

> TNO rapportage

Het nut van *Expert Judgement* voor analyse van sociale innovaties

TNO innovation
for life

4 november 2014 >

Het nut van *Expert Judgement* voor analyse van sociale innovaties

Datum	4 november
Auteurs	R. Vergeer W.M.G. Courage N.P. van Elst J.N. Huibregtse T. de Jong
Projectnummer	060.08177/01.05 en 060.07452/01.05.05
Rapportnummer	TNO 2014 R11530
Contact TNO	Steven Dhondt
Telefoon	088 866 61 55
E-mail	steven.dhondt@tno.nl

Gezond Leven
Schipholweg 77-89
2316 ZL LEIDEN
Postbus 3005
2301 DA LEIDEN
www.tno.nl

T 088 866 61 00
F 088 866 87 95
infodesk@tno.nl

© 2014 TNO

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 opdrachten 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.

Handelsregisternummer 27376655

Inhoudsopgave

Management samenvatting	i
1 Inleiding	1
1.1 Probleemstelling.....	1
1.2 Doel- en vraagstelling	1
1.3 Onderzoekopzet	1
1.4 Leeswijzer	2
2 Over <i>expert judgement</i>	3
2.1 Wat is <i>expert judgement</i> ?	3
2.2 Verschil met andere expertmethoden	4
2.3 Verrijking van data door Bayesiaans samenvoegen van data en informatie uit <i>expert judgement</i>	4
3 De sociale innovatie: functioneel meten bij Orionis	6
3.1 De case 'functioneel meten bij Orionis'	6
3.2 Stappen bij het opzetten van een <i>expert judgement</i>	7
3.3 Lukt het ons een model van de sociale innovatie te specificeren?.....	7
3.3.1 Modelleren functioneel meten en duurzame inzetbaarheid heeft moeilijkheden	7
3.4 Dilemma's als gevolg van deze moeilijkheden	9
4 Conclusies en discussie	11
4.1 Conclusies.....	11
4.1.1 Expert judgement kan maar beperkt worden toegepast in het domein van sociale innovatie	11
4.1.2 Als input voor de expert judgement is een concreet model nodig met daarin voor experts herkenbare concepten	11
4.2 Discussie	11
4.2.1 Hoe en onder welke voorwaarden kan expert judgement worden toegepast voor het bepalen van de effecten van een sociale innovatie?	11
4.2.2 Aan welke criteria moet een casebeschrijving van sociale innovatie voldoen, wil die zich lenen voor analyse met expert judgement?	12
Literatuur	13
Bijlage 1 Conceptualisering van duurzame inzetbaarheid en vragen voor een <i>expert judgement</i> sessie, versie 13-5-2014	15

Management samenvatting

Het project 'Expert Judgement en Sociale Innovatie' draagt bij aan

1. de ontwikkeling van een impact assessment van sociale innovatie in termen van onder andere het effect op de prestaties en duurzame inzetbaarheid van werknemers, en
2. de ontwikkeling van een expert methode, '*expert judgement*', waarmee effecten kunnen worden gekwantificeerd.

Expert judgement is een wijze voor het verzamelen van informatie van experts over een probleem of onderwerp, waarbij de verschillende experts op een valide wijze rekenkundig geaggregeerd worden.

Het project zoomt in op de vragen:

1. hoe en onder welke voorwaarden kan expert judgement worden toegepast voor het bepalen van de effecten van een Sociale Innovatie;
2. aan welke criteria moet een casebeschrijving van sociale innovatie voldoen, wil die zich lenen voor analyse met *expert judgement*;
3. welke input is nodig vanuit cases sociale innovatie om de impact van een sociale innovatie te kwantificeren?

Om deze vragen te beantwoorden wordt de bij TNO eerder toegepaste methode van *expert judgement* voor het eerst toegepast binnen de context van een te ontwikkelen model van de impact van een sociale innovatie. We verkennen hoe relaties binnen dat model voor een specifiek voorbeeld van een sociale innovatie gekwantificeerd kunnen worden op basis van *expert judgement*.

We stuiten tijdens deze verkenning op een aantal haken en ogen bij de toepassing van *expert judgement* binnen het domein van sociale innovatie.

We concluderen dat *expert judgement* beperkt dient te worden toegepast in het domein van sociale innovatie, namelijk tot factoren en relaties waarvoor geen tot weinig contextgebondenheid geldt. Met contextgebondenheid wordt bedoeld dat de impact van de sociale innovatie sterk afhankelijk is van de specifieke context waarin die plaatsvindt. Een voorbeeld van contextgebondenheid is het invoeren van voice (een kenmerk van workplace innovation, een type sociale innovatie) in een organisatie. De impact daarvan op de betrokkenheid van de medewerkers hangt af of medewerkers de ervaring hebben dat hun inbreng serieus wordt genomen. Dat hangt af van contextfactoren zoals organisatiecultuur en zichtbaarheid van de consequenties van de inbreng van medewerkers.

Contextgebondenheid beperkt *expert judgement* op twee wijzen. Ten eerste is er een beperkt gedeelte visie op hetgeen waarover de experts worden bevroegd. Afhankelijk van de context interpreteert de ene expert dit immers anders dan de andere. Ten tweede beperkt contextgebondenheid de generaliseerbaarheid van de resultaten van *expert judgement* naar een omgeving met andere contexten. Verder concluderen we dat als input voor de *expert judgement* een concreet model nodig is met daarin voor experts herkenbare concepten. Dat betekent dat glashelder en ondubbelzinnig moet zijn over welke factoren de experts bevroegd worden, en - indien naar relaties gevraagd wordt - hoe die met elkaar samenhangen. Voor de verdere verwerking van de resultaten dient glashelder en ondubbelzinnig te zijn op welke wijze de relaties en factoren die gekwantificeerd worden met *expert judgement* samenhangen met andere onderdelen van het gehele probleem. Verder moeten er experts zijn met ervaring met de verschillende relaties en/of concepten in het model, om de parameters

in het model te schatten. De relaties en factoren moeten dus duidelijk te vertalen zijn naar empirische verschijnselen.

1 Inleiding

1.1 Probleemstelling

Om de innovatiegraad van Nederlandse organisaties te verbeteren wordt gekeken naar sociale innovaties. Frank Pot (2012) beschrijft een sociale innovatie als: “*een sociale innovatie blijkt betrekking te hebben op het participatief en in onderlinge samenhang vernieuwen van arbeid, organisatie en personeelsbeleid om het functioneren van werkenden te verbeteren teneinde zowel de organisatieprestaties, de kwaliteit van de arbeid als de arbeidsrelaties op een hoger niveau te brengen*”. Voor het beleid is het van belang dat de impact van deze sociale innovaties helder is. Die impact wordt voornamelijk achteraf (ex post) gemeten aan de hand van surveys. Voor het beleid, maar ook voor de directe stakeholders bij sociale innovaties, is het van belang dat op voorhand (ex ante) helder zou zijn wat een sociale innovatie vooraf zou kunnen betekenen of opleveren. Dit deelproject van het Enabling Technology Project (ETP) “Gedrag en prestatie” onderzoekt of de methode, *expert judgement*, een instrument kan zijn om sociale innovaties ex ante te beoordelen. Met een dergelijke methode zouden we in staat moeten zijn om een volledige ex ante impact assessment uit te voeren. Voor de verschillende genoemde stakeholders bespaart een dergelijke methode onzekerheid en verkeerde investeringen. In vergelijking met een enquête heeft *expert judgement* als voordeel dat het sneller en goedkoper is.

1.2 Doel- en vraagstelling

Het doel van het project is om een verkenning uit te voeren naar de mogelijkheden om *expert judgement* bij sociale innovaties toe te passen om tot een ex ante impact assessment van dergelijke sociale innovaties te komen. Bij de beoordeling van sociale innovaties wordt als nevendoeel gekeken of de Bayesiaanse methode om nieuwe informatie bij beoordelingen toe te voegen helpt om een bredere en beter onderbouwde beoordeling te geven.

De vragen zijn:

1. hoe en onder welke voorwaarden kan *expert judgement* worden toegepast voor het bepalen van de effecten van een sociale innovatie;
2. aan welke criteria moet een casebeschrijving van sociale innovatie voldoen, wil die zich lenen voor analyse met *expert judgement*;
3. welke input is nodig vanuit cases sociale innovatie om de impact van een sociale innovatie te kwantificeren?

1.3 Onderzoeksopzet

Dit project is een eerste verkenning van *expert judgement*. We willen daarbij een sociale innovatie case ‘klaar maken’, opdat ze in een *expert judgement* zou kunnen worden meegenomen. De case die we daarbij gebruiken gaat over ‘functioneel meten bij Orionis’ (zie verder). Daarbij gaat het allereerst om na te gaan of we aan de voorwaarden van een *expert judgement* kunnen voldoen. Pas als deze stap is genomen, willen we de vervolgstappen

nemen: beoordeling en Bayesiaanse behandeling van data. Dit betekent dat we eerste vier stappen van het doen van een *expert judgement* doorlopen (zie voor alle stappen § 3.2):

1. identificatie van een model waarin de vorm van de relaties tussen de verschillende variabelen gespecificeerd zijn;
2. selectie van relaties en/of verdelingskenmerken van variabelen waarvoor expertmening gaat plaatsvinden;
3. selectie expertgroep;
4. opstellen vragenlijst

Vooraf willen we toelichten wat *expert judgement* is en hoe Bayesiaanse technieken erin passen.

1.4 Leeswijzer

Hoofdstuk 2 beschrijft de methode '*expert judgement*'. Hoofdstuk 3 gaat in op de case Ori-onis en de uitwerking van de methode. Omdat we vaststellen dat de methode niet is toe te passen, beschrijven we wat niet werkt. In hoofdstuk 4 worden de onderzoeksvragen beantwoord en conclusies getrokken.

2 Over expert judgement

2.1 Wat is expert judgement?

Expert judgement is een wijze voor het verzamelen van informatie van experts over een probleem of onderwerp, waarbij de verschillende experts op valide wijze op een rekenkundige wijze geaggregeerd worden. Bij *expert judgement* wordt “traditionele statistiek” vervangen door direct een (expert)opinie te vragen naar het gemiddelde en naar een spreidingsmaat. *Expert judgement* met behulp van de in dit onderzoek toegepaste methode van Cooke (zie hieronder) levert informatie van een ‘gecombineerde expert’ (zie hieronder) in de vorm van een kansverdeling en/of samenvattingsmaatstaven van die verdeling, bijvoorbeeld het gemiddelde en de betrouwbaarheid.

Het is gebruikelijk in de economie en milieuwetenschappen om voor (maatschappelijke) vraagstukken via experts tot een oordeel te komen. De methode *expert judgement* wordt vooral toegepast in het milieu- en klimaatdomein, bijvoorbeeld met betrekking tot rang correlaties met toepassing in luchtvervuiling (Morales et al., 2014a) en het onderzoeken van ontwikkelingen in regenbui karakteristieken in Nederland (Morales et al., 2014b).

Expert judgement is een methode die van waarde is bij gebrek aan gegevens over een onderwerp of probleem en/of bij gebrek aan (betaalbare, haalbare) methoden om die gegevens te verzamelen, terwijl er wel behoefte is aan een kwantificeerbare vorm van informatie, zodat bijvoorbeeld de betrouwbaarheid van uitspraken op basis van de informatie kan worden vastgesteld. Het gaat hierbij om het inbrengen van informatie en het ‘behandelen’ van die informatie tot een eensluidend antwoord op de gestelde vraag.

Bij het opzetten en uitvoeren van *expert judgement* worden grosso modo de volgende stappen doorlopen:

1. identificatie van een model waarin de vorm van de relaties tussen de verschillende variabelen gespecificeerd zijn;
2. selectie van relaties en/of verdelingskenmerken van variabelen waarvoor expertmening gaat plaatsvinden;
3. selectie expertgroep;
4. opstellen vragenlijst;
5. sessie(s) met experts: toelichting expertmening en afnemen vragenlijst;
6. verwerking van de resultaten.

Eén van de meest gebruikte methodes bij *expert judgement* is het zogeheten Classical Model van Cooke (Cooke, 1991; Cooke et al., 1999). De eerste stappen van deze methode zijn in dit project gezet. Bij deze methode wordt de zogeheten *decision maker* (de ‘gecombineerde expert’) bepaald door aan iedere expert een gewicht toe te kennen op basis van zijn prestaties op zogenaamde ‘controlevragen’ (variabelen die worden gekwantificeerd op deze wijze worden in de terminologie van *expert judgement* genoemd: de zogenaamde ‘seed variabels’). De ‘controlevragen’ worden gebruikt om voor iedere expert te bepalen hoe *informatief* de expert is en hoe goed *gekalibreerd* de expert is. Deze methode is alleen mogelijk, wanneer er voor iedere vraag niet alleen een enkele puntschatting wordt gevraagd maar een verdeling in de vorm van kwantiel of een kans schatting voor een bepaalde *event*. Veel gebruikte kwantielen zijn het 5%, 50% en 95% kwantiel. De geaggregeerde mening van de experts is dan de gewogen som van de individuele experts.

Een belangrijk vraagstuk bij *expert judgement* is natuurlijk, wanneer het samenvoegen van meningen van experts tot een valide mening over een (maatschappelijk) vraagstuk leidt. Bij het opzetten van een dergelijke beoordeling worden in de praktijk de volgende principes gesteld en bewaakt (Cooke, 1991):

- › reproduceerbaarheid - alle berekeningen moeten reproduceerbaar zijn;
- › accountability - de expertmeningen moeten herleidbaar zijn;
- › empirical control - de expertmeningen moeten in principe empirisch te toetsen zijn;
- › neutraliteit - de methode bevordert experts om hun “ware” opinie te geven;
- › fairness - alle experts worden a priori gelijkwaardig behandeld.

De neutraliteit van de experts is één van de belangrijkste uitgangspunten van een goede combinatie van methode en experts. Dit betekent dat de methode voor het combineren van de expertmeningen moet bevorderen dat experts hun echte inschatting geven. Een tweede belangrijk uitgangspunt is *fairness*. Dit betekent dat alle experts a priori gelijkwaardig behandeld moeten worden.¹

In deze principes ligt dan ook de zwakte van de *expert judgement* methode. Ze is minder gericht op het vormen van consensus, kan duurder zijn dan kwalitatieve methoden en ze kan een schijnzekerheid creëren als de betrouwbaarheid wordt overschat.

2.2 Verschil met andere expertmethoden

Expert judgement verschilt van een andere expertbenadering zoals de *Delphi methode*. In de Delphi methode wordt een groepsproces gestuurd (Helmer, 1966; Gustafson et al., 1973). De (anonieme) experts worden over meerdere ronden of iteraties bevraagd. Hierbij wordt in iedere volgende ronde op basis van feedback van de eerdere (groeps)inschattingen een nieuwe, al of niet bijgestelde, inschatting gevraagd. Op deze wijze wordt middels convergentie naar consensus gezocht. Voordeel van de Delphi methode ten opzichte van de meer mathematisch gefundeerde methoden is, dat deze in het algemeen laagdrempeliger is voor wat betreft de mathematische basis en benodigde kennis van kansrekening. Daarnaast kan de methode synergie tussen de experts bevorderen, omdat de experts van de feedback kunnen “leren”.

De Delphi ronde geeft echter geen methode voor het combineren van experts, als er geen consensus is. Nadeel is dan ook het mogelijk problematische en soms langdurige proces om naar een resultaat te komen, waarbij men er voor moet waken dat het uiteindelijk geen compromisemodel wordt, waarin de werkelijke tendensen en meningen zijn uitgemiddeld of waar experts met een afwijkende mening niet langer actief deelnemen omdat ze hun meningen niet meer in volgende rondes willen verantwoorden. Tot slot biedt de Delphi geen expliciete wijze om met de onzekerheid om te gaan die inherent is aan het doen van schattingen.

2.3 Verrijking van data door Bayesiaans samenvoegen van data en informatie uit *expert judgement*

De informatie van de ‘gecombineerde expert’ die via de methode van Cooke is verkregen, kan worden gecombineerd met data op basis van de regel van Bayes. Met behulp van zoge-

¹ De “controlevragen” maken het mogelijk om naar aanleiding van de informatie die experts geven, een inschatting te maken van hoe “goed” de expert is. In tweede instantie hoeven experts dus niet meer gelijkwaardig behandeld te worden.

naamde “Bayesian updating” kan de schatting van de gecombineerde experts “geupdate” worden op basis van (nieuwe) informatie uit andere bronnen, zoals externe data.

Bij deze methode wordt eerst de informatie van de geaggregeerde expert bepaald, op basis van de methode van Cooke. Vervolgens wordt die informatie gecombineerd met data op basis van de methode van Bayes.

Technisch gezien wordt de informatie van de geaggregeerde experts gebruikt om een *informative prior* op te stellen, die samen met de data wordt gebruikt om de Bayesiaanse schatter te geven op basis van data en expertmeningen (Albert et al., 2012). Kader 2.1 geeft meer informatie over de Bayesiaanse methode.

Kader 2.1 De Bayesiaanse methode om data te verrijken

De Bayesiaanse methode stelt ons in staat om andere verworven kennis te combineren met de informatie die in voorliggende data is vastgelegd. Normaliter beperkt de dataverzameling bij *expert judgement* zich tot aanwezige informatie die met traditionele statistiek wordt samengevoegd. Daarmee wordt bestaande, andere verworven kennis genegeerd. Het voordeel van een Bayesiaanse methode is dat:

- › ze minder eisen stelt aan data;
- › ze daardoor eerder, bijvoorbeeld voordat een beslissing over een interventie is genomen, ingezet kan worden;
- › ze daarmee een veelbelovende methode voor impact assessment is om besluitvorming te ondersteunen, zeker in een context waar nog weinig “harde” gegevens bekend zijn.

Met de methode zorg je ervoor dat de experts tijdens de beoordeling worden voorzien van nieuwe informatie. Zij dienen dan wel een verwachting te formuleren van het effect van deze informatie. Die verwachting kunnen ze op basis van ‘nieuwe feiten’ (bijvoorbeeld data over het effect van de sociale innovatie) bijstellen.

Bayes geeft hiervoor een rekenregel:

$$\text{Posterior} = \text{Prior} * \text{likelijkheid.}$$

Hierin is Prior de initiële verwachting van het effect, likelijkheid de op de data gebaseerde verwachting van het effect en Posterior de bijgesteld verwachting van het effect. Posterior, prior en likelijkheid zijn omgeven met onzekerheid.

In formules:

$$P(E|D) = P(E) * P(D|E) / P(D)$$

- › $P(D|E) / P(D)$: Likelihood = standaard statistiek: schatter parameter + verdeling. De likelijkheid kunnen we berekenen aan de hand van de data.
- › $P(E)$: Prior = een uitdrukking van jouw schatting van de parameter + onzekerheid. De Prior gaan we baseren op expert schattingen.

Samen bepalen die de posterior.

Toepassen van de Bayesiaanse methode vraagt dus om:

1. data analyse;
2. het uitvragen van expert schattingen (*expert judgement*).

3 De sociale innovatie: functioneel meten bij Orionis

3.1 De case ‘functioneel meten bij Orionis’

De sociale innovatie die is gekozen, is het invoeren van “functioneel meten” bij de Sociale Dienst op Walcheren (Orionis). “Functioneel meten” is een vorm van informatie-ondersteund werken, waarbij het informatiesysteem en het besturingssysteem bottom-up worden ontworpen. Hiermee wordt de basis gelegd voor een lerende organisatie, omdat medewerkers goede feedback kunnen krijgen op hun handelen. Parallel aan het ontwerpen en ontwikkelen van het informatiesysteem krijgen de medewerkers training in hoe zij hiervan kunnen leren en hun gedrag kunnen aanpassen. Functioneel meten is dus een sociale innovatie in de klassieke zin, omdat het zowel een technische als een sociale component heeft. Voor een uitgebreidere beschrijving van functioneel meten en de theoretische inbedding, zie kader 3.1.

Kader 3.1 Functioneel meten

Functioneel meten

De Interventie “functioneel meten” heeft als doel een feedbacksysteem in te richten, waarmee medewerkers kunnen leren over de effectiviteit van hun werk. De inrichting van dat feedbacksysteem heeft een technische en een sociale component. De technische component bestaat uit het herinrichten van het informatiesysteem en het ontwerpen van een analysekader. De sociale component beslaat het niveau van zowel de werknemers als het managementniveau. Op het niveau van de werknemers worden trainingen gegeven en intervisiebijeenkomsten georganiseerd die medewerkers moeten helpen om zich een “lerende houding” eigen te maken. Daarnaast wordt de inrichting van het informatiesysteem “bottom-up” ontworpen. Op managementniveau gaat het er bijvoorbeeld om dat experimenteerruimte aan medewerkers wordt geboden, en dat in de HR-cyclus of in teamoverleggen de “lerende houding” wordt gestimuleerd en geborgd.

Theoretische inbedding functioneel meten

Functioneel meten (FM) vertoont om te beginnen veel kenmerken van een sociale innovatie. Frank Pot (2012) beschrijft een sociale innovatie als *“sociale innovatie blijkt betrekking te hebben op het participatief en in onderlinge samenhang vernieuwen van arbeid, organisatie en personeelsbeleid om het functioneren van werkenden te verbeteren teneinde zowel de organisatieprestaties, de kwaliteit van de arbeid als de arbeidsrelaties op een hoger niveau te brengen”*. FM heeft een aantal van deze kenmerken, met name het participatieve aspect, het vernieuwen van arbeid, organisatie en personeelsbeleid en het verbeteren van het functioneren van de werkenden en de organisatie. Pot (2012) wijst op het Job Demand - Job Control (JD-JC) model (Karasek, 1979; Karasek & Theorell, 1990) als een belangrijke inspiratiebron voor het huidige denken over sociale innovatie, met name vanwege de gedachte achter het JD-JC model dat kwaliteit van de arbeid en kwaliteit van de organisatie in elkaars verlengde liggen. In dit project wordt voor het beschrijven van het effect van functioneel meten een model gebruikt dat is geïnspireerd op een JD-JC model (zie hieronder).

Functioneel meten kan ook gezien worden in de context van “gift exchange”. Uit onderzoek van Kube et al. (2012) blijkt dat vooral “non-monetary gifts” motiverend werken. Kube et al. onderscheiden als “gifts” slechts materiële zaken. Bij functioneel meten zijn de “gifts” echter immaterieel.

De leer- en experimenteerruimte die wordt geboden bij FM, kunnen gezien worden door medewerkers als een gift die motiverend kan werken. Uit het genoemde onderzoek blijkt, dat de perceptie van

“moeite” die de werkgever in de “gift” steekt positief bijdraagt aan de motivatie. In dat kader valt het te verwachten dat “immateriële” gifts een grote potentie hebben, omdat ze bijna per definitie een inspanning van het management en/of de organisatie vragen. Volgens Gächter et al. (2012) is ook de perceptie van de motivatie die collega's ontlenen aan de “gift” van groot belang. In dat kader levert ook de intervisie die wordt georganiseerd in het kader van FM een bijdrage.

Ten slotte kan functioneel meten ook worden gerelateerd aan de theorie van intrapreneurship. Centrale elementen in intrapreneurship zijn proactief gedrag, innovatief gedrag en risiconemend gedrag (Stam et al., 2013; Preenen et al., 2013) . De lerende houding en experimenteer ruimte die bij FM centraal staan, stimuleren deze elementen zowel op het niveau van de medewerkers als de faciliteiten vanuit het management. Het feedbacksysteem kan gezien worden als een belangrijke basis voor intrapreneurship, omdat scherper wordt welke vormen van gedrag - routinematig, proactief, innovatief of risiconemend - effectief zijn.

In het algemeen zijn de concepten sociale innovatie en intrapreneurship verwant aan elkaar. Dhondt (2012: p. 15) beschrijft in zijn oratie sociale innovatie als de organisatorische ruimte en ruggensteun voor intrapreneurship.

3.2 Stappen bij het opzetten van een *expert judgement*

Het uitvoeren van een *expert judgement* vergt de volgende stappen:

1. identificatie van een model waarin de vorm van de relaties tussen de verschillende variabelen gespecificeerd is;
2. selectie van relaties en/of verdelingskenmerken van variabelen waarvoor expertmening gaat plaatsvinden;
3. selectie expertgroep;
4. opstellen vragenlijst;
5. sessie(s) met experts: toelichting expertmening en afnemen vragenlijst;
6. verwerking van de resultaten.

Bij de vertaling van de Orionis-case hebben we de volgende stappen kunnen uitvoeren:

- ad 1. en 2. We proberen een model te specificeren.
- ad 3. Selectie van een expertgroep.
- ad 4. We hebben een concept vragenlijst opgesteld.

Bij de uitwerking zijn we tegen enkele belangrijke operationaliseringsproblemen aangelopen die maken dat we feitelijk *expert judgement* beperkt bruikbaar kunnen maken voor ex ante impact assessments voor sociale innovatie cases. We beschrijven de stappen die we hebben genomen en de problemen die we zijn tegengekomen.

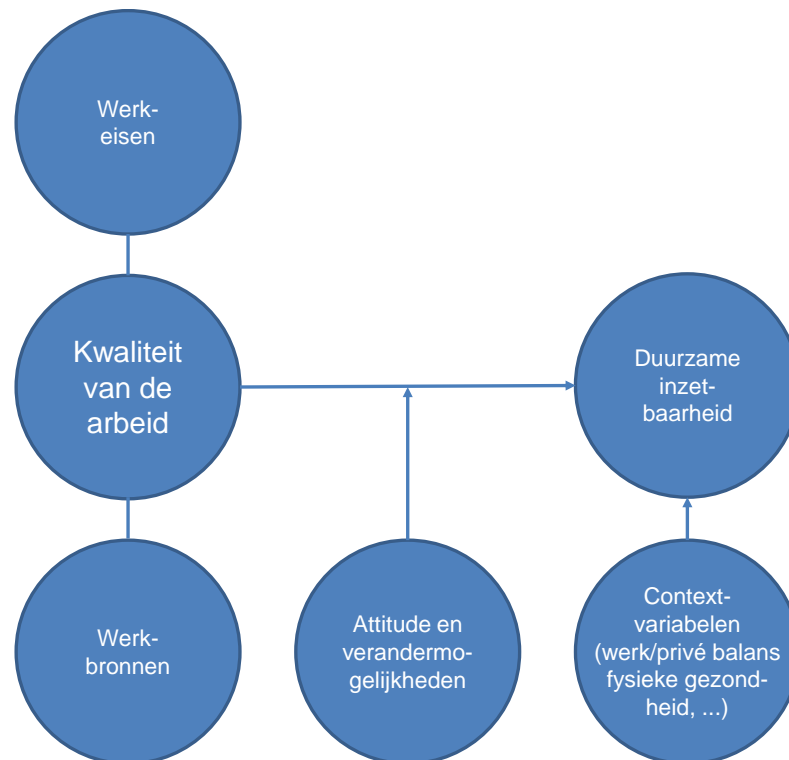
3.3 Lukt het ons een model van de sociale innovatie te specificeren?

3.3.1 *Modellering functioneel meten en duurzame inzetbaarheid heeft moeilijkheden*

Als kader voor het meten van de relatie tussen functioneel meten en duurzame inzetbaarheid (DI) is een conceptueel model opgesteld, dat is geïnspireerd op het Job Demand - Job Control model (Karasek, 1979; Karasek & Theorell, 1990). Zoals in kader 3.1 geschetst heeft het JD-JC model belangrijke componenten geleverd voor modellen op het gebied van sociale innovatie.

Het JD-JC model betreft objectieve kenmerken uit de organisatie/het werk. We voegen nog een subjectief element toe aan het DI-model. Van Dijk et al. (1990) geven aan dat ook co-

pingmogelijkheden en attitude van werknemers van belang zijn. Zie voor het model figuur 3.1 en bijlage 1.



Figuur 3.1 Specificatie van het model van sociale innovatie in de case Orionis

Duurzame inzetbaarheid wordt in het bovenstaand model gezien als een resultante van de kwaliteit van de arbeid, attitude en verandermogelijkheden van de medewerker en contextvariabelen zoals de fysieke gezondheid en de werk-privé balans. De interventie (functioneel meten bij Orionis) grijpt aan op de werkeisen, op de werkbronnen en op de attitudes en verandermogelijkheden.

Als we het model verder willen uitwerken en deels willen kwantificeren op basis van vragen die in een *expert judgement* sessie kunnen worden vastgelegd met het oogmerk dat de resultaten generaliseerbaar zijn, stuiten we op een aantal moeilijkheden.

Moeilijkheid 1: beperkte generaliseerbaarheid resultaten

Het model dat we hebben uitgewerkt, blijkt in de praktijk nog weinig hanteerbaar te zijn voor een verdere specificatie. Het concept sociale innovatie blijft ambigu en diffuus: er zijn meerdere definities in omloop die niet volledig overlappen. Een vervolprobleem is dat we niet in staat zijn om de toepassing los te maken van de specifieke context. Daardoor is het moeilijk om in te schatten in hoeverre een door het projectteam te kiezen casus van “sociale innovatie” genoeg overeenkomsten heeft met andere sociale innovaties, laat staan dat er iets gezegd kan worden over hoe transfereerbaar de lessen zijn die in het verband van onze casus geleerd worden. Door de contextgebondenheid wordt het gewenste principe van generaliseerbaarheid van resultaten sterk ingeperkt.

Moeilijkheid 2: bij selectie van de expertgroep

Parallel aan het modelleren van het model van sociale innovatie zijn we aan de slag gegaan met het formeren van een expertgroep (stap 3 uit de onderzoeksaanpak). Voor stap 3, de selectie van experts, zijn de volgende experts overwogen:

1. experts van buiten Orionis;
2. leidinggevenden binnen Orionis;
3. medewerkers van Orionis;
4. staf (bijvoorbeeld de bedrijfsarts) van Orionis.

Overwegingen om experts al dan niet te selecteren, zijn:

- › mate waarin de expert goed zicht heeft op de sociale innovatie;
- › neutraliteitsgehalte van de expert (belang bijvoorbeeld of de mate waarin de expert een subjectief referentiekader meebrengt ten aanzien van het onderwerp waarover hij wordt bevraagd);
- › kosten voor het betrekken van de expert (eventuele reiskosten, overnachtingskosten).

Het bleek dat de mate waarin de expert goed zicht heeft op de sociale innovatie bij Orionis, het neutraliteitsgehalte negatief beïnvloedt. Sterker nog: een expert die kan beoordelen wat het effect van de sociale innovatie is, maakt tevens deel uit van de organisatie en brengt dus een belang en zijn persoonlijk referentiekader mee.

3.4 Dilemma's als gevolg van deze moeilijkheden

Als gevolg van deze moeilijkheden hebben we als projectteam beoordeeld wat we feitelijk nog met *expert judgement* bij sociale innovatie zouden moeten aanvangen. We zien de volgende drie dilemma's.

Dilemma 1. Neutraliteitsgehalte versus expertgehalte

Een sociale innovatie is per definitie maatwerk, op zijn minst aangepast aan lokale omstandigheden van de organisatie en vaak ook aan persoonlijke eisen/wensen van medewerkers. Dat impliceert dat experts vaak moeilijk buiten de organisatie te vinden zijn, omdat mensen van buiten de organisatie een te beperkt zicht hebben op de sociale innovatie en dus niet kunnen inschatten wat de effecten hiervan zijn. Echter: medewerkers van de organisatie hebben belangen. Dat beperkt de validiteit van de schattingen die zij afgeven.

Dit dilemma speelt ook bij traditioneel onderzoek. Een manier om hiermee om te gaan, die wellicht ook in de context van *expert judgement* kansen biedt, is het opnemen van een schaal die "meet" in welke mate mensen neutraal antwoorden, zoals een Marlowe-Crowne Social Desirability Scale (zie Wikipedia) of een correctie voor "negative affectivity" (Van den Heuvel et al., 2005). Dergelijke manieren zijn (nog) niet ontwikkeld voor *expert judgement* met de methode van Cook.

Dilemma 2. Expertgehalte versus subjectief referentiekader

Zoals hierboven geschetst komen de experts die het beste zicht hebben op de interventie en de gevolgen, uit de organisatie. Maar daarmee ondervinden zij de gevolgen van de interventie persoonlijk. Als hen wordt gevraagd naar uitspraken over hun collega's, is de kans groot dat ze dergelijke uitspraken doen op basis van een projectie op henzelf. Daarmee beantwoorden ze dus eigenlijk een vraag over zichzelf, terwijl wordt gevraagd naar de populatie. Een mogelijke aanpak om dit dilemma te ontstijgen is in de vragen die aan de experts worden gesteld zowel vragen op te nemen over de expert zelf als over de populatie. Daarmee

wordt een expliciet onderscheid gemaakt tussen de eigen subjectieve beleving en die van de collega's.

Dilemma 3. Concreetheid versus externe validiteit

Het derde dilemma hangt samen met het punt dat, naarmate een sociale innovatie wint aan concreetheid, de contextgebondenheid toeneemt. Wanneer bijvoorbeeld lokaal maatwerk en participatie belangrijke verklaringen zijn voor het succes van functioneel meten, is het dan wel mogelijk om uitspraken over deze specifieke sociale innovatie bij deze specifieke organisatie te generaliseren naar de wereld buiten die organisatie?

In de context van *expert judgement*: als het concept van sociale innovatie impliceert dat experts alleen te vinden zijn binnen de organisatie, hoe kunnen uitspraken van deze experts over deze sociale innovatie bij hun organisatie dan gegeneraliseerd worden naar andere organisaties? Met andere woorden: hoe harder het nodig is dat experts over heel specifieke kennis beschikken over lokale omstandigheden (die lastig observeerbaar zijn voor buitenstaanders), hoe minder de generaliseerbaarheid van hun uitspraken wordt.

De dilemma's hebben als gevolg dat *expert judgement* een minder valide instrument wordt om uitspraken te doen over het effect van een sociale innovatie.

4 Conclusies en discussie

4.1 Conclusies

We hadden als doel om *expert judgement* toe te passen bij cases sociale innovatie. Indien we met onze test er in zouden slagen de methode toe te passen, dan zouden we aan stakeholders een instrument in handen leveren om op voorhand (*ex ante*) de verwachte impact van sociale innovaties in te schatten. We hebben daarvoor een case geselecteerd en geprobeerd om de regel der kunst van een *expert judgement* toe te passen. Onze ervaring is dat dergelijke cases niet te modelleren zijn volgens de eisen van een *expert judgement*.

Op basis van de resultaten trekken we de volgende conclusies.

4.1.1 *Expert judgement kan maar beperkt worden toegepast in het domein van sociale innovatie*

Expert judgement heeft slechts zin als problemen niet contextgebonden zijn. De met sociale innovatie samenhangende contextgebondenheid maakt het moeilijk om te abstraheren en inschattingen te maken over mogelijke impact. Dat wil zeggen dat *expert judgement* beperkt dient te worden tot gebieden/problemen waarvoor de met sociale innovatie samenhangende contextgebondenheid niet geldt. Het gaat dan om factoren als bijvoorbeeld inflatievoeten of energieprijzen.

4.1.2 *Als input voor de expert judgement is een concreet model nodig met daarin voor experts herkenbare concepten*

Dat wil zeggen dat het model, dat wordt gehanteerd, eenduidig afbakent welke relaties er toe doen en tussen welke factoren die relaties lopen. De factoren dienen herkenbaar te zijn. De relaties en factoren moeten dus duidelijk te vertalen zijn naar empirische verschijnselen en er dienen experts te zijn die daar ervaring mee hebben.

4.2 Discussie

4.2.1 *Hoe en onder welke voorwaarden kan expert judgement worden toegepast voor het bepalen van de effecten van een sociale innovatie?*

Kader 4.1 Beperking *expert judgement* in de context van sociale innovatie 1

De beperkte generaliseerbaarheid van de resultaten (dilemma 3).

Wijze waarop hiermee kan worden omgegaan

Aan deze beperking valt lastig te ontkomen. *Expert judgement* vraagt om concrete concepten waaraan de experts hun kennis/ervaring kunnen relateren. Hoe concreter de sociale innovatie wordt, hoe groter de context gebondenheid en daarmee hoe kleiner de generaliseerbaarheid.

Kader 4.2 Beperking *expert judgement* in de context van sociale innovatie 2

Het gevaar dat experts hun eigen, subjectieve referentiekader zwaar laten meewegen bij het geven van schattingen over de populatie (dilemma 2).

Wijze waarop hiermee kan worden omgegaan

Formuleer vragen over de persoon zelf, naast vragen over de populatie.

Kader 4.3 Beperking *expert judgement* in de context van sociale innovatie 3

Het belang dat experts hebben bij de uitkomsten van de analyse weegt zwaar mee (dilemma 1).

Wijze waarop hiermee kan worden omgegaan

Ook aan deze beperking valt lastig te ontkomen, omdat hij samenhangt met degenen die als “expert” gezien worden.

Wel zou onderzocht kunnen worden in hoeverre het mogelijk is om met slim gekozen “controle”vragen, die naast een objectieve maatstaf gelegd kunnen worden, te corrigeren voor de mate waarin experts hun eigen belang laten meewegen.

4.2.2 ***Aan welke criteria moet een casebeschrijving van sociale innovatie voldoen, wil die zich lenen voor analyse met expert judgement?***

Die input in een *expert judgement* bestaat uit een concreet model waarin de relaties tussen de verschillende aspecten rond de sociale innovatie, de effecten en de impact van overheidsbeleid in functionele vorm is beschreven. Pas als zo iets op te leveren is, dan kan de methode worden toegepast. Verder moeten er experts zijn met ervaring met de verschillende relaties en/of concepten in het model, om de parameters in het model te schatten.

De belangrijkste criteria zijn concreetheid en herkenbaarheid van de concepten. De experts moeten kunnen begrijpen waarop ze bevraagd worden en dat op een eenduidige wijze kunnen relateren aan hun ervaringen en/of kennis.

Literatuur

Albert, I., Donnet, S., Guihenneuc-Jouyauz, C., Low-Choy, S., Mengersen, K., & Rousseau, J. (2012). Combining expert opinions in prior elicitation. *Bayesian Analysis*, 7(3), 503-532.

Cooke., R. M. (1991). *Experts in Uncertainty*. New York: Oxford University Press.

Cooke, R. M., Goossens, L. H. J., & Kraan, B. C. P. (1999). *Probabilistic Accident Consequence Uncertainty Assessment Procedures Guide Using Expert Judgement*. Report prepared under contract No. ETNU-CT93-0104-NL for the Commission of European Communities Directorate-general XI (Environment and Nuclear Safety) Directorate D (DG 11 CCJH), Published as EUR 18820, June 1999.

Dhondt, S. (2012). Sociale Innovatie als voorwaarde voor een duurzame economie. Oratie KU Leuven.

Gächter, S., Nosenzo, D., & Sefton, M. (2012) The impact of Social Comparisons on Reciprocity. *Scand. J. of Economics*, 114(4), 1346–1367.

Gustafson, D., Shulka, R., Delbecq, A., & Walster, A. (1973). A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups and nominal groups. *Organizational Behaviour and Human Performance*, 9(2), 280-291.

Helmer, O. (1966). *Social Technology*. New York: Basic Books.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: implications for job redesign. *Administrative science quarterly*, 24, 285-308.

Karasek, R., & Theorell, T. (1990). *Healthy work, stress, productivity, and the reconstruction of working life*. New York: Basic Books.

Kube, S., Maréchal, M. A., & Puppe, C. (2012). The Currency of Reciprocity: Gift Exchange in the Workplace. *American Economic Review*, 102(4), 1644-1662.

Morales-Nápoles, O., Hanea, A. M., & Worm D. T. H. (2014a). Experimental results about the assessments of conditional rank correlations by experts: Example with air pollution estimates. In R. D. J. M. Steenbergen, P. H. A. J. M. van Gelder, S. Miraglia & A. C. W. M. Vrouwenvelder (Eds), *Safety, Reliability and Risk Analysis: Beyond the Horizon* (pp. 1359-1366). London: Taylor & Francis Group.

Morales-Nápoles, O., et al. (2014b). *Expert Judgement study on Tail Dependency in future Rain Showers in the Netherlands*. Delft: TNO. TNO ETP project Games2R 2014 (in progress).

Pot, F. (2012). Sociale innovatie: historie en toekomstperspectief. *Tijdschrift voor Arbeidsvraagstukken*, 28(1), 6-21.

Preenen, P. T. Y., Kraan, K., Van der Meulen, F., & Dhondt, S. (2013). *Factoren van invloed op intrapreneurship: Een verkennende studie*. Hoofddorp: TNO.

Stam, E., Bosma, N., Van Witteloostuijn, A., De Jong, J., Bogaert, S., Edwards, N., & Jaspers, F. (2012). *Ambitious Entrepreneurship. A review of the academic literature and new directions for public policy*. The Hague: Advisory Council for Science and Technology Policy (AWT).

Van den Heuvel, S. G., Van der Beek, A. J., Blatter, B. M., Hoogendoorn, W. E., & Bongers, P. (2005). Psychosocial work characteristics in relation to neck and upper limb symptoms. *Pain*, 114(1-2), 47-53.

Van Dijk, F.J.H., Van Dormolen, M., Kompier, M. A. J., & Meijman, T.F. (1990). Herwaardering model belasting-belastbaarheid. *Tijdschr Soc Gezondheidsz*, 68(1), 3-10.

Bijlage 1 Conceptualisering van duurzame inzetbaarheid en vragen voor een *expert judgement* sessie, versie 13-5-2014

Expert judgement richt zich op de volgende vraag:

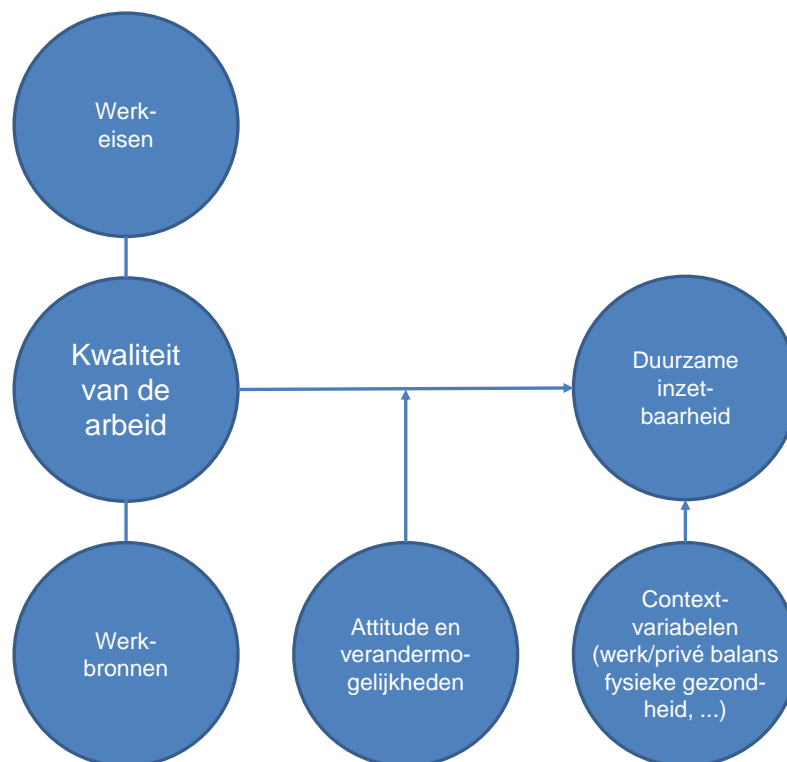
Wat is het effect van functioneel meten bij Orionis op de duurzame inzetbaarheid van medewerkers?

Functioneel meten is een TNO-interventie bij Orionis, gericht op:

1. creëren van een informatiesysteem dat bruikbare feedback geeft op het resultaat van het werk van medewerkers, en
2. creëren van een lerende attitude bij individuele en groepen medewerkers.

Deze bijlage gaat in op hoe we duurzame inzetbaarheid definiëren en hoe we dat operationaliseren. Dat is een eerste stap om te komen tot vragen die gesteld kunnen worden tijdens een *expert judgement* sessie.

We hanteren het volgende conceptuele model dat de duurzame inzetbaarheid van een medewerker verklaart.



We zien duurzame inzetbaarheid dus als een resultante van de kwaliteit van de arbeid, attitude en verandermogelijkheden van de medewerker en contextvariabelen zoals de fysieke gezondheid en de werk-privé balans.

Onze interventie bij Orionis grijpt aan op de werkeisen, op de werkbronnen en op de attitudes en verandermogelijkheden. We operationaliseren naast het concept duurzame inzetbaarheid ook deze concepten.

We hanteren verder de volgende definities:

- › **kwaliteit van de arbeid:** een goede match tussen werkeisen en werkbronnen (JDC-model);
- › **werkeisen:** eisen die het werk aan de medewerker stelt, zoals vakkennis, taakeisen of eisen op mentaal en/of emotioneel vlak;
- › **werkbronnen:** de regelmogelijkheden die medewerkers hebben om hun werk goed te kunnen doen, zoals autonomie/regelmogelijkheden, sociale en functionele steun, leer-mogelijkheden, etc.;
- › **attitude en verandermogelijkheden:** kenmerken op het niveau van de medewerker, zoals het belang dat wordt gehecht aan leren en ontwikkelen, stressbestendigheid, flexibiliteit, copingstijl, etc. Dit concept wordt ook wel persoonlijke bronnen genoemd;
- › duurzame inzetbaarheid is een combinatie van de intentie (willen) en de mogelijkheid (kunnen) om productief door te werken bij Orionis.

Contextvariabelen zijn factoren die invloed hebben op duurzame inzetbaarheid, maar niet door de interventie worden beïnvloed.

Operationalisering van de concepten

Bij de operationalisering is een keuze gemaakt uit verschillende items/vragen die in een aantal vragenlijsten (NEA, DIX, STREAM, WEBA) worden gebruikt om een concept te meten. Het criterium is of te verwachten valt dat de interventie iets verandert.

De vorm van de vragen kan verschillen tot het geven van een rapportcijfer (oordeel), het aangeven in hoeverre men het eens is met een stelling of het aangeven van de frequentie waarmee iets voorkomt. De vorm van de vraag is hier kortheidshalve weggelaten.

Werkeisen

- › Is uw functie het afgelopen jaar uitgebreid? (Ja, breder geworden; ja, moeilijkere taken gekregen, Nee) (DIX).
- › Mate waarin het werk intensief nadenken vereist (WEBA).

Werkbronnen

- › Mate waarin u feedback krijgt op de kwaliteit van uw werk (WEBA).
- › Heeft u invloed op de beslissingen van uw werkteam/taakgroep/afdeling?
- › Fouten worden in onze organisatie besproken om ervan te leren (DIX).
- › Verschillen van mening leiden in deze organisatie tot zinvolle gesprekken waar we met elkaar beter van worden (DIX).

Attitude en verandermogelijkheden

- › Als ik problemen heb met mijn werk dan zoek ik hulp (DIX).
- › Ik leer het liefst in de praktijk/tijdens het werk (DIX).
- › Ik ben goed in staat om mijn kennis en vaardigheden bij te houden (DIX).
- › Ik kan goed omgaan met spanning.

Duurzame inzetbaarheid

- › Als ik 's morgens opsta, heb ik zin om aan het werk te gaan (STREAM).
- › Zou u over 5 jaar nog bij deze organisatie willen werken (à la WEBA)?
- › Heeft u over 5 jaar meer moeite met de zwaarte van het werk (DIX)?
- › Presteert u over 5 jaar nog net zo goed als nu (DIX)?
- › Aan het einde van de werkdag voel ik me moe.

Mogelijke seed-variabele (variabelen die zich lenen voor “controle vragen”):

- › aantal verzuimdagen afgelopen 12 maanden;
- › frequentie verzuim afgelopen 12 maanden.