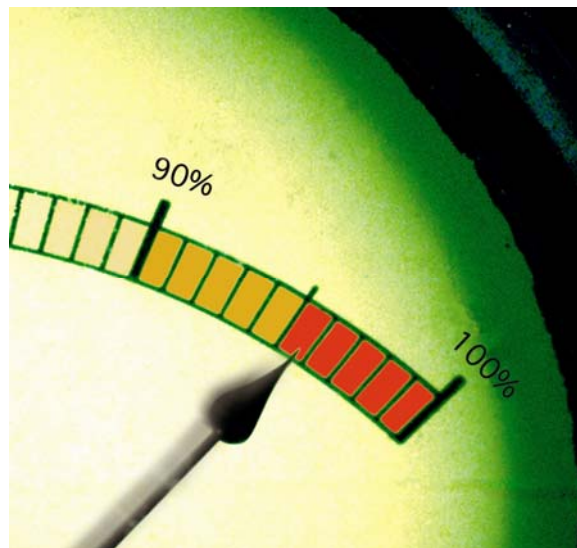




Panteia

Research to Progress

Research voor Beleid | EIM | NEA | IOO | Stratus | IPM



Naar dezelfde loonwaarde?

Een vervolgonderzoek naar de vergelijkbaarheid van loonwaardemethoden

Eindrapport

Een onderzoek in opdracht van het Ministerie van Sociale Zaken
en Werkgelegenheid

Suzanne de Visser, Sonja van der Kemp, Martine Van Ommeren
en Mirjam Engelen

Projectnummer: BA4007

Zoetermeer, 5 oktober 2012

De verantwoordelijkheid voor de inhoud berust bij Panteia/Research voor Beleid. Het gebruik van cijfers en/of teksten als toelichting of ondersteuning in artikelen, scripties en boeken is toegestaan mits de bron duidelijk wordt vermeld. Vermenigvuldigen en/of openbaarmaking in welke vorm ook, alsmede opslag in een retrieval system, is uitsluitend toegestaan na schriftelijke toestemming van Panteia/Research voor Beleid. Panteia/Research voor Beleid aanvaardt geen aansprakelijkheid voor drukfouten en/of andere onvolkomenheden.



Voorwoord

In mei 2011 voerde Panteia/Research voor Beleid in opdracht van het Ministerie van SZW een experiment loonwaarde uit. Er zijn verschillende methoden voor loonwaardebepaling op de markt, en eindgebruikers als gemeenten, SW-bedrijven, werkgevers en werknemers moeten ervan op aan kunnen dat de keuze voor een bepaalde methodiek geen invloed uitoefent op de hoogte van de vastgestelde loonwaarde. Doel van het experiment was daarom de vergelijkbaarheid van de uitkomsten van de verschillende beschikbare methoden voor loonwaardebepaling te onderzoeken.

Het experiment liet zien dat de methoden zowel qua inhoud en uitvoering als qua uitkomsten dusdanig van elkaar verschilden dat deze vergelijkbaarheid niet aan eindgebruikers gegarandeerd kon worden. Het bleek daarom nodig richtlijnen te ontwikkelen om de loonwaardebepaling te uniformeren.

In een vervollexperiment is het effect van deze richtlijnen op de vergelijkbaarheid van de loonwaardemethoden onderzocht. Voorliggend rapport doet verslag van de uitkomsten van dit nieuwe experiment en geeft daarmee inzicht in de vergelijkbaarheid van de aan de richtlijnen aangepaste methoden.

Een opmerking bij het interpreteren van de resultaten is vooraf op zijn plaats: in het experiment onderzoeken we in hoeverre de uitkomsten van de aangepaste methoden vergelijkbaar zijn. Op grond van het experiment kunnen echter geen uitspraken worden gedaan over welke methoden *het meest* (of juist *het minst*) valide en/of betrouwbaar zijn. Het experiment is nadrukkelijk niet bedoeld om een dergelijke rangorde aan te brengen. Dit is op grond van het experiment ook niet mogelijk.

Suzanne de Visser
Projectleider



Inhoudsopgave

1	Inleiding	7
1.1	Loonwaardebepaling: een nieuw experiment	7
1.2	Opzet experiment	8
1.3	Opbrengst van het onderzoek: interpretatie resultaten	10
2	Methoden aangepast	13
2.1	Bestanddelen	13
2.2	Norm	15
2.3	Aanpassingen methoden	16
3	Uitkomsten experiment	17
3.1	Overzicht uitkomsten	17
3.2	Resultaten beschouwd	17
3.3	Vergelijking tussen methoden	18
3.4	Conclusie	20
4	Verschillen verklaard	21
4.1	Verklaring 1. Niveau methode: berekeningswijze en weging	21
4.2	Verklaring 2. Niveau methode: inhoud methoden	24
4.3	Verklaring 3. Niveau uitvoering: vaststellen norm	26
4.4	Verklaring 4. Niveau uitvoering: vaststellen weging	27
4.5	Verklaring 5. Niveau uitvoering: rol van de uitvoerder	27
4.6	Conclusie	29
5	Conclusies en aanbevelingen	31
5.1	Bescheiden succes, maar...	31
5.2	Verdere uniformering methoden noodzakelijk	31
5.3	Niet alles op niveau van methoden op te lossen	31
5.4	Tot slot: validering	34
	Bijlage 1 Overzicht resultaten experiment 1	35



1 Inleiding

1.1 Loonwaardebepaling: een nieuw experiment

Loonwaardebepaling: het belang van validiteit en betrouwbaarheid

Loonwaarde en loonwaardebepaling staan volop in de (politieke) belangstelling. Werknemers met arbeidsbeperkingen moeten – ook wanneer zij niet in staat zijn zelfstandig het volledige wettelijk minimumloon te verdienen – in toenemende mate aan de slag bij reguliere werkgevers. Dat betekent dat werkgevers gecompenseerd moeten worden voor dat deel dat deze werknemers verminderd productief zijn: een werkgever zal in beginsel immers alleen willen betalen voor de door de werknemer geleverde arbeidsprestatie.

Het is belangrijk dat zo goed en objectief mogelijk wordt vastgesteld hoe groot de verminderde arbeidsprestatie van een werknemer in een specifieke functie precies is. Inzicht hierin kan worden verkregen door het vaststellen van de **loonwaarde** van een werknemer. De afgelopen jaren zijn verschillende methoden ontwikkeld waarmee de loonwaarde van een werknemer op een geobjectiveerde manier kan worden bepaald. Het is daarbij voor de afnemers van en belanghebbenden bij de loonwaardebepaling (gemeenten, SW-bedrijven, werknemers en werkgevers) van belang dat zij ervan op aan kunnen dat de keuze voor een bepaalde methodiek en/of uitvoerder geen invloed uitoefent op het resultaat van de loonwaardebepaling. Het zou voor de uitkomst van de loonwaardebepaling niet moeten uitmaken wie de loonwaardebepaling uitvoert en met welke methode dit gebeurt. Hiermee rijst de vraag naar de **validiteit en betrouwbaarheid** van de methoden die nu op de markt zijn.

Een nieuw experiment

Om meer zicht te krijgen op deze aspecten van validiteit en betrouwbaarheid van methoden voor loonwaardebepaling, heeft Panteia/Research voor Beleid in opdracht van het Ministerie van SZW eerder een experiment loonwaarde uitgevoerd.¹ In dit experiment is de vergelijkbaarheid van de uitkomsten van zeven loonwaardemethoden onderzocht, door van negen werknemers met een arbeidsbeperking de loonwaarde vast te stellen met behulp van verschillende methoden. Door de uitkomsten van deze metingen voor elke werknemer naast elkaar te zetten, ontstond inzicht in de vergelijkbaarheid van de uitkomsten van de methoden.

Deze vergelijkbaarheid bleek beperkt: het experiment liet een flink aantal grote verschillen tussen de methoden zien. Deze verschillen konden niet worden toegeschreven aan één of enkele methoden in het bijzonder: vrijwel elke combinatie van twee methoden liet grote verschillen zien. We konden op basis van dit experiment dan ook *niet* concluderen dat de methoden valide en betrouwbaar waren. De uitkomsten van *alle* methoden verschilden (sterk) van elkaar. Wel bood het experiment verschillende aanknopingspunten om de vergelijkbaarheid van de methoden te verbeteren. Een deel van deze aanknopingspunten is door het ministerie gebruikt om richtlijnen te ontwikkelen die moesten leiden tot een uniformering van de loonwaardebepaling. Deze richtlijnen en aanpassingen zijn beschreven in hoofdstuk 2. De onderzochte methoden zijn aan deze richtlijnen aangepast. In een **nieuw**

¹ Research voor Beleid (2011), *Van deelnemer naar werknemer. Loondispensatie en loonwaarde in de praktijk*.

experiment loonwaarde onderzoeken we nu of en in hoeverre deze aanpassingen inderdaad leiden tot minder grote verschillen in de uitkomsten van methoden en daarmee tot een betere validiteit en betrouwbaarheid.

Doelstelling en vraagstelling

Het Ministerie van SZW wil met dit nieuwe experiment zicht krijgen op het effect van het aanpassen en uniformeren van methoden voor loonwaardebepaling op basis van door het ministerie gestelde richtlijnen.

De vraagstelling van dit nieuwe experiment is driedelig:

- A. Leidt het aanpassen van de methodieken op basis van de door het ministerie gestelde richtlijnen tot vergelijkba(a)r(d)e(r) uitkomsten tussen de methodieken?
- B. Kunnen we op basis hiervan concluderen dat de methoden na aanpassing aan deze richtlijnen inderdaad valide en betrouwbaar zijn?
- C. Wat is hiervoor de verklaring?

Doel van het experiment is uitdrukkelijk *niet* een rangorde in de methoden aan te brengen of te beoordelen wat de meest betrouwbare of valide methode is. Dit is op basis van de gekozen onderzoeksopzet ook niet mogelijk.

De begrippen validiteit en betrouwbaarheid voeren verder dan vergelijkbaarheid alleen. De experimenten loonwaarde zijn daarom onderdeel van een breder onderzoek naar de validiteit en betrouwbaarheid van de loonwaardemethoden in het kader van de evaluatie van de Tijdelijke wet pilot loondispensatie. De resultaten hiervan worden gepresenteerd in de halfjaarlijkse voortgangsrapportages over deze evaluatie.

1.2 Opzet experiment

Uitgangspunten bij het onderzoek

Het basisprincipe van het experiment is dat van een werknemer met een arbeidsbeperking de loonwaarde wordt vastgesteld met behulp van verschillende methoden. Door de uitkomsten van deze metingen naast elkaar te zetten, ontstaat inzicht in de vergelijkbaarheid van de uitkomsten van de methoden. We doen dit voor verschillende werknemers. Op deze manier is objectief vast te stellen in hoeverre de aanpassingen van de methoden inderdaad leiden tot vergelijkbare uitkomsten tussen de verschillende methoden.

Bij het experiment zijn we uitgegaan van de volgende uitgangspunten:

1. Deelnemers met een arbeidsbeperking

Deelnemers aan het experiment zijn werknemers met een arbeidsbeperking waarvan de loonwaarde naar verwachting lager is dan 100% van het wettelijk minimumloon (WML). De keuze voor deelnemers met een arbeidsbeperking is inhoudelijk gemotiveerd: de loonwaarde van reguliere werknemers (dat wil zeggen: werknemers zonder arbeidsbeperking) zal naar verwachting met vrijwel elke methode op of rond 100% uitkomen. Er zijn immers geen arbeidsbeperkingen die leiden tot een verminderde arbeidsprestatie. Alle methoden hanteren een bovengrens van 100% loonwaarde, waardoor het niet mogelijk is de loon-

waarde verder te differentiëren. Eventuele verschillen in de uitkomsten van de loonwaardebepaling zullen daarom eerder zichtbaar worden bij cliënten met een loonwaarde van minder dan 100% WML. Dit maakt het makkelijker de vergelijkbaarheid van de loonwaardebepalingen te bepalen. Bij het meten van de loonwaarde van een reguliere werknemer ontstaat bovendien een methodologisch probleem: volgens het principe van loonwaardebepaling is de reguliere werknemer de norm waartegen de arbeidsprestatie moet worden afgezet.¹ De loonwaardemeting zou zich in dit geval dus feitelijk toespitsen op het vergelijken van de werknemer met zichzelf: een vergelijking van regulier met regulier.

2. Deelnemers met een werkplek en een werkgever

De deelnemers aan het experiment hebben daarnaast een (reguliere) werkplek. De loonwaarde is immers per definitie situationeel en kan alleen in relatie tot een concrete functie worden vastgesteld. Het is niet mogelijk de loonwaarde vast te stellen buiten een specifieke werkplek om.² Bovendien is voor vrijwel alle methodieken voor de loonwaardebepaling ook input van de werkgever nodig. Het is dus noodzakelijk dat deelnemers een werkgever hebben, die bovendien bereid is aan het experiment mee te doen.

3. Loonwaarde op een bepaald tijdstip: metingen kort na elkaar

Loonwaarde is niet statisch: de loonwaarde van een werknemer kan in de loop der tijd veranderen. Factoren als opleiding en ontwikkeling, maar ook veranderingen in de organisatie, kunnen ertoe leiden dat de loonwaarde van een werknemer mettertijd stijgt of daalt. Om de vergelijkbaarheid van de uitkomsten te waarborgen, is het daarom van belang dat de metingen van dezelfde werknemer binnen korte tijd (bij voorkeur binnen een week) na elkaar plaatsvinden. Dit is gebeurd.

4. Optimum belasting en voldoende waarnemingen

Deelname aan het experiment is relatief belastend voor werknemer, werkgever en uitvoerder van de loonwaardebepaling. Een loonwaardebepaling is – afhankelijk van de methode – behoorlijk tijdsintensief en vraagt de nodige openheid en (psychologische) inspanning van zowel werkgever als werknemer. De korte doorlooptijd van de metingen, waarbij de loonwaardebepalingen kort na elkaar worden uitgevoerd, vergroot de belasting. Het is daarom zaak het aantal loonwaardebepalingen per werknemer zoveel mogelijk te beperken. Tegelijkertijd is het noodzakelijk het aantal waarnemingen niet teveel te beperken, om voldoende basis voor de vergelijking te kunnen garanderen. Dit maakt het mogelijk de uitkomsten van de verschillende methodieken op betrouwbare wijze rechtstreeks en kruislings met elkaar te vergelijken om afwijkende patronen op het spoor te komen. In de opzet van het experiment is daarom gestreefd naar een optimum tussen het beperken van de belasting voor deelnemende cliënten en uitvoerders enerzijds, en het garanderen van voldoende waarnemingen om betrouwbare conclusies te kunnen trekken anderzijds.

¹ Kemper, R., L. de Ruig en S. de Visser (2009), *Aan de slag met loonwaardemeting. Inventarisatie van methoden om loonwaarde te meten.*

² Idem.

Praktische aanpak

In het experiment zijn de zes methodieken, die worden gebruikt binnen de pilot loondispen-
satie, meegenomen. Met deze methodieken is al enige praktische ervaring opgedaan in de
pilot. Het betreft:

- Actíva Loonwaardemethodiek
- Arbolabmethode
- Dariuz
- Competensys Loonbalans
- Matchcare
- VTA-methodiek van Pantar

De UWV Loonwaardemethodiek is hieraan toegevoegd omdat deze bij de Wajong veel wordt
toegepast¹.

Om deelnemers aan het experiment te werven, is contact gezocht met verschillende ge-
meenten en SW-bedrijven. Hen is gevraagd de deelnamebereidheid van hun cliënten met
een (naar verwachting) beperkte loonwaarde én een werkplek te peilen. Hierbij was van be-
lang dat ook de betreffende werkgevers bereid waren aan het experiment mee te werken.
Na de nodige inspanningen zijn op deze manier negen werknemers en hun werkgevers uit
verschillende sectoren bereid gevonden aan het experiment deel te nemen. Van elk van de-
ze werknemers is de loonwaarde op een bepaald tijdstip bepaald met behulp van twee tot
vijf verschillende loonwaardemethoden. Door voor verschillende werknemers deels dezelfde
methoden te gebruiken, kunnen de uitkomsten ook kruislings met elkaar worden vergele-
ken.

Om de invloed van uitvoeringsverschillen te minimaliseren en zo de betrouwbaarheid van
de resultaten te waarborgen, is de ontwikkelaars van de methodieken gevraagd de loon-
waardebepalingen binnen het experiment zelf op locatie uit te voeren. Bij de loonwaarde-
bepaling was ook steeds een onderzoeker van Panteia/Research voor Beleid aanwezig om
de loonwaardebepaling te observeren. De observatieverslagen bieden aanknopingspunten
om in de analysefase eventuele verschillen in de vastgestelde loonwaardes te kunnen ver-
klaren.

1.3 Opbrengst van het onderzoek: interpretatie resultaten

Het resultaat van dit loonwaarde-experiment geeft een eerste inzicht in de vergelijkbaar-
heid van de volgens de richtlijnen aangepaste methoden. Theoretisch zijn hierbij de vol-
gende uitkomsten mogelijk:

- a. De uitkomsten van de loonwaardebepalingen verschillen weinig van elkaar. Wanneer de
loonwaardebepalingen van de verschillende methoden voor dezelfde cliënten dicht bij
elkaar liggen, geeft dit experiment voldoende grond uitspraken te doen over de be-
trouwbaarheid van de methoden.

¹ Inmiddels zijn er andere methoden op de markt om loonwaarde te bepalen.

- b. De uitkomsten van de loonwaardebepaling verschillen van elkaar. In dit geval kunnen we niet zeggen dat de methoden betrouwbaar en/of valide zijn, maar ook niet welke methode onbetrouwbaar of niet-valide is: het aantal waarnemingen is hiervoor te klein. We concluderen dan dat de methoden kennelijk niet hetzelfde meten en dat de resultaten niet vergelijkbaar zijn en geven hiervoor verklaringen.

Op grond van het experiment kunnen dan ook *geen* uitspraken worden gedaan over welke methoden het meest, of juist het minst valide en/of betrouwbaar zijn. Het experiment is nadrukkelijk niet bedoeld om een dergelijke rangorde aan te brengen. Bij het interpreteren van de resultaten dient hier rekening mee gehouden te worden: een methodiek die systematisch van andere methodieken afwijkt, is daarmee niet per definitie minder valide of betrouwbaar dan de andere methodieken. Mogelijk is het immers juist deze methodiek die de loonwaarde op de 'goede' manier meet.

Wel maakt het experiment (deels) inzichtelijk waarom eventuele verschillen tussen methodieken ontstaan. We onderzoeken op welke punten de metingen van elkaar afwijken. Dit biedt aanknopingspunten om tot een meer uniforme (en daarmee meer valide en betrouwbare) loonwaardebepaling te komen.

2 Methoden aangepast

Het voorgaande experiment loonwaarde bood vier concrete aanknopingspunten om de validiteit en betrouwbaarheid van de methodieken te verbeteren. Aanknopingspunten konden worden gevonden in:

- Inhoud methoden: een uniforme keuze ten aanzien van de *bestanddelen* waaruit de loonwaarde bestaat
- Inhoud methoden: een uniforme keuze ten aanzien van de *norm* waartegen de arbeidsprestatie van de werknemer moet worden afgezet
- Uitvoering methoden: het waarborgen van de *objectiviteit* van de metingen
- Uitvoering methoden: het waarborgen van de *betrouwbaarheid* van de metingen

De inhoud van de methoden bleek uiteen te lopen, het ministerie vond uniformering op dit vlak noodzakelijk en heeft daarom besloten de (inhoudelijke) aanknopingspunten over te nemen. Dat betekende dat de methodieken zodanig moesten worden aangepast dat zij dezelfde bestanddelen meten en dat zij binnen deze metingen dezelfde norm hanteren. Hiervoor zijn nadere richtlijnen ontwikkeld.

2.1 Bestanddelen

Prestatieniveau en additionele kosten

Het voorgaande experiment laat zien dat de verschillende bestanddelen die methodieken gebruikten, kunnen worden geclusterd naar twee hoofdelementen:

- Prestatieniveau
- Additionele kosten

Alle methodieken meten het *prestatieniveau* van de werknemer. Dat wil zeggen: zij brengen allemaal in kaart welke prestatie de werknemer op de werkvloer levert, en zetten deze prestatie vervolgens af tegen die van een reguliere collega in dezelfde functie.

Vijf van de zeven methodieken namen daarnaast ook *additionele kosten* in de meting mee. Additionele kosten hebben niet direct betrekking op de prestaties van de werknemer 'an sich', maar op de kosten die de werkgever moet maken om de werknemer in dienst te hebben en houden. Het gaat hier bijvoorbeeld om zaken als begeleiding, ziekteverzuim en werkplekaanpassingen. Niet alle methoden lieten additionele kosten meewegen in de berekening van de loonwaarde en de methodieken die dit wel deden, namen verschillende (combinaties van) kostensoorten mee. Dit leidde tot verschillen in de vastgestelde loonwaardes: reden om hierin een uniforme keuze te maken.

Het ministerie heeft in overleg met verschillende pilotgemeenten besloten dat **alleen het prestatieniveau** in de loonwaardebepaling mag worden meegewogen. Additionele kosten mogen tijdens de loonwaardebepaling in kaart worden gebracht, maar mogen niet worden meegenomen in de berekening van de loonwaarde. Gemeenten gaven aan dat zij eventuele additionele kosten liever op een andere manier aan werkgevers vergoeden: het is niet 'fair' deze in de loonwaarde – en daarmee uiteindelijk in het loon van de werknemer – te verrekenen. Voor het vergoeden van deze kosten zijn andere instrumenten beschikbaar. Boven-

dien biedt juist de (mate van) compensatie van additionele kosten de ruimte om met de werkgever te onderhandelen, iets dat niet wenselijk is wanneer het gaat om de hoogte van de loondispensatie. Door de additionele kosten buiten de loonwaarde te houden, behouden gemeenten de ruimte om toch met werkgevers te onderhandelen, zonder dat de werknemer hiervan de dupe wordt. De werknemer kan zo rekenen op een eerlijk loon, dat uitsluitend gebaseerd is op zijn geleverde prestaties op de werkvloer.

Prestatieniveau

Deze beslissing betekende dat de methoden zodanig moesten worden aangepast, dat zij niet langer additionele kosten in de berekening van de loonwaarde meenamen. Alleen het prestatieniveau bepaalt de loonwaarde. Dit prestatieniveau bestaat uit de volgende aspecten:

- 1) Hoe *snel* voert de werknemer zijn werkzaamheden uit?
Een eerste belangrijk aspect van het prestatieniveau is de snelheid waarmee de taken worden uitgevoerd. De methoden brengen daarom vrijwel allemaal (impliciet of expliciet) in kaart hoe het werktempo van de werknemer zich verhoudt tot dat van een reguliere collega in dezelfde functie.
- 2) Hoe *goed* voert de werknemer zijn werkzaamheden uit?
Daarnaast is het voor het prestatieniveau belangrijk dat de werkzaamheden niet alleen op een redelijk tempo worden uitgevoerd, maar dat dit ook goed gebeurt. Methoden kijken daarom ook naar de kwaliteit van het geleverde werk en de nauwkeurigheid waarmee de taken worden uitgevoerd.
- 3) Hoe *inzetbaar* is de werknemer?
De inzetbaarheid van werknemers met een arbeidsbeperking wijkt soms af van die van hun reguliere collega's. Zij zijn daardoor minder flexibel in de organisatie in te zetten, bijvoorbeeld omdat zij alleen kunnen werken met regelmaat en/of een afgebakend (soms vereenvoudigd) takenpakket. Dit houdt in dat zij soms dezelfde output en kwaliteit leveren, maar alleen op bepaalde afgebakende taken. Dit heeft invloed op het prestatieniveau. Methoden kijken daarom ook naar de inzetbaarheid van de werknemer om het prestatieniveau te bepalen. De inzetbaarheid heeft zowel betrekking op de inhoud van het werk (functieaanpassingen, zoals bijvoorbeeld taakvereenvoudiging of aanpassing van het takenpakket) als op de randvoorwaarden waaronder het werk wordt uitgevoerd (werkaanpassingen, zoals aanpassing van de werktijden of een aangepast pauzeregime).

Richtlijnen samengevat:

Loonwaarde meet de prestatie die de werknemer levert op de werkvloer en de manier waarop deze prestatie zich verhoudt tot die van een reguliere werknemer in dezelfde functie. Dit prestatieniveau wordt bepaald door de volgende aspecten:

- Hoe *snel* worden de werkzaamheden uitgevoerd?
 - ✓ Tempo
- Hoe *goed* worden de werkzaamheden uitgevoerd?
 - ✓ Kwaliteit
- Hoe *inzetbaar* is de werknemer?
 - ✓ Functieaanpassingen (inhoud)
 - ✓ Werkaanpassingen (randvoorwaarden)

2.2 Norm

Ook ten aanzien van de te hanteren norm waartegen dit prestatieniveau van de werknemer moest worden afgezet, moest een uniforme keuze worden gemaakt. Dit gold met name voor werknemers in (sterk) aangepaste of gecreëerde functies. We zagen dat methoden hierin verschillende keuzes maakten:

- De arbeidsprestatie werd afgezet tegen de eisen die aan een reguliere werknemer (zouden) worden gesteld op het eigen, aangepaste, functieniveau. De norm werd dan bepaald door het *aangepaste functieniveau*.
- De arbeidsprestatie werd afgezet tegen de eisen die aan een reguliere werknemer worden gesteld op het reguliere functieniveau. De norm wordt dan bepaald door het *reguliere functieniveau*.

Door het ministerie is hierop besloten dat **voor de norm uitgegaan moet worden van het reguliere functieniveau**. Uit een voorbeeldcasus die door alle methoden is uitgewerkt, bleek echter dat hierbij nog nadere duiding nodig was.

Voorbeeldcasus

Hugo werkt bij een hekkenmaker en heeft hier een ondersteunende functie op assistent-niveau. Hoewel de functie van assistent binnen de branche vaker voorkomt en ook in de CAO is beschreven, kent het bedrijf zelf deze functie niet: alle reguliere medewerkers werken minimaal op het niveau van lasser/monteur. De functie van assistent komt onder reguliere werknemers binnen dit bedrijf niet voor.

De vraag is in dit geval daarom tegen welke reguliere normfunctie de arbeidsprestatie van Hugo moet worden afgezet. Een aantal methoden kiest ervoor hierbij het *bedrijf* als uitgangspunt te nemen en uit te gaan van binnen het bedrijf gebruikelijke (hogere) reguliere functieniveau. De arbeidsprestatie van Hugo wordt daarmee afgezet tegen de eisen die aan de reguliere normfunctie van lasser/monteur worden gesteld. De loonwaarde komt daarmee relatief laag uit.

Andere methoden gaan voor het bepalen van de normfunctie uit van de werkzaamheden van de *werknemer*, en zoeken daarbij de meest dichtbij gelegen 'erkende' reguliere functie. Omdat in de CAO de functie van assistent wordt beschreven, wordt de arbeidsprestatie van Hugo afgezet tegen de eisen die binnen dit (lagere) reguliere functieniveau worden gesteld. De loonwaarde komt daarmee relatief hoog uit.

Bovenstaand voorbeeld laat zien dat het vaststellen van de juiste normfunctie geen sinecure is. In beide gevallen gaat het om een reguliere functie, en hebben de methoden zich dus gehouden aan de richtlijnen van het ministerie. Niettemin kiezen zij voor verschillende normfuncties van verschillend niveau, waardoor alsnog verschillen in de vastgestelde loonwaarde optreden. Een verdere invulling van deze richtlijn bleek dus noodzakelijk.

De reguliere normfunctie is daarom nader gespecificeerd als: de reguliere functie die **het dichtst tegen de werkzaamheden van de werknemer aan** ligt. Uitgangspunt bij het vaststellen van de normfunctie is dus de werknemer en niet het bedrijf. Dat betekent concreet dat de normfunctie niet noodzakelijkerwijs regulier hoeft voor te komen binnen het bedrijf, wanneer binnen de branche of sector een meer op de functie lijkende 'erkende' of 'gangbare' (bijvoorbeeld omdat deze is beschreven in de CAO of het geldende functiegebouw) functie bestaat. Indien dit het geval is, moet deze – op de functie van de werknemer lijkende – functie als uitgangspunt en norm voor de meting worden gehanteerd.

Richtlijnen samengevat:

Voor het vaststellen van de norm waartegen de arbeidsprestatie van de werknemer moet worden afgezet, moet worden uitgegaan van het *reguliere functieniveau*. Het gaat hier om de binnen de branche of sector 'gangbare' functie die *het dichtst* tegen de werkzaamheden van de werknemer aan ligt.

2.3 Aanpassingen methoden

Deze richtlijnen betekenen concreet dat de methoden – indien nodig – als volgt moesten worden aangepast:

- Op het niveau van de methode: aanpassen van bestanddelen
 - Verwijderen van additionele kosten uit de berekening van de loonwaarde
 - Meten en berekenen van het prestatieniveau op basis van de bestanddelen werktempo, kwaliteit en inzetbaarheid.
- Op het niveau van de uitvoering: het vaststellen van de juiste norm. Hiervoor moeten voor elke meting twee stappen worden doorlopen:
 - Wat zijn de werkzaamheden die de werknemer uitvoert?
 - Welke – binnen de branche of sector 'gangbare' – reguliere functie lijkt hier het meest op?

3 Uitkomsten experiment

3.1 Overzicht uitkomsten

Wat is nu het resultaat van de geformuleerde richtlijnen? Leidt de implementatie van deze richtlijnen er inderdaad toe dat de vastgestelde loonwaardes dichterbij elkaar liggen? In onderstaande tabel zijn de uitkomsten van de metingen per cliënt en methodiek weergegeven.

Tabel 3.1 Overzicht percentage loonwaarde per cliënt naar methode

	Dariuz	VTA methodiek	Activa	Arbolab	Loonbalans	Match-care	Methode UWV	Vershil maximaal
Cliënt 1	60%		48%	59%		48%		12%
Cliënt 2	58%				62%			4%
Cliënt 3	61%				58%	59%	100%	42%
Cliënt 4		93%	92%		71%			22%
Cliënt 5	45%	66%	57%			53%	41%	25%
Cliënt 6		97%		62%	85%		74%	35%
Cliënt 7		63%	38%	58%			30%	33%
Cliënt 8			66%		100%	73%		34%
Cliënt 9	64%	82%		35%		75%		47%
<i>Aant. waarnemingen</i>	5	5	5	4	5	5	4	

3.2 Resultaten beschouwd

Het overzicht laat nog altijd een aantal grote verschillen zien. Ontwikkelaars geven zelf aan dat het vaststellen van loonwaarde geen exacte wetenschap is: de loonwaarde is niet tot op de procent nauwkeurig te bepalen. Dit betekent dat verschillen in de vastgestelde loonwaardes kunnen voorkomen en zelfs verwacht mogen worden. Ontwikkelaars noemen een

marge van ongeveer 10%¹. Dit betekent dat de verschillen in vastgestelde loonwaardes voor één cliënt (dat wil zeggen: het verschil tussen de hoogste en de laagste vastgestelde loonwaarde) binnen deze marge van 10% zouden moeten vallen.

Uit het overzicht blijkt echter dat deze verschillen in een aantal gevallen (aanzienlijk) groter zijn. Van slechts één cliënt vallen alle vastgestelde loonwaardes binnen de marge van 10%. Dit is zelfs een slechter resultaat dan tijdens het voorgaande experiment, toen dit voor twee cliënten het geval was.²

Tegelijkertijd lijken de verschillen tijdens deze ronde van het experiment vaker toe te schrijven aan 'uitschieters'. Hoewel van slechts één cliënt alle waarnemingen binnen de marge van 10% vallen, is dit voor drie cliënten voor alle waarnemingen behalve één het geval. Van nog eens één cliënt is het verschil tussen de hoogste en de laagste waarneming bovendien slechts 12% en vallen alle waarnemingen dus bijna binnen de toegestane marge. Al met al zijn de verschillen – op uitschieters na – in dit tweede experiment minder groot dan bij het eerste experiment.

Verder zijn uitschieters in deze ronde vaker toe te schrijven aan enkele methoden. Waar in het voorgaande experiment vrijwel alle methoden uitschieters vertoonden (dat wil zeggen: het verschil met de meest dichtbij gelegen waarneming is (veel) groter dan 10%), geldt dat binnen de huidige ronde nog maar voor vier methoden van de zeven. Dat wijst erop dat de overige methoden dichter naar elkaar zijn toegegroeid en daarmee vergelijkbaarder zijn geworden.

Om overeenkomsten en verschillen tussen methoden nader te kunnen duiden, gaan we in de volgende paragraaf dieper in op de vergelijking tussen methoden onderling.

3.3 Vergelijking tussen methoden

Vrijwel elke combinatie van twee methoden komt minimaal twee keer voor.³ Cliënt 2 en cliënt 3 zijn bijvoorbeeld door zowel Dariuz als Loonbalans gemeten. Door voor alle combinaties de loonwaardes met elkaar te vergelijken, krijgen we inzicht in de mate waarin bepaalde methodieken vaker hogere of lagere loonwaardes hebben gemeten, of juist gelijk zijn aan elkaar. We hanteren hierbij de door ontwikkelaars genoemde marge van 10%: als de vastgestelde loonwaardes binnen deze marge van 10% vallen, beschouwen we de uitkomsten als gelijk. Omdat meerdere combinaties van twee waarnemingen deze marge van 10% heel dicht naderen (dat wil zeggen: het verschil tussen twee waarnemingen is slechts 11% of 12%), hebben we ook gekeken naar de mate waarin de loonwaardes bijna binnen deze marge vallen. We hebben er bewust voor gekozen geen statistische analyses uit te voeren. Het aantal waarnemingen is hiervoor te beperkt.

In onderstaande tabellen zijn de resultaten van de vergelijking zichtbaar. In tabel 3.2 zijn de gemaakte vergelijkingen te zien.

¹ Kemper, R., L. de Ruig en S. de Visser (2009), *Aan de slag met loonwaardemeting. Inventarisatie van methoden om loonwaarde te meten*.

² Voor een overzicht van de resultaten van het eerste experiment: zie bijlage 1.

³ Uitzondering is de combinatie Arbolab-Loonbalans.

Tabel 3.2 Vergelijking tussen methoden

	Dariuz	VTA	Activa	Arbolab	Loonbalans	Matchcare	UWV
Dariuz		--	≈ ≈	=+	=	≈ ≈ ≈ ≈	-
VTA	++ ¹		==+	++	+ ≈	+=	+++
Activa	≈ ≈	==		≈-	+ -	==	+=
Arbolab	=-	- -	≈+		-	≈-	≈+
Loonbalans	=	- ≈	-+	+		=+	- ≈
Matchcare	≈ ≈ ≈ ≈	-	==	≈+	=-		- ≈
UWV	+=	- - -	-	≈-	+ ≈	+ ≈	

In de cellen van de tabel is steeds het resultaat van de vergelijking van twee methoden weergegeven. Elke rij geeft voor de betreffende methode aan of de vastgestelde loonwaardes in vergelijking met de andere methodieken hoger (+), lager (-), of niet verschillend (=) zijn. We houden hierbij steeds een marge aan van 10%: als de verschillen tussen twee loonwaardes niet groter zijn dan 10%, beschouwen we de loonwaardes als niet verschillend. In meerdere gevallen valt het verschil tussen twee methoden nagenoeg binnen de marge van 10%. Dat is in de tabel aangegeven met een ≈. Het verschil is in dit geval niet groter dan 12%. In de rij 'Activa' is bijvoorbeeld te zien dat de loonwaarde die door Activa is vastgesteld één keer lager uitkomt dan de door Arbolab vastgestelde loonwaarde, en dat de loonwaardes voor een andere cliënt voor beide methodieken nagenoeg binnen een marge van 10% vallen. In vergelijking met UWV komt Activa één keer hoger uit en zijn de loonwaardes één keer aan elkaar gelijk. Voor elke combinatie van twee methodieken is op deze manier te zien hoe de vastgestelde loonwaardes zich tot elkaar verhouden.

Tabel 3.3 bevat een overzicht van de totalen.

Tabel 3.3 Resultaten vergelijking

	Dariuz	VTA	Activa	Arbolab	Loonbalans	Matchcare	UWV	Totaal
+	1	10	2	2	3	1	3	22
-	3	0	3	5	3	3	5	22
=/≈	10	5	9	5	5	11	5	50
Totaal	14	15	14	12	11	15	13	94

Per methodiek is aangegeven hoe vaak de loonwaarde hoger (+), lager (-), of (nagenoeg) niet verschillend (=/≈) is van de met de andere methodieken vastgestelde loonwaardes. Dariuz is bijvoorbeeld veertien keer vergeleken met een andere methode, waarvan Dariuz één keer een hogere loonwaarde had, drie keer een lagere loonwaarde en tien keer een loonwaarde had die (nagenoeg) gelijk was aan de loonwaarde die met een andere methode is vastgesteld.

¹ ++ betekent dat de vergelijking tussen deze methoden twee keer is gemaakt en dat in beide gevallen de methode in de rij tot een hogere uitkomst leidde dan de methode in de kolom.

Interpretatie vergelijkingen

In totaal hebben we 94 vergelijkingen kunnen maken. Uit deze vergelijkingen blijkt dat:

- De vastgestelde loonwaardes in ruim de helft van de gevallen (50 keer, of 53%) (nagenoeg) aan elkaar gelijk zijn. Dit is een verbetering ten opzichte van de voorgaande ronde van het experiment, toen dit voor slechts een derde van de vergelijkingen gold;
- Dariuz, Actíva en Matchcare merendeels loonwaardes geven die (nagenoeg) gelijk zijn aan de met de andere methoden vastgestelde loonwaardes. Deze methoden geven ten opzichte van elkaar bovendien voor *elke* vergelijking (nagenoeg) dezelfde loonwaarde. Ook dit is een verbetering ten opzichte van het voorgaande experiment, toen we concluderden dat de uitkomsten van *alle* methodieken (sterk) van elkaar verschilden;
- VTA een overwegend hogere loonwaarde geeft dan de andere methoden, Arbolab vaker een hogere loonwaarde geeft dan de andere methoden, terwijl de richting van het verschil voor de overige methodieken (Loonbalans en UWV) verschilt.¹

3.4 Conclusie

Er kan dus worden geconcludeerd dat het aanpassen van de methoden **resultaat lijkt te hebben gehad**. De verschillen tussen de methoden zijn – een aantal uitschieters daargelaten – minder groot en de vastgestelde loonwaardes zijn vaker dan in het voorgaande experiment (nagenoeg) aan elkaar gelijk. Een aantal methoden geeft nu bovendien vergelijkbare uitkomsten. Tegelijkertijd moeten we concluderen dat dit **nog niet genoeg** is: de uitschieters veroorzaken nog altijd grote verschillen en over de hele linie zijn de uitkomsten van *alle* methoden nog niet altijd vergelijkbaar. De uitkomsten van het experiment laten, kortom, nog steeds veel verschillen zien.

Dat betekent dat we op grond van het experiment nog steeds *niet* kunnen concluderen dat de methoden valide en betrouwbaar zijn. Hoewel de uitkomsten overwegend dichterbij elkaar liggen dan tijdens het voorgaande experiment, meten nog steeds niet alle methoden hetzelfde, of meten zij niet allemaal op dezelfde manier.

In het volgende hoofdstuk gaan we op zoek naar verklaringen voor deze resterende verschillen.

¹ Aan deze gegevens kunnen – zoals al in de inleiding aangegeven – *geen* conclusies worden verbonden over wat de 'betere' en wat de 'minder goede' methoden zijn. De cijfers geven slechts weer in hoeverre de methoden vergelijkbaar zijn, maar zeggen niets over de kwaliteit van de door de methoden vastgestelde loonwaardes.

4 Verschillen verklaard

Om verklaringen te vinden voor de nog resterende verschillen, zijn de uitkomsten van de loonwaardemetingen op cliëntniveau geanalyseerd. Bij elke loonwaardemeting is een onderzoeker van Panteia/Research voor Beleid aanwezig geweest. De observatieverslagen bieden aanknopingspunten om de verschillen te verklaren. Daarnaast zijn per cliënt de rapportages van de loonwaardemetingen bestudeerd.

Op basis van deze analyses hebben we een aantal verklaringen voor de verschillen in loonwaardes kunnen vaststellen. Twee hiervan hebben betrekking op de **inhoud van de methoden**, overige verklaringen zijn gerelateerd aan de **manier waarop de methoden worden uitgevoerd**.

Op het niveau van de methoden zien we verschillen in:

- Berekeningswijze en weging
- Inhoud methoden

Op het niveau van de uitvoering zien we verschillen in:

- Vaststellen van de norm
- Vaststellen van de weging
- Rol van de uitvoerder

We lichten deze verklaringen in onderstaande paragrafen nader toe.

4.1 Verklaring 1. Niveau methode: berekeningswijze en weging

De methoden meten in deze tweede ronde van het experiment dezelfde bestanddelen om de loonwaarde van een cliënt te bepalen: zij brengen allemaal het tempo, de kwaliteit en de inzetbaarheid in kaart en bepalen op basis daarvan de loonwaarde. Dit maakt het in vergelijking met het voorgaande experiment makkelijker de methoden met elkaar te vergelijken. Deze grotere vergelijkbaarheid legt nieuwe verschillen bloot. De belangrijkste daarvan heeft betrekking op de **berekeningswijze en weging** van de loonwaarde: op welke manier bepaalt de combinatie van de afzonderlijke bestanddelen tempo, kwaliteit en inzetbaarheid de loonwaarde van een cliënt? Methoden maken hierin verschillende keuzes.

Verschillende berekeningswijzen

Een eerste verschil betreft de berekeningswijze van de loonwaarde: hoe volgt de loonwaarde uit de bestanddelen tempo, kwaliteit en inzetbaarheid? We zien vier verschillende berekeningswijzen:

1. De loonwaarde is het **(gewogen) gemiddelde** van tempo, kwaliteit en inzetbaarheid. Ofwel: $\text{loonwaarde} = (\text{weging} \times \text{tempo}) + (\text{weging} \times \text{kwaliteit}) + (\text{weging} \times \text{inzetbaarheid})$.
2. De loonwaarde is het **(gewogen) product** van tempo, kwaliteit en inzetbaarheid. Ofwel: $\text{loonwaarde} = (\text{weging} \times \text{tempo}) \times (\text{weging} \times \text{kwaliteit}) \times (\text{weging} \times \text{inzetbaarheid})$.
3. De loonwaarde is het **verschil** van tempo, kwaliteitsverlies en inzetbaarheidbaarheidsverlies.
Ofwel: $\text{loonwaarde} = \text{tempo} - \text{kwaliteitsverlies} - \text{inzetbaarheidsverlies}$.
4. De loonwaarde is een **combinatie** van bovenstaande berekeningswijzen.
Bijvoorbeeld: $\text{loonwaarde} = (\text{tempo} - \text{kwaliteitsverlies}) \times \text{inzetbaarheid}$.

We zien voor meerdere cliënten dat de deelscores op de bestanddelen tempo, kwaliteit en inzetbaarheid wel binnen elkaars marge liggen (en daarmee dus nagenoeg aan elkaar gelijk zijn), maar dat hier door een andere berekeningswijze toch een geheel andere loonwaarde uit volgt. Een fictief rekenvoorbeeld maakt dit inzichtelijk.

Rekenvoorbeeld

De loonwaarde van een cliënt wordt vastgesteld met methode A, B, C en D. Alle methoden komen uit op de volgende deelscores:

Tempo: 40%
Kwaliteit: 80%
Inzetbaarheid: 90%

- Methode A gaat voor de berekening van de loonwaarde uit van het *gemiddelde* van deze bestanddelen. De loonwaarde komt daarmee uit op: $(40\% + 80\% + 90\%) / 3 = 70\%$
- Methode B gaat voor de berekening van de loonwaarde uit van het *product* van deze bestanddelen. De loonwaarde komt daarmee uit op: $0,4 \times 0,8 \times 0,9 = 29\%$
- Methode C gaat voor de berekening van de loonwaarde uit van het *verschil* van deze bestanddelen. De loonwaarde komt daarmee uit op: $40\% - 20\% - 10\% = 10\%$
- Methode D gaat voor de berekening van de loonwaarde uit van een *combinatie* van bovenstaande berekeningswijzen. De loonwaarde komt daarmee uit op: $(40\% - 20\%) \times 90\% = 18\%$

Opvallend is dat de verschillende berekeningswijzen er soms toe leiden dat een cliënt met de ene methode hogere deelscores behaalt dan met de andere methode, maar dat de eerste methode (in het voorbeeld methode B) door een verschil in berekeningswijze uiteindelijk toch op een (veel) lagere loonwaarde uitkomt. We hebben dit in het experiment verschillende keren gezien. De verschillende berekeningswijzen zijn zo zowel belangrijke veroorzakers van verschillen in loonwaardes, maar hebben hierdoor soms juist ook een dempend effect op dit verschil.

Verskillende weging

Daarnaast maken methoden verschillende keuzes als het gaat om de **weging** van de verschillende aspecten. Feitelijk zou voor de berekening van de loonwaarde een tweeledige weging moeten worden toegepast: er moet – indien voor de functie opportuun – worden gewogen op **taken en op bestanddelen**. Immers: niet elke taak is binnen een functie even belangrijk. Wanneer een cliënt heel goed scoort op taken die voor de functie erg belangrijk zijn, moet dat in de loonwaarde zwaarder doorwegen dan een slechte score op taken die binnen de functie veel minder belangrijk zijn. Met andere woorden: aan de belangrijkste taken moet het grootste gewicht worden toegekend. De belangrijkste taken bepalen immers voor het belangrijkste deel de productiviteit van een werknemer.

Hetzelfde geldt voor de bestanddelen tempo, kwaliteit en inzetbaarheid. Een voorbeeld: als een werknemer heel nauwkeurig is, maar totaal niet onder druk kan presteren, zal hij in een functie als archiefmedewerker een hogere loonwaarde genereren dan in een functie als afwashulp in de horeca. Kwaliteit is in de functie van archiefmedewerker immers erg belangrijk, terwijl in de functie van afwashulp tempo juist weer heel belangrijk is. Voor een functie als archiefmedewerker zal kwaliteit in de loonwaarde daarom zwaarder moeten meewegen dan voor een functie als afwashulp, waar aan tempo juist een groter gewicht moet worden gegeven.

In de praktijk zien we dat deze weging op taken en bestanddelen **niet door alle methoden op dezelfde manier wordt toegepast**. Sommige methoden passen in het geheel geen weging toe, en laten alle taken en bestanddelen in de loonwaarde even zwaar meewegen. Andere methoden voorzien wel in een weging van de taken, maar hanteren geen weging voor de bestanddelen tempo, kwaliteit en inzetbaarheid. Eén methode hanteert weliswaar een weging voor deze bestanddelen, maar deze weging staat los van de functie en is voor elke meting hetzelfde (er geldt altijd een verhouding van 50:25:25 voor de bestanddelen tempo, kwaliteit en inzetbaarheid, ook wanneer tempo binnen een functie eigenlijk niet zo belangrijk is). Tot slot zien we methoden die zowel wegen op taken als op bestanddelen. De verschillende keuzes die methoden hierin maken, leiden eveneens tot verschillen in loonwaarde.

Rekenvoorbeeld

Een werknemer werkt op een project in een stadsarchief en heeft hier als taak om alle persoonsnamen uit een collectie oude handschriften te filteren en deze in de computer in te voeren. Op deze manier bouwt het archief een database van belangrijke historische figuren uit de stad op. Binnen deze functie is het vooral erg belangrijk dat het werk goed en nauwgezet wordt uitgevoerd: het is erg belangrijk dat de medewerker geen namen in de tekst over het hoofd ziet, omdat dit de database incompleet zou maken. Tempo is voor deze taak veel minder van belang: de subsidie voor dit project loopt nog een aantal jaar door. Volgens de werkgever wordt de productiviteit van deze functie voor ongeveer 20% bepaald door het tempo. Kwaliteit bepaalt de resterende 80% van de productiviteit.

Methode A, B en C hebben elk de loonwaarde van deze cliënt bepaald. Zij constateren allen dat het tempo van de cliënt op ongeveer 10% van het reguliere niveau ligt, maar dat de cliënt wel erg goede kwaliteit levert en binnen de functie volledig inzetbaar is.

Methode A past *geen weging* van de bestanddelen toe. Het tempoverlies weegt binnen deze methode even zwaar als de goede score op kwaliteit en bestanddelen. De loonwaarde komt daarmee uit op $(10\% + 100\% + 100\%) / 3 = 70\%$

Methode B hanteert een *vaste weging* voor tempo, kwaliteit en inzetbaarheid, waarbij tempo dubbel meeweegt. Het tempoverlies weegt daarmee binnen deze methode twee keer zo zwaar mee als de goede score op kwaliteit en inzetbaarheid. De loonwaarde komt daarmee uit op $(10\% \times 0,5) + (100\% \times 0,25) + (100\% \times 0,25) = 55\%$.

Methode C past de weging aan de vereisten van de functie aan. Het tempoverlies bepaalt daarmee slechts 20% van de loonwaarde. Dit resulteert in een loonwaarde van $(10\% \times 0,2) + (100\% \times 0,8) = 82\%$.

Voor bovenstaand voorbeeld zijn we ervan uitgegaan dat de methoden weliswaar een verschillende weging toepassen, maar dat zij wel dezelfde berekeningswijze hanteren (in dit geval het (al dan niet gewogen) gemiddelde van de bestanddelen). Bovendien zijn we voor dit voorbeeld uitgegaan van een functie die bestaat uit één taak, zodat niet ook verschillen in de weging van taken ontstaan. In de praktijk lopen **verschillende wegings- en berekeningswijzen** echter door elkaar heen. Het is op basis van bovenstaande rekenvoorbeelden niet moeilijk voor te stellen hoe dit tot grote verschillen in loonwaarde kan leiden, ook wanneer de deelscores op de bestanddelen tempo, kwaliteit en inzetbaarheid nagenoeg aan elkaar gelijk zijn.

4.2 Verklaring 2. Niveau methode: inhoud methoden

Met de richtlijnen zoals die door het ministerie zijn neergelegd, zijn de meeste methoden inhoudelijk meer op elkaar gaan lijken. Zij meten dezelfde bestanddelen en doen dit in grote lijnen op dezelfde manier. Eén methode wijkt voor wat betreft de inhoud af van de andere methoden. Verschillen doen zich voor op het gebied van:

- Visie en uitgangspunten: wat voor *'type' informatie* wordt er gemeten?
- Meetinstrumenten: *hoe* wordt dit gemeten?
- Bestanddelen: wat wordt daarmee *concreet in beeld gebracht*?

Waar het bij het *type informatie* voor het merendeel van de methoden gaat om het gedrag dat de werknemer op de werkvloer laat zien (output), kijkt één methode vooral naar de aanleg en vermogens van de werknemer (input). Bij deze verschillen in visie hoort ook een ander type *meetinstrument*. Het merendeel van de methoden maakt gebruik van meer 'zachte' methoden van dataverzameling: werkgever en werknemer worden – in een gesprek en/of op basis van een vragenlijst – gevraagd de arbeidsprestatie van de werknemer te beoordelen. Hoewel op verschillende manieren wordt getracht deze informatie te objectiveren, blijft deze informatie per definitie altijd subjectieve elementen bevatten. De 'andere' methode ondervangt dit door gebruik te maken van een serie psychologische tests, waarbij de kenmerken van de werknemer letterlijk worden gemeten. Deze verschillen in type informatie en meetinstrumenten kunnen leiden tot verschillen in de vastgestelde loonwaardes. Dat blijkt uit onderstaande voorbeelden uit de cases.

Voorbeeld

Een cliënt werkt al meer dan tien jaar in dezelfde functie. Zijn loonwaarde wordt vastgesteld met vier verschillende methoden. Alle methoden constateren dat de cliënt een aantal beperkingen heeft die tot productieverlies leiden: de cliënt is verstandelijk beperkt en is daardoor weinig flexibel, het kost hem veel tijd om nieuwe taken aan te leren en hij is niet goed in staat zelf problemen op te lossen.

Voor methode A wegen deze zaken echter veel zwaarder mee dan voor methoden B, C en D. Dat komt omdat methode A vooral kijkt naar de input die de cliënt levert: wat zijn zijn capaciteiten, persoonlijke eigenschappen en competenties en in hoeverre past dit bij de eisen die aan de functie worden gesteld? Uit de loonwaardebepaling blijkt dat de functie-eisen een relatief groot beroep doen op juist die capaciteiten die de cliënt in aanleg niet bezit: flexibiliteit, probleemoplossend vermogen en instrueerbaarheid. De loonwaarde komt daarmee laag uit.

Methoden B, C en D kijken vooral naar output- en gedragkant (wat gebeurt er op de werkvloer?) en constateren dat de cliënt op dit gebied weliswaar behoorlijk wat beperkingen heeft, maar dat cliënt en werkgever in de tien jaar dat de cliënt binnen deze functie werkzaam is een goede modus hebben gevonden om met deze beperkingen om te gaan. De cliënt heeft in deze tien jaar veel routine opgedaan en kent alle voorkomende werkzaamheden door en door, waardoor er in de praktijk nog slechts zelden een beroep op zijn flexibiliteit, probleemoplossend vermogen en instrueerbaarheid wordt gedaan. Deze methoden constateren daarom dat het productieverlies – juist doordat de cliënt al jaren binnen deze functie werkt – binnen de huidige functie beperkt is. Zij komen daarmee op een aanzienlijk hogere loonwaarde uit dan methode A.

Voorbeeld

De loonwaarde van een cliënt wordt bepaald met verschillende methoden. Eén methode maakt hiervoor gebruik van een serie psychologische tests waarmee de capaciteiten, persoonlijke eigenschappen en competenties van de cliënt worden gemeten. De methode stelt daarmee vast dat de cliënt problemen heeft op het gebied van aanpassing ('een zeer geringe flexibiliteit') en sociale vaardigheden ('niet gespist op de signalen van anderen', 'kan zich moeizaam in anderen verplaatsen').

De andere methoden maken geen gebruik van psychologische tests, maar vragen in een gesprek met werkgever en cliënt naar de manier waarop de cliënt op de werkvloer functioneert. Opvallend is dat het door methode A geconstateerde gebrek aan flexibiliteit en sociale vaardigheden in deze gesprekken niet aan de orde komt. In plaats daarvan zit het productieverlies volgens de werkgever vooral in een gebrek aan zelfstandigheid (omdat de cliënt erg onzeker is), iets waarvan uit de psychologische tests van methode A juist bleek dat de cliënt hier 100% op scoort. Het gebruik van verschillende meetinstrumenten leidt zo tot de identificatie van verschillende 'probleemgebieden' die productieverlies veroorzaken en daarmee tot verschillende loonwaardes.

Psychologische tests zijn gevalideerd om bepaalde variabelen (bestanddelen) te meten. Dit maakt het minder makkelijk om methoden die hiervan gebruik maken aan de nieuwe richtlijnen aan te passen. De resultaten van deze tests zijn niet zondermeer te vertalen naar de bestanddelen tempo, kwaliteit en inzetbaarheid, omdat dit ten kosten kan gaan van de interne validiteit. Voor methoden die gebruik maken van 'zachte', niet gevalideerde meetinstrumenten is dit aanzienlijk gemakkelijker: zij stappen gemakkelijker over dit vraagstuk van interne validiteit heen en passen hun vragenlijsten eenvoudigweg aan door enkele vragen te schrappen of toe te voegen.¹

Voorbeeld

Van een cliënt wordt de loonwaarde vastgesteld met verschillende methoden. De loonwaarde van methode A komt vrij hoog uit: de andere methoden stellen een groter productieverlies vast. Nadere analyse laat zien dat het productieverlies dat de andere methoden vaststellen, voor een groot deel is toe te schrijven aan de inzetbaarheid van de cliënt, en dan met name aan de *functieaanpassingen* die voor hem zijn gedaan. De cliënt voert in de praktijk lang niet alle taken die bij de functie horen uit en komt daardoor uit op een lage loonwaarde. Dit aspect van inzetbaarheid wordt met de psychologische tests van methode A echter niet gemeten. Het is voor deze methode daarom lastig dit bestanddeel in de loonwaarde mee te wegen, waardoor de loonwaarde vrij hoog uitkomt.

Al met al zien we dat het **verschil** ten opzichte van de andere methoden door deze facetten ten opzichte van het eerste experiment oploopt. De andere methoden zijn door de doorgevoerde aanpassingen sterker op elkaar gaan lijken. Het verschil met de methode die minder sterk is aangepast, loopt hiermee op.

¹ Dit maakt het vraagstuk van de interne validiteit juist ook voor deze methoden echter niet minder relevant en urgent. Juist omdat zij geen gebruik maken van gevalideerde meetinstrumenten, is het voor hen van belang de interne validiteit van hun methode aan te tonen: meten zij wel echt wat zij zeggen te meten? Dat is door deze methoden niet aangetoond (zie hiervoor ook hoofdstuk 5).

4.3 Verklaring 3. Niveau uitvoering: vaststellen norm

In het voorgaande experiment zagen we dat verschillen in de gehanteerde norm tot grote verschillen kunnen leiden in de loonwaarde. Daarom zijn hiervoor door het ministerie nadere richtlijnen vastgesteld. Het vaststellen van de norm die voor een meting gebruikt moet worden, is echter niet iets dat op het niveau van de methode is in te regelen: het is uiteindelijk aan de uitvoerder om de 'juiste' norm vast te stellen. Dit blijkt geen sinecure: hoewel verschillen in norm in vergelijking met het voorgaande experiment minder vaak voorkomen en daarmee dus ook minder vaak tot verschillen in de vastgestelde loonwaardes leiden, maken uitvoerders hierin soms nog steeds **verschillende keuzes**.

Voorbeeld

Een cliënt werkt bij een autoschadebedrijf. Zijn voornaamste taak is het wassen en schoonmaken van auto's: een taak die ook door zijn collega's en de werkgever wordt uitgevoerd. Volgens de werkgever doet de cliënt in vergelijking met reguliere collega's en met de werkgever zelf ongeveer de helft langer over het uitvoeren van deze taak: "de andere jongens en ik doen er grofweg een kwartier over, hij een half uur." De uitvoerders van methoden A, B en C concluderen hierop dat het tempo van de werknemer hiermee dus op 50% ligt. De uiteindelijke loonwaarde komt daarmee uit op ongeveer 60%. De uitvoerder van methode D concludeert eveneens dat het tempo van de cliënt in vergelijking met reguliere collega's ongeveer de helft minder is. Deze uitvoerder gaat voor de loonwaardebepaling echter uit van een andere norm en vergelijkt de werknemer niet met zijn collega's. Deze collega's zijn allen werkzaam als geschoolde reparateurs en verdienen daarmee ongeveer 2 keer het Wettelijk Minimumloon (WML). De werkzaamheden van de cliënt hebben volgens de uitvoerder echter een waarde die rond WML ligt, waarmee de functie-eisen dus ook twee keer zo laag zouden moeten liggen. Deze uitvoerder constateert dat de functie van autowasser een reguliere functie is, en dat de cliënt dus op basis van deze norm moet worden beoordeeld. Het tempoverlies ten opzichte van reguliere collega's wordt volgens deze uitvoerder dus opgeheven door de minder strenge norm die voor deze cliënt geldt, waarmee de loonwaarde uitkomt op 100%.

In een enkel geval blijkt dat uitvoerders de door het ministerie vastgelegde **richtlijnen niet toepassen**:

Voorbeeld

Een cliënt werkt als productiemedewerker in een 'privaat' SW-bedrijf. Dit bedrijf kent geen reguliere functies: de functie van productiemedewerker is binnen dit bedrijf aangepast, zodat deze geschikt is voor de SW-doelgroep. Niet alle taken die bij de reguliere functie van productiemedewerker horen, worden binnen dit bedrijf dus uitgevoerd. Deze 'versmalde' functie is voor de betreffende cliënt nog verder aangepast. De cliënt voert maar een deel van de nog overblijvende taken van de functie uit. Om de loonwaarde van deze cliënt te kunnen bepalen, moet daarmee in feite een dubbele correctie voor de reguliere norm plaatsvinden. De cliënt voert niet alle taken uit die binnen het bedrijf voorkomen, maar dit geheel aan taken vormt op zichzelf ook nog steeds geen reguliere functie. De norm is dus die van productiemedewerker in een *regulier* bedrijf, en niet die van productiemedewerker binnen het bedrijf waar de cliënt werkt. Eén van de uitvoerders past dit principe niet (goed) toe. Hij berekent een inzetbaarheidverlies in de loonwaarde door omdat de cliënt niet alle taken 'die binnen het bedrijf voorkomen, kan uitvoeren'. Omdat de taken die binnen het bedrijf voorkomen bij elkaar echter ook geen volledige reguliere functie vormen, is het feitelijke productieverlies ten opzichte van een reguliere functie echter nog groter. De uitvoerder laat dit buiten beschouwing: de norm is die van de (aangepaste) functie binnen het bedrijf. De andere methoden hanteren wel de (strengere) reguliere norm en komen op het bestanddeel inzetbaarheid dus aanzienlijk lager uit.

4.4 Verklaring 4. Niveau uitvoering: vaststellen weging

Eerder lieten we al zien dat het voor de vergelijkbaarheid van de vastgestelde loonwaardes van belang is dat de gemeten bestanddelen op dezelfde manier worden gewogen. Deels is dit iets dat – zoals in het gegeven rekenvoorbeeld in paragraaf 4.1 – op het niveau van de methoden is ingeregeld: de weging ligt dan vooraf al vast of de weging wordt bepaald op basis van een door de werkgever in te vullen vragenlijst. Soms is dit echter niet op het niveau van de methode vastgelegd, en is het aan de uitvoerder om het relatieve belang van de bestanddelen tempo, kwaliteit en inzetbaarheid te bepalen. Bovendien is niet alles op het niveau van de methode vast te leggen: soms blijkt het relatieve belang van elk van de bestanddelen pas tijdens een gesprek of de observatie. Het is dan aan de uitvoerder om de weging vast of bij te stellen. Uitvoerders maken hierin soms **verschillende keuzes**. Dat dit kan leiden tot verschillen in de vastgestelde loonwaardes, blijkt uit onderstaand voorbeeld uit één van de cases.

Voorbeeld

Van een cliënt wordt de loonwaarde bepaald met behulp van drie methoden. In geen van de methoden is een weging van de bestanddelen 'ingebouwd'. Het is dus aan de uitvoerders om te bepalen welk gewicht aan de verschillende bestanddelen moet worden toegekend.

In alle gesprekken komt aan de orde dat de cliënt langzamer werkt dan reguliere collega's in dezelfde functie. Tijdens alle metingen blijkt echter ook dat tempo voor deze functie niet van groot belang is: kwaliteit is aanzienlijk belangrijker, en juist hierop scoort de cliënt erg goed. Ook is de cliënt voor 100% inzetbaar. De uitvoerders van methoden A en B passen de weging hierop aan: de minder goede score op tempo weegt minder zwaar mee dan de goede score op kwaliteit. De loonwaarde komt daarmee hoog uit. De uitvoerder van methode C past deze weging echter niet toe. Het tempoverlies drukt daarom relatief zwaar op de loonwaarde, die daardoor 20% lager uitkomt dan de door de andere methoden vastgestelde loonwaardes.

4.5 Verklaring 5. Niveau uitvoering: rol van de uitvoerder

In het voorgaande experiment stelden we vast dat uitvoerders verschillen in de manier waarop zij de meting afnemen. We zagen toen tijdens de observaties **drie verschillende rollen** van uitvoerders:

1. Een *begeleidende* rol, waarbij de uitvoerder de vragenlijsten doorneemt en/of toelicht en beschikbaar is bij vragen van werknemer en/of werkgever, maar verder geen invloed uitoefent op de gegeven scores.
2. Een *controlerende* rol, waarbij de uitvoerder de vragenlijsten doorneemt en/of toelicht, in gesprek met werkgever en werknemer de antwoorden bespreekt en de keuze voor een score controleert op juistheid en consistentie en deze indien nodig bijstelt.
3. Een *sturende* rol, waarbij de uitvoerder de vragenlijsten doorneemt en/of toelicht, in gesprek met werkgever en werknemer de antwoorden bespreekt en vervolgens zelf de best passende score kiest. Het definitieve antwoord wordt niet bij werkgever en werknemer gecontroleerd.

Het voorgaande experiment liet al zien dat deze verschillende rollen kunnen leiden tot verschillen in de vastgestelde loonwaardes. Wanneer een uitvoerder sturend is en zelf de best passende score bepaalt, kan het oordeel van de uitvoerder afwijken van dat van werkgever en werknemer. Wanneer de meting dan wordt herhaald met een uitvoerder die zich beperkt tot een begeleidende of controlerende rol, leidt dit tot een verschil in loonwaarde.

Binnen het huidige experiment zien we deze verschillende rollen van uitvoerders nog steeds terugkomen. Dat dit ook nu weer leidt tot verschillen in loonwaarde, blijkt uit onderstaand voorbeeld uit één van de cases.

Voorbeeld

De loonwaarde van een cliënt wordt bepaald door twee uitvoerders. Tijdens beide metingen wordt vastgesteld dat de cliënt niet erg assertief is. Aan de cliënt wordt in beide gevallen gevraagd of dit volgens hem leidt tot productieverlies. De cliënt geeft aan dat dit niet het geval is. De uitvoerder van methode A heeft een begeleidende rol en neemt dit antwoord over: de cliënt scoort op dit aspect 100%. De uitvoerder van methode B legt dit antwoord ter controle aan de werkgever voor. De werkgever geeft aan dat dit in bepaalde gevallen wel degelijk leidt tot tempoverlies, omdat hij zich regelmatig voor het karretje van collega's laat spannen. "Hij is altijd degene die koffie gaat halen: hij zal niet snel nee zeggen. Of als andere collega's vragen hebben, dan sturen ze hem naar mij toe: 'ga jij dat eens even vragen'. Dat zijn dan niet zijn vragen, maar hij is daardoor wel steeds wel van zijn werk weg." De cliënt beaamt dat dit inderdaad wel eens voorkomt, en dat hij daardoor soms minder snel werkt dan zijn collega's. De score wordt daarop naar beneden bijgesteld.

Validiteit: bijsturing en controle

De rol van uitvoerder zoals hierboven beschreven heeft betrekking op de (directe) invloed die de uitvoerder uitoefent op de gegeven scores. De mate van invloed varieert van zeer klein (in het geval van een begeleidende rol) tot zeer groot (in het geval van een sturende rol). Dit neemt niet weg dat een zekere mate van **bijsturing en controle** door de uitvoerder – onafhankelijk van zijn of haar rol bij de totstandkoming van de scores – vrijwel bij elke methodiek noodzakelijk is om de **validiteit** van de meting te waarborgen. Het gaat er dan niet om te controleren of werkgever en werknemer de 'juiste' scores geven, maar om vast te stellen of alle partijen 'het (nog) steeds over hetzelfde hebben'. Interpreteren werkgever en werknemer de norm, het scoresysteem en de gebruikte bestanddelen steeds op de juiste manier? Meet de loonwaardebepaling daarmee inderdaad wat deze wordt verondersteld te meten? Onderstaande voorbeelden uit de cases laten zien waarom bijsturing en controle op dit punt belangrijk is.

Voorbeeld

Voor een valide loonwaardemeting is het belangrijk dat de uitvoerder goed uitlegt hoe werkgever en werknemer het score-systeem moeten interpreteren. Een score van 100% betekent in het geval van een loonwaardebepaling: presenterend op regulier niveau (ofwel: functioneert voldoende). Een loonwaardebepaling onderscheidt zich daarmee nadrukkelijk van een beoordeling, waarin 100% zoveel betekent als: voldoet volledig aan de functie-eisen (ofwel: functioneert perfect). Met andere woorden: 100% is in de context van een loonwaardebepaling niet een score van 10 uit 10, maar geeft weer dat een werknemer voor 100% gelijk functioneert aan een andere (reguliere) collega. Voor werkgevers en werknemers is dit onderscheid niet altijd even duidelijk. Het is daarom belangrijk dat de uitvoerder dit goed uitlegt en hier ook tussentijds op stuurt of achteraf voor corrigeert.

Eén van de cases illustreert hoe verschillen in de vastgestelde loonwaardes kunnen optreden wanneer dit niet consequent gebeurt. De werkgever binnen deze case geeft aan: "er is altijd wel wat te ontwikkelen, bij iedereen. Daarom geef ik gewoon niet zo snel 100%". Bij het invullen van de vragenlijsten, geeft zij inderdaad opvallend vaak een score van 60% tot 80%. De uitvoerders van methoden A en B vragen hier verder niet op door en corrigeren dit niet. Er wordt door beide methoden daarom een verminderde loonwaarde vastgesteld.

De uitvoerder van methode C controleert wél consequent of de werkgever het scoresysteem op de juiste manier interpreteert. Zo vindt de werkgever dat de werknemer op verschillende gebieden erg goed presteert: "ze komt afspraken altijd heel goed na. Ze komt altijd ruim op tijd, is er vaak zelfs een kwartier van tevoren. Heel netjes en positief. Ik geef haar daar 80% voor." Uitvoerder: "Hoe is dat dan anders dan voor een reguliere kracht? Komen die nog vaker op tijd?" Werkgever: "Nee, ze doet dat juist heel goed, beter zelfs dan de meeste van haar collega's." De uitvoerder legt uit dat de werknemer daarmee dus op of zelfs boven regulier

niveau presteert, en stelt de score bij naar 100%. De werknemer blijkt verder nog wat problemen te hebben met kunnen aangeven van grenzen. Werkgever: "ze is eigenlijk bijna te lief, in elk geval weinig dominant. Ik zou zeggen: 50%." De uitvoerder van methode C vraagt hierop door: "leidt dit dan ook tot problemen op het werk? Is ze daardoor minder productief dan andere collega's?" De werkgever antwoordt hierop dat dit niet het geval is: "het is meer voor haarzelf vervelend, maar haar werk heeft er inderdaad niet onder te lijden. Het is ook eigenlijk niet echt nodig voor de functie. Het is ook wel een beetje kenmerkend voor de branche: heel veel collega's hebben dat, je moet je ook dienstbaar opstellen." De uitvoerder corrigeert de score op basis hiervan naar 100%. Aan het eind van de meting blijkt op deze manier dat de werknemer – hoewel de werkgever in eerste instantie steeds scores onder 100% heeft gegeven – in vergelijking met haar collega's helemaal niet verminderd productief is. Methode C komt, door de adequate controle en bijsturing van de uitvoerder, daarom uit op een loonwaarde van 100%.¹

Voorbeeld

Ook ten aanzien van de norm is een dergelijke controle en – indien nodig – bijsturing door de uitvoerder noodzakelijk.

Een cliënt werkt op de postkamer van een ministerie. Aan het begin van het gesprek is vastgesteld dat het hier een aangepaste functie betreft, en dat de norm van reguliere medewerker post en logistiek moet worden gehanteerd. Tijdens het gesprek blijkt echter dat de werkgever deze norm niet altijd toepast: "qua kwaliteit is er eigenlijk niet echt productieverlies. Onder druk maakt hij wel fouten, maar dat soort werkzaamheden vraag ik ook niet aan hem. Dat doen zijn collega's dan. Niet echt productieverlies dus." De uitvoerder van methode A stelt dit niet bij: kwaliteit wordt beoordeeld op 100%. De uitvoerder van methode B vraagt hier wel op door: "Maar in het werk komt dat dus wel voor? Dan zou er dus wel kwaliteitsverlies zijn als hij hier in de volledige functie zou werken?" De werkgever beaamt dit, en de score op kwaliteit wordt naar beneden bijgesteld.

4.6 Conclusie

Samenvattend zien we de volgende verklaringen voor de verschillen in de vastgestelde loonwaardes:

- De **richtlijnen** zoals die door het ministerie zijn vastgelegd, zijn in het algemeen goed geïmplementeerd door de methodieken. In een enkel geval gebeurde dit niet helemaal goed, dit ligt soms op het niveau van de methode (bestanddelen) en soms op het niveau van de uitvoering (norm).
- De grotere vergelijkbaarheid tussen methoden brengt – paradoxaal genoeg – ook **nieuwe verschillen** aan het licht: de berekeningswijze (hoe volgt uit de vastgelegde bestanddelen de loonwaarde), maar ook de weging (hoe moeten de verschillende bestanddelen ten opzichte van elkaar gewogen worden) varieert.
- **Uitvoeringsverschillen** tussen uitvoerders.

¹ Deze case illustreert daarmee dat uitschieters niet per definitie een 'onjuiste' loonwaarde weergeven en dat op basis van het experiment dus geen uitspraken kunnen worden gedaan over welke methode de 'beste' of juist de 'slechtste' loonwaarde vaststelt (zie ook de opmerkingen hierover in hoofdstuk 1).

5 Conclusies en aanbevelingen

5.1 Bescheiden succes, maar...

Het experiment geeft aanleiding te veronderstellen dat het vaststellen van richtlijnen en criteria waaraan alle methoden voor loonwaardebepaling moeten voldoen, leidt tot een meer uniforme loonwaardebepaling en daarmee tot vergelijkbaarder uitkomsten tussen de methoden. Het aanpassen en uniformeren van de methoden lijkt resultaat te hebben: verschillen in de vastgestelde loonwaardes zijn binnen de huidige ronde van het experiment minder vaak de regel en lijken vaker te wijten aan uitschieters, terwijl de uitkomsten vaker (nagenoeg) aan elkaar gelijk zijn. Een aantal methoden geeft nu bovendien vergelijkbare uitkomsten.

Het is dus zinvol aan deze richtlijnen vast te houden. Dat betekent dat een uniforme keuze wordt gemaakt ten aanzien van de **bestanddelen** die moeten worden gemeten om de loonwaarde van een werknemer te bepalen en ten aanzien van de **norm** waartegen deze arbeidsprestatie moet worden afgezet.

Dit betekent daarmee nog niet dat we nu kunnen concluderen dat de methoden betrouwbaar zijn en dat het dus niet meer uitmaakt wie de loonwaardebepaling uitvoert en welke methodiek hiervoor wordt gebruikt. Daarvoor zijn de verschillen over de hele linie nog altijd te groot.

5.2 Verdere uniformering methoden noodzakelijk

De grotere inhoudelijke overeenkomsten tussen methoden brengen – door de grotere vergelijkbaarheid – nieuwe verschillen aan het licht. De methoden meten nu vrijwel allemaal dezelfde bestanddelen, maar hoe zij vervolgens met deze bestanddelen omgaan om op basis hiervan de loonwaarde te bepalen, varieert. Door verschillen in berekeningswijze en weging ontstaan alsnog (grote) verschillen in de uitkomsten van de loonwaardebepalingen.

Het is daarom aan te bevelen de methodieken op deze punten verder te uniformeren, door richtlijnen te ontwikkelen voor de **berekeningswijze** (hoe volgt de loonwaarde uit de verschillende bestanddelen) en **weging** (hoe moeten de verschillende bestanddelen ten opzichte van elkaar worden gewogen) van de loonwaarde. Momenteel is nog niet duidelijk welke berekeningswijze en weging tot de meest valide uitkomsten leidt.

5.3 Niet alles op niveau van methoden op te lossen

Tegelijkertijd laat het experiment zien dat niet alles op het niveau van de methoden is op te lossen: uniformering lost weliswaar verschillen in berekening op, maar het is vervolgens aan de uitvoerders om de methodiek op uniforme wijze (dat wil zeggen: conform de richtlijnen) uit te voeren. De **persoon van de uitvoerder** is daarmee een **belangrijke schakel** in de verdere uniformering van de loonwaardebepaling.

We zien dit het duidelijkst terug in het **vaststellen van de norm** waartegen de arbeidsprestatie moet worden afgezet. Deze norm varieert per cliënt en daarmee per meting (elke

functie kent immers zijn eigen norm) en kan daarom niet in de methode worden 'ingebouwd'. Het is daarom aan de uitvoerder om deze norm op de juiste manier vast te stellen. Wanneer uitvoerders dit niet op dezelfde wijze doen, kan dit tot verschillen in de vastgestelde loonwaardes leiden. De invloed van de uitvoerder strekt echter verder dan het vaststellen van de norm alleen: ook **andere uitvoeringsverschillen** vormen een belangrijke verklaring voor de vastgestelde resterende verschillen in loonwaarde.

Het is daarom belangrijk om uitvoerders goed op hun taak toe te rusten en toegerust te houden. Dit onderstreept het belang van een goede opleiding en mogelijk ook certificering. In de voorgaande ronde van het experiment hebben we hier al aanbevelingen voor gedaan.

Het verplichtstellen van een **training** waaraan enkele minimale kwaliteitseisen worden gesteld, blijft als belangrijke aanbeveling staan. Uitvoerders moeten daarbij in elk geval worden getraind in het op uniforme wijze uitvoeren van de methodiek. Hierbij dient aandacht te worden besteed aan het vaststellen van de norm, de verschillende bestanddelen waaruit loonwaarde bestaat en de weging hiervan, de verschillende rollen die een uitvoerder tijdens de loonwaardebepaling kan aannemen en hoe deze (kunnen) doorwerken in de vaststelling van de loonwaarde en het belang van tussentijdse bijsturing en controle.

Urgentie buiten het experiment nog groter

In de alledaagse uitvoeringspraktijk – dat wil zeggen: buiten de experimentele setting – zal de urgentie van deze aanbeveling alleen maar toenemen. Binnen het experiment zijn de loonwaardebepalingen steeds uitgevoerd door een aan de methodiek gelieerde uitvoerder. In de praktijk kan de loonwaardebepaling vaak ook worden uitgevoerd door een andere, door de methode getrainde, uitvoerder van een externe organisatie (zoals de gemeente, het SW-bedrijf of een RIB). Deze uitvoerders staan niet alleen verder van de methode af, waarmee de kans op uitvoeringsverschillen toeneemt, maar hebben vaak ook een belang bij een bepaalde uitkomst van de loonwaardebepaling. We zien in de pilot loondispensatie dat dit belang ten koste kan gaan van de kritische en onafhankelijke houding van de uitvoerder. Dit heeft effect op de betrouwbaarheid van de loonwaardebepaling.

Aanwijzen één methode is niet de oplossing

De invloed van de uitvoerder betekent ook dat het aanwijzen en verplicht stellen van één of enkele loonwaardemethode(n) niet de oplossing is om tot vergelijkbare en uniforme resultaten te komen. De sleutelrol van de uitvoerder betekent dat verschillen zich niet alleen kunnen voordoen *tussen* methoden, maar dat de kans dat verschillen zich *binnen* methoden voordoen (bijna) net zo groot is.

De experimenten laten inderdaad zien dat uitvoeringsverschillen zich ook binnen methoden voordoen. Dit geldt voor alle onderzochte methoden.

Voorbeeld

Eén van de methoden is binnen het experiment uitgevoerd door drie verschillende uitvoerders. De manier waarop deze uitvoerders de methodiek uitvoerden, verschilt sterk.

Uitvoerder 1 baseert de loonwaardebepaling op een gesprek met werkgever en werknemer en inventariseert daarbij eerst alle deeltaken waaruit de functie van de cliënt bestaat. Vervolgens vraagt hij voor al deze deeltaken afzonderlijk naar het productieverlies op de bestanddelen tempo, kwaliteit en inzetbaarheid. Nadat hij heeft uitgelegd wat hiermee wordt bedoeld ("tempoverlies betekent dus: hoeveel procent werkt hij op deze taak minder snel dan zijn reguliere collega's"), neemt hij de scores van werkgever en werknemer zonder uitzondering over. Hij stelt zich daarmee begeleidend op.

Uitvoerder 2 maakt voor de loonwaardebepaling eveneens gebruik van een gesprek met werkgever en werknemer, maar maakt daarbij een minder duidelijke uitsplitsing naar de bestanddelen tempo, kwaliteit en inzetbaarheid. Hij vraagt in algemene termen naar het productieverlies op de verschillende taken ("wat is het productieverlies op deze taak?"), maar vraagt daarbij echter wel kritisch door ("U zegt, 50% productieverlies. Dat wil dus zeggen: ze doet de helft van de werkzaamheden die ze zou moeten doen. Dat ze in uren de helft doet van een valide soortgelijke en dat ze in de helft van de gevallen een beroep doet op een ander. Dat is vrij fors. Doen we haar daarmee niet tekort?"). De scores die werkgever en werknemer geven, worden op deze manier steeds door de uitvoerder gecontroleerd en indien nodig bijgesteld. Dit gebeurt altijd in overleg met werkgever en werknemer: zij moeten zich in de score kunnen vinden.

Uitvoerder 3 voert in het kader van de loonwaardebepaling ook gesprekken met werkgever en werknemer, maar vult deze gegevens aan met een observatie van de werkzaamheden van de werknemer. Zij meet daarbij hoeveel een reguliere werknemer in een bepaalde tijdseenheid presteert en zet dit af tegen de output die de cliënt in dezelfde tijd levert. Zij heeft bovendien een duidelijk sturende rol: zij bepaalt op basis van het gesprek en haar observaties zelf het juiste percentage voor elk van de uitgevraagde bestanddelen, zonder dit naar werkgever en werknemer terug te koppelen. In een aantal gevallen wijkt dit percentage tot soms wel de helft af van de door werkgever en werknemer genoemde percentages.

Dezelfde methode wordt zo door verschillende uitvoerders op geheel verschillende wijze uitgevoerd. We zien hierbij niet alleen een verschil in de rol die de verschillende uitvoerders aannemen (begeleidend, controlerend én sturend), maar ook in de methoden van dataverzameling die zij gebruiken (gesprekken versus gesprekken en observatie). Als de loonwaardemeting van cliënt 1 door uitvoerder 2 of 3 was uitgevoerd en vice versa, waren er mogelijk andere loonwaardes uitgekomen. Dit illustreert dat verschillen zich niet alleen voordoen *tussen* methoden, maar (soms minstens even sterk) ook *binnen* methoden.

Het experiment geeft daarmee geen aanleiding te veronderstellen dat het aanwijzen en verplicht stellen van één of enkele methode(n) zal leiden tot vergelijkba(a)r(d)e(r) loonwaardes. Het aanwijzen van één methode is daarmee **niet de oplossing** voor de problemen rondom het meten van loonwaarde. De oplossing moet deels worden gezocht in het verder uniformeren van de methoden onderling, maar ligt vooral ook *binnen* methoden in het opleiden en trainen van uitvoerders in het (blijvend) op uniforme wijze uitvoeren van de methodiek. Bovendien is het nodig meer inzicht te krijgen in de **validiteit** van de loonwaardemeting (weerspiegelt de uitkomst van een meting de werkelijke loonwaarde van een cliënt?). In paragraaf 5.4 gaan we hier nader op in.

Naleving

Naleving van de richtlijnen is ook een belangrijk aandachtspunt. Zelfs binnen de experimentele setting – waarin methoden extra aandacht aan de implementatie van de richtlijnen hebben besteed – zien we dat de richtlijnen niet altijd (goed) worden geïmplementeerd. Dit zal in de alledaagse uitvoeringspraktijk niet minder vaak het geval zijn. Een zeker toezicht en controle op de naleving van de richtlijnen is daarom geen overbodige luxe.

Dit **toezicht** zou zich minimaal moeten richten op de volgende punten:

- **Implementatie van de richtlijnen** met betrekking tot norm, bestanddelen, berekeningswijze en weging: wordt de loonwaardebepaling volgens de richtlijnen uitgevoerd?
- **Opleiding, training en certificering** van de uitvoerders: worden uitvoerders voldoende toegerust om de loonwaardebepaling volgens de richtlijnen te kunnen uitvoeren?

5.4 Tot slot: validering

De uitkomsten van het experiment wijzen erop dat het formuleren van en toezien op de implementatie van richtlijnen kan leiden tot meer uniforme en vergelijkbare uitkomsten tussen de methoden onderling. De kans dat de vastgestelde loonwaarde daarmee minder afhankelijk wordt van de gebruikte methode (dat wil zeggen: dat het geen verschil maakt met welke methode de loonwaarde wordt vastgesteld) neemt daarmee toe.

Dat leidt op zichzelf echter nog niet tot meer valide loonwaardebepalingen. Het experiment geeft immers geen antwoord op de vraag of de vastgestelde loonwaardes – hoewel deze nu beter vergelijkbaar zijn – ook echt de *werkelijke* loonwaarde van een cliënt weerspiegelen. Ofwel: is de loonwaarde die door de methoden wordt vastgesteld eigenlijk wel de *juiste* loonwaarde? Een verdere verbeterslag kan daarom worden gemaakt op het niveau van elke methode afzonderlijk: methoden hebben zelf een belangrijke verantwoordelijkheid om deze validiteit van hun loonwaardebepaling aan te tonen. Op dit moment is nog geen van de onderzochte methoden op deze manier gevalideerd. Enkele methoden zijn hier wel mee bezig door het aanleggen van een database waarop analyses kunnen worden uitgevoerd, maar dit is een langdurig proces dat voorlopig nog niet zal worden afgerond. Het merendeel van de methoden is hier nog niet mee begonnen en/of heeft hiervoor ook geen plannen.

Een laatste aanbeveling is daarom om aandacht te besteden aan dit proces van interne validering. In dit kader is te overwegen als aanvullend **kwaliteitscriterium** aan methoden te stellen dat zij inzichtelijk maken op welke manier zij werken aan de **validering** van hun methode. Zolang dit proces van validering nog niet is afgerond, is het belangrijk dat methoden inzichtelijk maken op welke manier zij de **objectiviteit en betrouwbaarheid** van hun metingen waarborgen. Juist omdat een loonwaardebepaling in veel gevallen per definitie subjectieve elementen bevat en sterk afhankelijk is van de persoon van de uitvoerder, is dit cruciaal om tot vergelijkbare, betrouwbare en uiteindelijk ook valide loonwaardebepalingen te komen.

Bijlage 1 Overzicht resultaten experiment 1

Tabel B.1 Overzicht percentage loonwaarde per cliënt naar methode

	Dariuz	VTA- methodiek	Activa	Arbolab	Loon- balans	Match- care	Methodiek UWV	Maximaal verschil
Cliënt 1	62%		41%	75%				34%
Cliënt 2	91%			78%	50%			41%
Cliënt 3	43%	65%			23%	61%	67%	44%
Cliënt 4		55%	72%		59%			17%
Cliënt 5	64%	70%	67%			72%		8%
Cliënt 6		60%		73%	43%		70%	30%
Cliënt 7		60%	79%	91%			100%	40%
Cliënt 8			54%		55%	77%		23%
Cliënt 9				84%		74%		10%
<i>Aant. Waarne- mingen</i>	4	5	5	5	5	4	3	3

Research voor Beleid
Bredewater 26
Postbus / P.O. Box 7001
2701 AA ZOETERMEER
tel: + 31 079 3 222 222
fax: + 31 079 3 222 212
The Netherlands
info@research.panteia.nl
www.research.panteia.nl

