van zijn geheugen. Men heeft beter toegang tot informatie uit het geheugen dat in een gelijke stemming of 'state' is onthouden dan informatie dat in een andere stemming of 'state' is onthouden. Dus als mensen even verminderd productief zijn, ze zich ook vaker meer bewust zijn van andere momenten waarop ze verminderd productief waren. Het tegenovergestelde geldt ook; mensen die tijdelijk meer productief zijn, zijn zich ook vaak meer bewust van andere momenten waarop ze meer productief waren (Kenealey 1997; Meloy 2002; Karney & Coombs 2000). Een ander fenomeen dat opspelt wanneer men vraagt naar het verleden, is dat men de neiging heeft om vooruitgang te zien. Het blijkt zo te zijn dat men de neiging heeft oneigenlijke verbetering te zien, of de verbetering te overdrijven (Karney & Coombs 2000). Nisbett et al. (1977) concluderen in hun studie dat mensen gaten in het verleden naar eigen idee invullen. Alleen duidelijke stimuli en duidelijk oorzaken worden goed onthouden, terwijl bij minder in het oog springende stimuli, of onduidelijke oorzaken, dit niet het geval is. Het is niet helemaal duidelijk vanaf wanneer al deze verschijnselen gaan optreden.

Een vierde factor is het gebrek aan kennis bij het vertalen van de testcores naar productiviteitsverliezen. Alleen de WLQ bevat een algoritme waarmee productiviteitsverliezen kunnen worden berekend. Bij het vertalen van de HPQ scores naar productiviteitsverliezen is er vanuit gegaan dat een score van 8 gelijk staat aan een productiviteitsverlies van 20%, een score van 7 met een productiviteitsverlies van 30% etc. Over het omzetten van EWPS scores naar productiviteitsverliezen is zover wij weten niks bekend. De productiviteitsverliezen berekend met de WLQ waren veel lager dan de productiviteitsverliezen berekend met de HPQ. Het berekenen van

productiviteitsverliezen is echter heel complex, omdat idealiter ook rekening moet worden gehouden met de mate waarin iemands verminderde functioneren invloed heeft op het grotere geheel van functioneren van de hele afdeling of organisatie.

Een vijfde factor is het gebruik van samengestelde vragen. Uit de hardop-denken sessie kwam naar voren dat sommige werknemers een gemiddeld antwoord geven of alleen antwoord geven op het eerste deel van de vraag en niet verder lezen. Het belang van eenduidige vragen bij het construeren van een vragenlijst komt hier dan ook duidelijk naar voren. Bij samengestelde vragen zijn er verschillende manieren om een vraag te interpreteren en er zijn verschillende manieren om tot een antwoord te komen. Interpretatie van de scores is hierdoor lastig, men kan achteraf niet achterhalen welke vraag er beantwoord is, omdat men niet weet welke strategie men gekozen heeft om tot een antwoord te komen.

Een zesde factor is een hoge inter-item correlatie. Bij het invullen van de EWPS bleek dat men moeilijk onderscheid kon maken tussen twee items (item 16 & 17). De inter-item correlatie tussen de betreffende items is ook erg hoog ($r=0.80$), zeker vergeleken met de gemiddelde inter-item correlatie ($r=0.26$). Het betreft hier de hoogste inter-item correlatie in de test. Dit geeft aan dat de toegevoegde waarde van het tweede item heel laag is. Men zou kunnen overwegen om één van deze items daarom ook weg te laten.

Met behulp van de hardopdenkprotocollen wilden we onder meer achterhalen of de WLQ ook inzetbaar is bij gezonde werknemers. Deze vragenlijst vraagt immers specifiek naar productiviteitsverliezen ten gevolge van gezondheidsklachten. Gezonde medewerkers zouden dus eigenlijk allemaal “nooit” moeten invullen op deze vragenlijst. Opvallend genoeg was dit wel het geval bij de deelnemers aan de hardopdenkmethoden maar niet bij de deelnemers die alleen de vragenlijst invulden. Het is echter onduidelijk of deze bevinding verklaard kan worden doordat de deelnemers aan de hardopdenkmethode de vragen beter lasen en overdachten voordat zij een antwoord gaven of dat zij niet durfden toe te geven tegenover de persoon die de test afnam dat zij gezondheidsproblemen hadden.

2.5 Conclusie

Geen van de drie onderzochte vragenlijst is in de huidige vorm inzetbaar bij projecten van de groep bewegen en arbeid. De overeenstemming tussen de drie vragenlijsten was matig tot groot voor de hele onderzoekspopulatie. De invloed van impression management lijkt groter te zijn bij abstracte vragenlijsten (HPQ) dan bij concrete vragenlijsten (EWPS/WLQ), zowel op itemniveau als op het niveau van de hele vragenlijst. De resultaten uit de hardop-denken protocollen ondersteunen deze conclusie. Het ligt op basis van de bevindingen uit deze studie voor de hand om een nieuwe vragenlijst te ontwikkelen. Deze vragenlijst zou bij voorkeur zowel concrete als abstracte vragen moeten bevatten, onder meer omdat concrete vragenlijsten moeilijk te vertalen zijn naar productiviteitsverliezen. Ook dient rekening te worden gehouden met de genoemde belemmerende factoren bij het meten van productiviteitsverliezen met behulp van een vragenlijst.

3 Belangrijkste bevindingen en vervolgstappen

3.1 Belangrijkste bevindingen

De belangrijkste bronnen van informatie voor het formuleren van vervolgstappen zijn de resultaten van de studies die voorafgingen aan het stageonderzoek (o.a. RSI@Work, Alive@Work), het stageonderzoek zelf, en een recente review van Schultz en Edington (2007). Ook hebben we een model gemaakt waarin de relatie tussen bewegen en productiviteit inzichtelijk wordt gemaakt. Dit model zal ten grondslag liggen aan alle vervolgstappen. Aan de hand van deze bevindingen is een stappenplan voor de komende jaren opgesteld waarvan het uiteindelijke doel is een vragenlijst te ontwikkelen die gebruikt kan worden om werkprestatie te meten in een overwegend gezonde werknemerspopulatie.

Onderzoek dat voorafging aan het stageonderzoek:

1. Bij kantoormedewerkers is werkprestatie meestal niet goed te meten met behulp van objectieve methoden.
2. De HPQ, WLQ en de EWPS lijken kansrijke instrumenten om werkprestatie mee te meten.
3. De HPQ is gevalideerd, maar in de groep “white collar workers” was de validiteit het laagst.
4. Er is onvoldoende bekend over de samenhang tussen bovenstaande vragenlijsten, de manier waarop medewerkers tot een antwoord komen op deze vragenlijsten en de problemen die zij hierbij ondervinden.
5. De spreiding op de HPQ scores is erg laag waardoor er vermoedelijk geen relaties met lichamelijke activiteit en fitheid kunnen worden aangetoond. Dit blijkt uit de resultaten van zowel het RSI@Work onderzoek als het Alive@Work onderzoek.
6. Uit een tweetal studies blijkt dat de HPQ niet gevoelig is voor verandering (RSI@Work studie en studie groep Gisela Sjøgaard).

Stageonderzoek:

1. De HPQ bevat abstracte vragen waarbij het aan medewerkers zelf wordt overgelaten wat zij onder werkprestatie verstaan. De WLQ en EWPS bevatten concrete vragen over omstandigheden die tot verminderde werkprestatie kunnen leiden.
2. De WLQ vraagt specifiek naar verminderde werkprestatie ten gevolge van gezondheidsproblemen, terwijl de HPQ en EWPS dat niet doen.
3. De samenhang tussen de drie vragenlijsten is matig tot goed. De laagste correlatie werd gevonden tussen de HPQ en WLQ, terwijl deze vragenlijsten beiden als de “beste vragenlijsten” worden gezien in de review van Schultz en Edington. De reden voor deze lage samenhang kan zijn dat zij niet hetzelfde beogen te meten. De WLQ meet prestatievermindering ten gevolge van ziekte, terwijl de HPQ verminderde werkprestatie veel breder in kaart probeert te brengen.
4. De correlatie tussen de vragenlijsten is hoger voor mannen dan voor vrouwen. Het is onbekend wat hiervan de oorzaak is. Mogelijk hebben mannen meer de neiging dan vrouwen om consistent te antwoorden.
5. Impression management (zoals het geven van sociaal wenselijke antwoorden) lijkt een grotere invloed te hebben op de spreiding in scores van de abstracte vragenlijst

(HPQ) dan op de spreiding in scores van de concrete vragenlijsten (WLQ en EWPS).

6. Medewerkers die de WLQ invulden terwijl zij hardop uitspraken wat ze dachten (hardop-denken protocol) (n=6) gaven vrijwel allemaal aan dat zij géén verminderde werkprestatie hadden omdat ze geen gezondheidsklachten hadden. Medewerkers die de WLQ invulden zonder dit hardop-denken protocol, gaven veel vaker aan dat ze wel verminderde werkprestatie hadden ten gevolge van gezondheidsklachten. Het is onbekend of dit verschil verklaard kan worden doordat de medewerkers die het hardop-denken protocol volgden de vragen beter lasen dan de andere medewerkers of doordat de eerste groep meer de neiging had om sociaal wenselijke antwoorden te geven door de aanwezigheid van de observator. Indien de eerste verklaring juist is, zou dat betekenen dat medewerkers de vragen uit de WLQ over het algemeen niet beantwoordden op de manier zoals het de bedoeling is.
7. Concrete vragen werden kort en snel beantwoord aan de hand van concrete informatie. De abstracte vragen werden ook kort en snel beantwoord, maar bij navraag worden deze vragen op gevoel te worden beantwoord zonder gebruik te maken van concrete feiten.
8. De EWPS bevat enkele samengestelde vragen die verwarring opleveren voor medewerkers. Indien de samengestelde vragen worden aangepast, lijken een aantal concrete vragen uit de EWPS in de toekomst inzetbaar bij een nieuw te ontwikkelen vragenlijst.

Review Schultz en Edington 2007:

1. De kosten die samengaan met productiviteitsverlies ten gevolge van verzuim en lichamelijke beperkingen vormen slechts een deel van de totale kosten aan productiviteitsverliezen.
2. Veel vragenlijsten om productiviteit te meten zijn ontwikkeld om productiviteitsverliezen ten gevolge van ziekte (presenteïsme) te meten en niet om productiviteitsverliezen te meten bij gezonde medewerkers.
3. Bij het meten van presenteïsme worden vaak de kosten die samenhangen met verminderde werkoutput, fouten op het werk, en het onvermogen om te voldoen aan de productiedoelen van de organisatie in kaart gebracht.
4. Er zijn vele redenen van verminderde productiviteit die niks te maken hebben met verminderde gezondheid.
5. De vragenlijsten die het meest gebruikt worden om presenteïsme te meten zijn de Work Limitations Questionnaire (WLQ) en de Health and Performance Questionnaire (HPQ). Beide vragenlijsten zijn relatief goed gevalideerd.
6. Van de meeste vragenlijsten op dit terrein is de validiteit en betrouwbaarheid bepaald, maar de kwaliteit van deze studies varieert.
7. Het is nog steeds onbekend hoe en of de resultaten uit de vragenlijsten kunnen worden uitgedrukt in geld. In een recente review waarin 20 instrumenten werden onderzocht werd er geen vragenlijst gevonden die gevalideerd was wat betreft het bepalen van kosten ten gevolge van productiviteitsverliezen.
8. Er is nog steeds geen “beste methode” beschikbaar om productiviteit te meten.
9. Sommige werknemerspopulaties hebben in onderzoek meer aandacht gekregen dan anderen omdat bepaalde werknemers populaties bij uitstek geschikt zijn om productiviteit bij te meten (bijv. call-center medewerkers).
10. Onderzoek heeft aangetoond dat zowel verschillende ziekten en aandoeningen als gezondheidsrisicofactoren geassocieerd zijn met verminderde werkprestatie ten gevolge van gezondheidsklachten (presenteïsme).

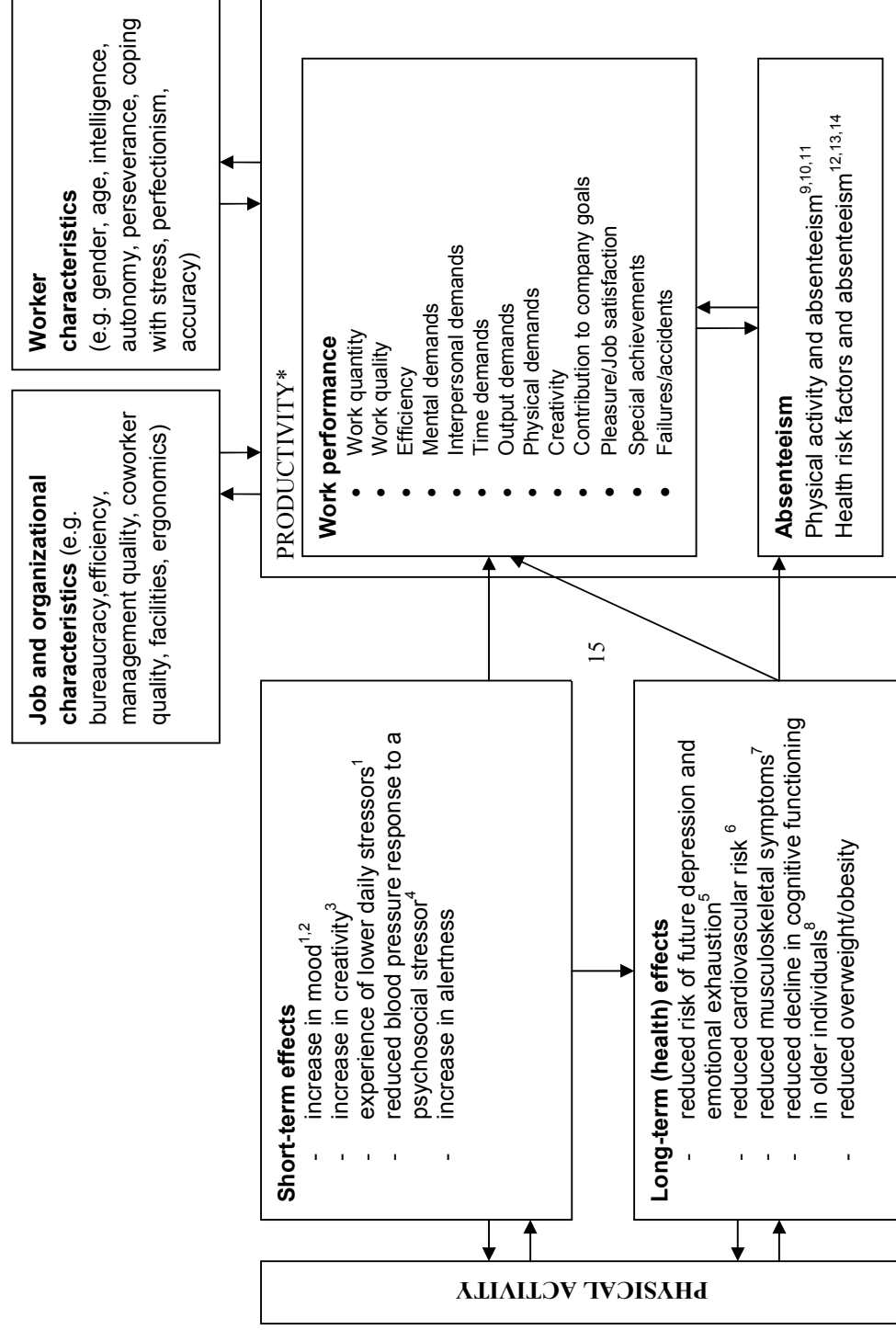
Conclusie: Er is behoefte aan een instrument dat inzetbaar is in een gezonde werknemerspopulatie en gevoelig is voor veranderingen in werkprestatie. Geen van de bestaande vragenlijsten voldoet aan deze behoefte. De drie vragenlijsten die onderzocht zijn hebben alle drie een aantal nadelen. Het nadeel van de WLQ is dat deze vragenlijst alleen prestatievermindering ten gevolge van gezondheidsklachten meet en hiermee zonder aanpassingen niet inzetbaar is in een populatie van (voornamelijk) gezonde werknemers. Het nadeel van de HPQ is dat deze vragenlijst abstracte vragen bevat die gevoelig zijn voor impression management. Het nadeel van de EWPS is dat deze vragenlijst samengestelde vragen bevat. Daarnaast bevatten zowel de EWPS als de WLQ concrete vragen die niet gemakkelijk te vertalen zijn naar kosten ten gevolge van productiviteitsverlies.

Een nieuw te ontwikkelen vragenlijst zou bij voorkeur zowel abstracte als concrete vragen moeten bevatten, zodat het effect van impression management zo laag mogelijk is en de spreiding in productiviteit zo hoog mogelijk. Abstracte vragen zijn nodig om de vertaling naar kosten te kunnen maken. De WLQ, HPQ en EWPS lijken alledrie kansrijk om als uitgangspunt te dienen voor het ontwikkelen van een nieuwe vragenlijst.

3.2 Criteria voor een nieuwe vragenlijst

1. In te vullen door gezonde werknemers en ongezonde werknemers
2. Voldoende spreiding in de antwoorden
3. Begrijpelijke vragen
4. Eenduidige vragen
5. Geen samengestelde vragen (twee vraagstellingen ineen)
6. Relevante vragen
7. Ongevoelig voor impression management (sociaal wenselijke antwoorden)
8. Gevalideerd in een overwegend gezonde Nederlandse werknemerspopulatie
9. Betrouwbaar
10. Een hoge interne consistentie
11. Gevoelig voor verandering
12. Gevoelig genoeg om kleine verschillen in werkprestatie te meten
13. Toepasbaar bij verschillende beroepen en branches maar in ieder geval bij kantoormedewerkers
14. Beschikbaar in het Nederlands
15. Niet te lang (binnen 10 minuten in te vullen)
16. Terugvraagperiode van twee weken
17. Bevat zowel abstracte als concrete vragen
18. Totaalscore op de vragenlijst te vertalen naar kosten

3.3 Model bewegen en productiviteit



* Lost productivity = reduced work performance and absenteeism

REFERENCES (in model)

1. Short-term (intermediate) effects

¹ Steptoe A, Kimbell J, Basford P. Exercise and the experience and appraisal of daily stressors: a naturalistic study. *Journal of Behavioural Medicine* 1998;21(4):363-374.

² Hoffman MD, Hoffman DR. Does aerobic exercise improve pain perception and mood? A review of the evidence related to healthy and chronic pain subjects. *Curr Pain Headache Rep.* 2007;11(2):93-7.

Conclusions: Aerobic exercise can cause an acute improvement in mood as well as a reduction in the perception of pain from a painful stimulus. Regular exercise training also may offer some protection from depression, is clinically useful in treating certain psychiatric and chronic pain conditions, and may allow for an enhancement of the acute improvements in mood from a single exercise session. The utility of aerobic exercise training for improving mood disturbances and pain perception among patients with chronic pain requires further investigation.

³ Gondola JC, Tuckman BW. Effects of a systematic program of exercise on selected measures of creativity. *Percept Mot Skills* 1985;60(1):53-4. Results: 23 students ran for 20 min. for 16 sessions. During the first and last class sessions, before running, they and a control group, were tested on 3 measures of creative thinking (verbal adaptability, diversity and originality). After the exercise sessions there were small but significant gains in measures of Remote Consequences and Alternate Uses.

⁴ Hamer M, Taylor A, Steptoe A. The effect of acute aerobic exercise on stress related blood pressure responses: A systematic review and meta-analysis. *Biological Psychology* 2006;71(2):183-190.

2. Long-term (health) effects

⁵ Bernaards CM et al. Can strenuous leisure time physical activity prevent psychological complaints in a working population? *Occup Environ Med.* 2006;63(1):10-6.

Conclusions: Strenuous leisure time physical activity (1-2 times per week) was associated with a lower risk of long term absenteeism (>21 days), whereas physical activity at a higher frequency was not.

⁶Salonen JT et al. *Am J Epidemiol* 1988;127(1):87-94.

⁷Van den Heuvel SG, Heinrich J, Jans MP, van der Beek AJ, Bongers PM. The effect of physical activity in leisure time on neck and upper limb symptoms. *Prev Med* 2005;41(1):260-7.

⁸Hillman CH, Weiss EP, Hagberg JM, Hatfield BD. The relationship of age and cardiovascular fitness to cognitive and motor processes. *Psychophysiology* 2002;39:303-312.

3. Physical activity and absenteeism

⁹ Heuvel SG van den et al. Effect of sporting activity on absenteeism in a working population. *Br J Sports Med* 2005;39:e15.

Conclusions: Employees practicing sports take sick leave significantly less often than their colleagues not practicing sports, while their periods of sick leave are shorter, especially when their work is sedentary

¹⁰ Proper KI et al. Dose-response relation between physical activity and sick leave. *Br J Sports Med* 2006;40(2):173-8.

Results: In two databases, workers meeting the recommendation of vigorous physical activity (active at a vigorous level for at least three times a week) had significantly less sick leave: more than one day over two months and more than four days over a year. The duration of vigorous physical activity was not associated with sick leave.

¹¹ Van Amelsvoort LGPM et al. Leisure time physical activity and sickness absenteeism; a prospective study. *Occupational Medicine* 2006;56:210-212.

Deelnemers werden als actief beschouwd indien zij tenminste 2 dagen per week aan lichamelijke activiteit in de vrije tijd deden. Het artikel maakt niet duidelijk welke vormen van lichamelijke activiteiten worden meegenomen.

4. Health risk factors and absenteeism

¹² Bertera RL. The effects of behavioral risks on absenteeism and health care costs in the workplace. *J Occup Med* 1991;33:1119 – 1124.

¹³ Yen LT, Edington DW et al. Prediction of prospective medical claims and absenteeism costs for 1284 hourly workers from a manufacturing company. *J Occup Med* 1992;34:428-435.

¹⁴ Burton WN et al. The role of health risk factors and disease on workers productivity. *J Occup Environ Med* 1999;41:863-877.

5. Health risks and work performance

¹⁵ Schultz AB and Edington DW. Employee Health and Presenteeism: A systematic Review. *J Occup Rehabil* 2007;17:547-579.

Description of work performance parameters

Performance parameter	Description
Work quantity	-
Work quality	e.g. Work without mistakes (output demands: WLQ/EWPS)
Efficiency	-
Mental performance	Think clearly, work carefully, concentrate on work, lose train of thought, easily read/use eyes, keep mind on work (WLQ), handle workload (output demands: WLQ)
Interpersonal performance	Speak in person/on phone, help others to work, control temper (WLQ), communication (VPM), keep promises (VPM), work with others/good fellow-ship (VPM), customer friendliness
Meet time demands	Work required hours, get going beginning of work day, start on work soon after arriving, work without breaks or rest, stick to routine/schedule (time demands: WLQ), work fast enough, finish work on time (output demands: WLQ/EWPS)
Meet physical demands	Walk/move around work locations, Sit/stand/stay in 1 position, Repeat motions, Bend/twist/or reach, Use handheld tools/equipment, Lift/carry/move objects ≥ 10 lb (WLQ)
Creativity	create new ideas, find new solutions, create new opportunities
Contribute to company goals	In order to use this parameter of work performance, company goals first need to be specified
Pleasure/job satisfaction	Lost interest in work, work = boring, felt excited by the challenge of my work
Special achievements	
Failures/accidents	

Questionnaires

1. VPM = Vragenlijst Prestatie Mogelijkheden (TNO Hoofddorp)
2. EWPS = Endicott Work Productivity Scale
3. HPQ = Health and Work Performance Questionnaire
4. AHM-HRA = Health Risk Appraisal (Musich)
5. PRODISQ (part E) = Productivity and Disease Questionnaire
6. WLQ = Work Limitations Questionnaire
7. HWQ = Health and Work Questionnaire
8. HLQ = Health and Labour Questionnaire
9. WPAI = Work productivity and Activity Impairment Questionnaire (WPAI)
10. SAHAPS or SPS = Stanford/American Health Association Presenteeism Scale
11. WPSI = Wellness Inventory

OPMERKINGEN

Sommige vragenlijsten maken gebruik van schalen (bijvoorbeeld VPM, WLQ). Andere vragenlijsten gebruiken enkele vragen om een bepaald concept te meten (bijvoorbeeld HWQ, HPQ) of maken alleen gebruik van een totaalscore (EWPS). Er is regelmatig overlap tussen de schalen uit verschillende vragenlijsten.

3.4 Stappenplan voor de komende jaren

1. Kritisch bekijken en aanvullen model bewegen en productiviteit
2. Voorleggen model aan experts (in Nederland)
3. Ontwikkelen nieuwe vragenlijst voor het meten van werkprestatie in een overwegend gezonde populatie kantoormedewerkers. Hierbij wordt gebruik gemaakt van concepten die in bestaande vragenlijsten worden gebruikt (zoals werkkwantiteit, werkkwaliteit, efficiency) aangevuld met opmerkingen van experts. De nieuwe vragenlijst omvat zowel concrete als abstracte vragen.
4. Voorleggen nieuwe vragenlijst aan experts
5. Uitzetten van pilots voor het testen van de nieuwe vragenlijst waarin o.a. de hardopdenk methode wordt toegepast
6. Validatie onderzoek. Werkprestatie wordt gemeten bij werknemers waarvan bekend is dat zij tijdelijk verminderd productief zijn (hierbij denken we aan werknemers die net teruggekeerd zijn op het werk na een ziekteperiode of zwangerschapsverlof). Een follow-up meting zal plaatsvinden op het moment dat de medewerker weer volledig productief zou moeten zijn (na enkele maanden). Met de nieuwe vragenlijst wordt nagegaan of er daadwerkelijk veranderingen kunnen worden gemeten in werkprestatie. Een tweede vraagstelling zou kunnen zijn of er verschillen zijn in verminderde werkprestatie en veranderingen in werkprestatie tussen branches/functiegroepen.
7. Inzetten van het nieuwe instrument in studies naar de relatie tussen bewegen en productiviteit en interventiestudies naar de effectiviteit van bedrijfsbewegingsprogramma's.

4 Referenties

- BERNAARDS, C., PROPER, K., HILDEBRANDT, V. (2007). Physical Activity, Cardiorespiratory Fitness, and Body Mass Index in Relationship to Work Productivity and Sickness Absence in Computer Workers With Preexisting Neck and Upper Limb Symptoms. *Journal of Occupational & Environmental Medicine*, 49, 633-640.
- ENDICOTT, J. & NEE, J. (1997). Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacology Bulletin*, 33, 13-16.
- KARNEY, B, COOMBS, R. (2000). Memory Bias in Long-Term Close Relationships: Consistency or Improvement? *Personality and Social Psychology Bulletin*, 26, 959-970.
- KENEALY, P. (1997). Mood-state-dependant Retrieval. The Effects of Induced Mood on Memory Reconsidered. *The Quarterly Journal of Experimental Psychology*, 50A, 290-317.
- KESSLER, R., BARBER, C., BECK, A., BERGLUND, P., MCKENAS, D., PRONK, N., SIMON, G., STANG, P., USTUN, T. & WANG, P. (2003). The World Health Organization Health and Work Performance Questionnaire (HPQ). *Journal of Occupational & Environmental Medicine*, 45, 156-174.
- KESSLER, R., AMES, M., HYMEL, P., LOEPKKE, R., MCKENAS, D., RICHLING, D., STANG, P. & USTUN, T. (2004). Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to Evaluate the Indirect Workplace Costs of Illness. *Journal of Occupational & Environmental Medicine*, 46, S23-S37.
- LERNER, D. BENJAMIN, C. & AMICK, B. (1998). *Work Limitations Questionnaire*. GlaxoWelcome, Inc
- LERNER, D., AMICK, B., ROGERS, W., MALSPEIS, S., BUNGAY, K. & CYNN, D. (2001). Work Limitations Questionnaire. *Medical Care*, 39, 72-85
- MELOY, M. (2002). Mood Driven Distortion of Product Information. *Journal of Consumer Research*, 27, 345-359.
- MUSICH S, HOOK D, BAANER S, SPOONER M, EDINGTON D. (2006). The association of Corporate Work Environment Factors, Health Risks, and Medical Conditions with Presenteeism among Australian Employees. *American Journal of Health Promotion*, 20, 353-363.
- NISBETT, R, DECAMP WILSON, T. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84, 231-259.
- PODSAKOFF, P. & ORGAN, D. (1986). Self-Reports in Organizational Research: Problems and Prospects. *Journal of Management*, 12, 531-544.
- RIJNAARD, M. & JANS, M. (2006). *De samenhang van fysieke conditie en bewegen met arbeidsproductiviteit*. Leiden. TNO Kennis voor zaken.

SCHULTZ AB, EDINGTON DW. Employee Health and presenteeism: A systematic review. *J Occup Rehabil* 2007; 17: 547-579.

VAN DEN BERG J. Productiviteit meten met een vragenlijst. Rapportage 1. Stageverslag TNO Kwaliteit van Leven 2007.

VAN DEN BERG J. Overzicht vragenlijsten productiviteit. Rapportage 2. Stageverslag TNO Kwaliteit van Leven 2007.

A Hardop-denken protocollen

Samenvatting van de belangrijkste bevindingen per vragenlijst

EWPS

In de EWPS (bijvoorbeeld item 9 & 12) komen een aantal samengestelde vragen voor. Het blijkt dat er verschillende manieren zijn waarop men deze vragen beantwoordt. Een respondent antwoordt op de volgende vraag; *“Heeft u zich geërgerd aan of bent u geïrriteerd geraakt door collega’s, baas / leidinggevende, klanten / leveranciers of anderen?”*, “Ik heb me vaak geërgerd aan ... en zelden geërgerd aan... , dus soms”. Hierbij geeft de respondent een gemiddeld antwoord.

Bij een andere vraag (item 13) *“Bent u tijdens het werk zomaar in slaap gevallen of erg slaperig geworden?”* zijn er respondenten die de vraag niet verder hardop lezen dan “... in slaap gevallen” en antwoorden hier meteen “nooit” en er waren respondenten die wel de hele vraag hardop oplazen en dan een antwoord gaven. De eerste soort respondenten antwoorden allemaal “nooit”, bij de tweede soort waren er wel een aantal die “zelden” of “soms” beantwoordden.

Een ander opvallend fenomeen waar twee respondenten tegen aanliepen is dat communicatie tegenwoordig veel via de mail verloopt. De vraag; *“Heeft u gemerkt dat u iemand vergeten was te bellen?”* was op hen niet van toepassing. De ene loste dit op door ‘nooit’ te antwoorden de ander was zo vrij om dit te vertalen naar vergeten iemand te mailen en zei daar soms last van te hebben en vulde ‘soms’ in.

Bij de vragen; *“Heeft u gemerkt dat uw productiviteit lager was dan verwacht, gezien de tijd dat u met iets bezig was?”* en *“Heeft u gemerkt dat uw efficiëntie lager was dan verwacht, gezien de tijd dat u met iets bezig was?”* bleek dat een aantal mensen daar geen onderscheid in konden maken.

HPQ

Bij de vraag; *“Op een schaal van 0 tot 10, waarin de score 0 overeenkomt met de slechts mogelijke prestatie en de score 10 met de best mogelijke prestatie in uw werk, hoe beoordeelt u de werkprestatie van de meeste werknemers die vergelijkbaar werk doen?”*. Twee respondenten antwoordde dat de vraag niet van toepassing was omdat men geen collega’s had die vergelijkbaar werk doen. De ene vulde “n.v.t.” op het antwoordblad terwijl dit geen antwoord mogelijkheid was, de ander ging collega’s beoordelen die geen vergelijkbaar werk doen. De respondenten die wel cijfers gaven deden dit op gevoel, zo zijn een respondent “Mijn collega’s zijn wel een 7”.

De vraag; *“Hoe zou u op dezelfde schaal van 0 tot 10 uw gebruikelijke werkprestatie beoordelen over de afgelopen 2 jaar?”* werd door een aantal ingevuld aan de hand van de beoordeling die de werkgever had gedaan. Zo zei een respondent “Mijn baas zei dat ik een 9 was”. Een tweede respondent zei; “Ik had een 8 op mijn beoordelingen”. De andere respondenten bleken bij navraag (hoe komt u aan dat antwoord?) een antwoord op gevoel te hebben gegeven. De periode “afgelopen 2 jaar” werd door drie personen die nog geen twee jaar aan het werk waren geïnterpreteerd als zolang ik aan het werk ben. Zo zei een respondent “Ik ben nog geen twee jaar aan het werk, maar in de periode dat ik werk ...”

De hoofdvraag van de HPQ; *“Hoe zou u op dezelfde schaal van 0 tot 10 uw algehele werkprestatie beoordelen over de afgelopen 2 weken?”*, daarbij vergeleken de meeste mensen zich met hun prestatie in de afgelopen twee jaar. Zo zei een respondent “de laatste twee weken waren iets minder, een 6” en andere respondent zei “net zo goed als altijd een 8”. De personen die zich niet vergeleken met hun gebruikelijke prestatie

waren de personen die nog niet zolang aan het werk waren. Bij navraag hoe men tot een antwoord komt antwoordde deze respondenten dat ze dat op gevoel deden.

WLQ

Het doen van de hardop denk methode lijkt daar invloed te hebben op het antwoord patroon. Twee respondenten begonnen de vragen te beantwoorden met; "Ik heb geen emotionele en lichamelijke problemen". Ze vulden "nooit moeilijk" in bij de eerste vragen, vervolgens gingen ze toch "soms moeilijk" of "zelden moeilijk" antwoorden. Ze gingen dit echter verbeteren omdat ze geen emotionele en lichamelijke problemen zeiden te hebben. Uiteindelijk vulden ze overal nooit in. De andere vier hardop denk respondenten, daarvan waren er twee ziek geweest, had er één relatie problemen (op het werk) en de andere last van stress en zei vertelden dit ook duidelijk. De respondent met relatie problemen op het werk zei na het derde item van de eerste reeks "Ik heb geen emotionele en lichamelijke problemen" en volgde het bovengenoemde antwoordpatroon, maar bij de vraag; *"In de afgelopen 2 weken, hoeveel van uw tijd maakten uw lichamelijke gezondheid of emotionele problemen het u moeilijk om het onderstaande te doen? werken zonder uw gedachtegang te verliezen"* vertelde hij vanwege zijn relatie hier wel eens last van te hebben en antwoordde; "soms moeilijk". Toen vervolgde hij zijn antwoordpatroon met "nooit moeilijk" om vervolgens nog één maal "zelden moeilijk" in te vullen. Ook deze respondent had dus ook een uitzonderlijk antwoord patroon.