



Validering loonwaardevaststelling

Deelrapport Resultaten, conclusies en aanbevelingen onderzoek
loonwaardevalidering (2013-2014). AKC, mei 2014



Inhoudsopgave

■	Voorwoord	4
■	Samenvatting resultaten loonwaardevalideringsonderzoek AKC	6
■	1 Totstandkoming rapport	9
■	2 Uitvoering loonwaardevalideringsproject AKC	11
■	3 Wetenschappelijke verantwoording	14
■	4 Resultaten	16
■	5 Aanbevelingen	20
■	Bijlage De hoofdelementen van de AKC-basissystematiek voor loonwaardebepaling	22

Dr. A.J.M. Schellart MBA, wetenschappelijk onderzoeker/bedrijfskundige
Drs. T.P.M. Hulsman, programmadirecteur AKC/projectleider loonwaardevalidering
M.T.F. Schouten MHD, arbeidsdeskundige/projectleider loonwaardevalidering AKC

© 2014, Arbeidsdeskundig Kennis Centrum, Nijkerk

Alle rechten voorbehouden.

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt,
in enige vorm of op enige wijze, zonder voorafgaande toestemming van het
Arbeidsdeskundig Kennis Centrum.

Over het AKC

Als onafhankelijk kennispartner verzamelt, ontwikkelt, verspreidt en evalueert het Arbeidsdeskundig Kennis Centrum (AKC) onderbouwde (arbeidsdeskundige) kennis en stimuleert het gebruik van die kennis. Het AKC bevordert de ontwikkeling van gevalideerde methoden voor gecertificeerde professionals op het terrein van Mens, Werk en Inkomen (preventie, begeleiding, participatie en claimbeoordeling). Het AKC ontwikkelt praktijkgerichte methoden en instrumenten. Ook valideert het AKC instrumenten en methoden door middel van een (wetenschappelijke) aanpak. Zie voor meer informatie website www.arbeidsdeskundigen.nl/akc/akc.php

Over het onderzoek

Het AKC is vanuit deze positie onder de titel 'Validering loonwaardemethoden' in mei 2013 een project gestart als vervolg op AKC Onderzoekscahier 6 Arbeidsdeskundige methodiek voor de vaststelling van gerealiseerde loonwaarde (mei 2012). Het AKC bouwt in dit project voort op sw adviezen van Panteia in het rapport Naar dezelfde Loonwaarde (oktober 2012). Hierin zijn de eerste stappen op weg naar meer uniformiteit gezet en werden basisafspraken voorgesteld over Tempo, Kwaliteit en Inzetbaarheid. De bandbreedte in de loonwaardevaststelling door verschillende methoden bleek met toepassing van deze begrippen echter nog zeer groot. Het AKC is daarom met een grote groep experts in gesprek gegaan over wat er nodig is om deze bandbreedte te verkleinen.

In het uitgevoerde onderzoek analyseert het AKC loonwaardevaststellingsmethoden voor de borging van een betrouwbare en valide vaststelling van loonwaarde, onafhankelijk van de uitvoerder en de methode.

Voorwoord

Deze beknopte projectrapportage presenteert de aanleiding, opzet, uitvoering en resultaten van het loonwaardevalideringsproject van het Arbeidsdeskundig Kennis Centrum (AKC). Dit project is uitgevoerd van mei 2013 tot en met mei 2014.

De rapportage is opgebouwd uit de volgende onderdelen:

- Samenvatting resultaten
- 1 Totstandkoming rapport
- 2 Uitvoering loonwaardevalideringsproject AKC
- 3 Wetenschappelijke verantwoording
- 4 Resultaten
 - Wetenschappelijke resultatenanalyse
 - Conclusies
- 5 Aanbevelingen

Samen met deze rapportage worden nog twee andere projectdocumenten gepubliceerd. Hierbij gaat het om:

- *AKC Cahier 13 A Validering loonwaardevaststelling. Deelrapport Inhoudsvalidatie methoden en opzet onderzoek loonwaarde validering. Standaardisering en uniformering van de kernbegrippen en module van loonwaardebepaling.* AKC, oktober 2013 (publicatie mei 2014). Hierin wordt de onderzochte basissystematiek van loonwaardevaststelling uitgebreid uitgelegd.
- *Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse.* PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014. Hierin worden de onderzoeksopzet en de resultaten wetenschappelijk gepresenteerd.

Het AKC bedankt de arbeidsdeskundigen, loonwaarde-experts, UWV, Dariuz, DWI en de onderzoekers, die samen dit unieke project mogelijk hebben gemaakt. Er is nuttig en baanbrekend werk verricht. De doelstelling van het valideringonderzoek was het verhogen van de inter-beoordelaarsbetrouwbaarheid en de mate van validiteit van loonwaardevaststelling. De resultaten geven aan dat dit doel gehaald is. Het AKC biedt andere loonwaardenmethoden ook de mogelijkheid aan om deel te nemen aan een dergelijke validering.

Met de resultaten van dit loonwaardevalideringsproject draagt het AKC bij aan de landelijke borging van betrouwbare en valide loonwaardebepaling. Alle regionale Werkbedrijven in wording en gemeenten kunnen er hun voordeel mee doen bij de uitvoering van de Participatiewet. Loonwaardebepaling kan ingezet worden als instrument om de arbeidsparticipatie van mensen met een grotere afstand tot de arbeidsmarkt te bevorderen.

*Mr. M. Klompé,
voorzitter bestuur AKC*

Samenvatting resultaten loonwaardevalideringsonderzoek AKC

Er zijn in Nederland nogal wat methoden in omloop om de loonwaarde vast te stellen. Hierbij gaat het om methoden die de loonwaarde vaststellen van medewerkers die als gevolg van beperkingen een lagere arbeidsproductiviteit hebben dan collega's zonder beperking in dezelfde functie. Er is echter te veel discussie over de vraag in hoeverre deze methoden hetzelfde meten en of ze allemaal in gelijke mate valide zijn. Dat dient noch het belang van de werknemers met beperkingen die economische waarde willen toevoegen, noch het ondernemerschap van de werkgevers die hun kansen willen bieden.

Met dit wetenschappelijk verantwoorde onderzoek van het Arbeidsdeskundig Kennis Centrum (AKC) ligt de weg open voor een valide (je meet wat je wil meten) loonwaardebepaling. Invoering van deze geteste nieuwe basissystematiek verhoogt de validiteit van de loonwaardevaststelling en de inter-beoordelaarsbetrouwbaarheid² – twee goed getrainde professionals komen binnen een beperkte bandbreedte tot eenzelfde resultaat – significant. Zo hebben de belangrijkste stakeholders – werknemers en werkgevers – de zekerheid dat de loonwaardevaststelling juist en vooral eenduidig plaatsvindt. En daar hebben ze ook recht op: de werknemer om de waardering te ontvangen die hem toekomt, de werkgever om te betalen voor wat wordt geleverd.

Een gezamenlijk overeengekomen basissystematiek voor loonwaardebepaling in combinatie met een gedegen training van de uitvoerders leidt tot een significante verbetering van de inter-beoordelaarsbetrouwbaarheid en validiteit van de loonwaardevaststellingen. Dat heeft het AKC nu wetenschappelijk verantwoord kunnen vaststellen. Dit is een doorbraak in de validering van loonwaardemethoden. De inhoud van deze basissystematiek is ontstaan door een bundeling van de kennis van enkele tientallen loonwaarde-experts – van onder meer UWV, Dariuz en VTA/DWI – door het AKC.

¹ *Validiteit: de mate waarin de methode meet wat hij zou moeten meten. Bij het onderzoeken van de validiteit wordt gekeken naar de mate waarin de resultaten van een methode (meetlat) en het te meten verschijnsel met elkaar overeenkomen. Validiteit is een gradatie; het is niet zo dat de ene methode valide is en de andere niet. Wel is de ene methode meer valide dan een andere. In dit onderzoek wordt de validiteit gemeten met behulp van een 'gouden standaard'. De resultaatcores zijn daarom in de mate van procentuele overeenstemming of in mate van procentuele afstand ten opzichte van die gouden standaard uitgedrukt.*

² *Inter-beoordelaarsbetrouwbaarheid (IBB): de mate waarin de scores van alle beoordelaars op eenzelfde meetvariabele overeenkomen. Daarbij geeft 1 volledig gelijke scores van alle beoordelaars aan en 0 volledig ongelijke scores. In onderzoek naar inter-beoordelaarsbetrouwbaarheid wordt een IBB van 0.7 of hoger bij professionals als een goede score gezien, omdat professionals een bepaalde beleidsruimte hebben voor de wijze waarop zij hun werk uitvoeren. Immers, niet ieder aspect van een door een professional te beoordelen situatie kan worden gestandaardiseerd.*

De inter-beoordelaarsbetrouwbaarheid (de variatie tussen beoordelaars) is met het gebruik van de basissystematiek zeer hoog (0,82 tot 0,94); de variatie is dus zeer klein. Voor dit type professionele beoordelingen is dit een uitzonderlijk goede score. Het percentage scores dat binnen een 10%-bandbreedte blijft ten opzichte van de gouden standaard stijgt bij de toepassing van de nieuwe basissystematiek met gemiddeld 52 procent. Dariuz en UWV zitten bij toepassing van de nieuwe basissystematiek gemiddeld circa 10 procent van de gouden standaard (juiste normprestatie Pi4) af. Zonder de nieuwe basissystematiek was dit voor UWV 13,3 procent en voor Dariuz 24 procent. VTA scoort 23,4 procent (dat was 28 procent). Als de prestatie wordt doorberekend naar een geldbedrag (Pi5) zit UWV gemiddeld 13,5 procent (was 22,5 procent) van de gouden standaard (juiste loonwaarde) af, Dariuz 18,6 procent (was 23,6 procent) en VTA 26,0 procent (was 36,9 procent). Forse verbeteringen dus bij de drie deelnemende loonwaardemethoden.

De resultaten tonen aan dat de (verminderde) prestatie van een werknemer met beperkingen – dankzij het gebruik van de nieuwe basissystematiek – objectiever en meer valide vast te stellen is door goed getrainde loonwaardedeskundigen. De basissystematiek werkt dus positief in de gewenste richting. De loonwaardebeoordelingen van dezelfde cases kennen een kleinere bandbreedte, een hogere beoordelaarsbetrouwbaarheid en zitten dicht bij de gouden standaard.

Aanbevelingen

- Regel voorafgaand aan de invoering van de Participatiewet (de decentralisatie naar gemeenten), de officiële landelijke introductie van de overeengekomen basissystematiek voor loonwaardevaststelling. Dit borgt de validiteit van de loonwaardevaststelling en de inter-beoordelaarsbetrouwbaarheid zo optimaal mogelijk. Alle gemeenten die de Participatiewet uitvoeren hebben hier baat bij.
- Stel training in de nieuwe basissystematiek verplicht en koppel de training aan certificering van loonwaardedeskundigen. Goed getrainde deskundigen bereiken namelijk een hogere beoordelaarsbetrouwbaarheid.
- Herhaal het onderling vergelijken van de kwaliteit van de methoden en de uitvoering van loonwaardebepalingen periodiek. Doe dit op de manier zoals het AKC heeft gedaan.
- De loonwaardemethoden die nog niet hebben deelgenomen aan dit onderzoek, zouden dat op korte termijn alsnog kunnen doen. Op deze manier ontstaat er een transparante vergelijking van de loonwaardebepaling door de verschillende methoden. Het AKC kan dit faciliteren.
- Gemeenten kunnen, als uitvoerder van de Participatiewet, bij hun keuze van de te hanteren loonwaardemethode een transparante vergelijking maken.

Leeswijzer

Alvorens de resultaten te presenteren, wordt de wetenschappelijke opzet van het onderzoek toegelicht om de lezer in de gelegenheid te stellen de waarde van dit loonwaarde-valideringsonderzoek te beoordelen.

Hoofdstuk 1 licht de inhoudsuniformering van loonwaardebepaling toe. *Hoofdstuk 2* beschrijft de uitvoering van het onderzoek: wie deed wat en hoe? *Hoofdstuk 3* gaat in op de verantwoording van de probleemstelling, vraagstelling en gekozen uitkomstmaten.

In *Hoofdstuk 4* worden de resultaten en conclusies op meerdere manieren gepresenteerd. Allereerst geeft *paragraaf 4.1* een wetenschappelijke resultatenanalyse en een toelichting. In *paragraaf 4.2* volgt een populaire versie van de weergave van de resultaten.

In *Hoofdstuk 5* worden de aanbevelingen geformuleerd voor beleidsmakers, gemeenten en (regionale) Werkbedrijven in wording die loonwaardebepalingsmethoden (gaan) hanteren.

1 Totstandkoming rapport

Sinds 2012 is duidelijk dat loonwaardebepaling een belangrijke rol kan gaan spelen in het vernieuwde uitvoeringsconcept Werk en Inkomen. Met het wegvallen van grote delen van de beschutte arbeidsplaatsen in de WSW, het niet meer kunnen detacheren vanuit de beschermende WSW-cao en het ontbreken van gesubsidieerde arbeid blijft de reguliere arbeidsmarkt over om de maatschappelijke arbeidsparticipatiedoelstellingen waar te maken. Het 'zuiver' kunnen bepalen van de waarde van de arbeidsprestatie van werknemers met arbeidsbeperkingen is – onder bovenstaande omstandigheden – een noodzakelijke randvoorwaarde voor succesvolle plaatsingen. Een eventuele quotumregeling kan ook effectiever zijn als er een gezonde balans is tussen de geleverde prestatie en het te betalen loon. Werkgevers en werknemers die samen zo'n beoogd quotum naleven, zien dan een fair loon voor een faire prestatie geborgd.

Voor de arbeidsdeskundige beroepsgroep (al jaar en dag specialisten op het gebied van loonwaardevaststelling) vormde dit de aanleiding om het AKC te vragen om een valideringstraject te starten voor de (arbeidsdeskundige) loonwaardesystematiek. Het AKC heeft de rijke expertervaring van arbeidsdeskundigen gecombineerd met de kennis en ervaring van experts van loonwaardemethoden en daar tevens een onafhankelijke en kritische wetenschappelijke blik aan toegevoegd. Drie groepen gebruikers van loonwaardemethodieken – Dariuz, VTA/DWI en UWV – bleken bereid om samen een basissystematiek op te stellen waaraan de vaststelling van loonwaarde moet voldoen. Zo ontstond een stevige basis voor de al langer gewenste consensus over de begripsinhoud van de loonwaardebepaling.

In vier sessies (29 mei, 4 juni, 11 juni en 11 juli 2013) is met een twintigtal loonwaarde-experts het consensusrapport met de basissystematiek voor loonwaardebepaling opgesteld. Dit rapport *Standaardisering en uniformering van de kernbegrippen en module van loonwaardebepaling* (AKC, oktober 2013) wordt als afzonderlijk rapport gelijktijdig gepubliceerd nu het loonwaarde-valideringsproject van het AKC in mei 2014 is afgerond. Het rapport bevat een op consensus van experts gebaseerde basissystematiek, die kan worden gezien als een expertnorm voor loonwaardebepaling. De basissystematiek is zodanig beschreven dat de systematiek overdraagbaar en trainbaar is.

De basissystematiek omvat zes deelelementen³:

- 1 Aantal uitgeoefende taken.
- 2 Normfunctie op basis van 60 procent-criterium en cao.
- 3 Normloon afgeleid uit normfunctie cao
- 4 Prestatiescore ten opzichte van de normprestatie (per taak en/of totale functie) vastgesteld met de begrippen (Tempo x Kwaliteit x Inzetbaarheid) x % tijd.
- 5 Additionele kosten.
- 6 Loonwaarde = normloon x prestatiescore x deeltijdfactor.

Het AKC veronderstelde dat deze afspraken set zou bijdragen aan het bereiken van meer inter-beoordelaarsbetrouwbaarheid en meer validiteit van (de) loonwaarde vaststellingsmethode(n).

De in het onderzoek betrokken loonwaardemethoden (Dariuz, UWV en VTA/DWI) hebben hun eigen methode aangevuld met deze basissystematiek. Daarmee conformeerden deze methoden zich aan de externe expertnorm. Zo kon worden onderzocht of het toepassen van de zes deelelementen van de basissystematiek de mate van inter-beoordelaarsbetrouwbaarheid en validiteit verhoogt.

Het AKC-project beoogde niet tot één methode te komen, maar tot uitwisselbaarheid van loonwaardemethoden op basis van het gezamenlijk overeengekomen afsprakenstelsel (de basissystematiek) die een betrouwbaardere⁴ loonwaarde vaststelling borgt. De resultaten van het AKC-onderzoek tonen aan dat dit is gelukt.

³ Zie voor een beknopte beschrijving van deze zes deelelementen Bijlage 3. De uitgebreidere beschrijving voor trainingsdoeleinden is te vinden in het separate rapport *Standaardisering en uniformering van de kernbegrippen en modules van loonwaardebepaling*. AKC, oktober 2013 (gepubliceerd juni 2014).

⁴ *Betrouwbaarheid (statistiek): de kans dat de gemeten waarde in een statistische proef gelijk is aan de werkelijke waarde. Veelal wordt daarbij uitgegaan van een 95%-betrouwbaarheid, dat wil zeggen een kans op een afwijking van 5%.*

2 Uitvoering loonwaardevalideringsproject AKC

De opzet van het AKC loonwaardevalideringsproject bood de mogelijkheid om in een laboratoriumsituatie de inter-beoordelaarsbetrouwbaarheid en de convergente validiteit⁵ te onderzoeken van loonwaardenmethodieken.

De eerste stap was, zoals hiervoor aangegeven, het beschrijven van de inhoud en de begrippen van de basissystematiek loonwaardebepaling. Dit werd vastgelegd in het rapport *Standaardisering en uniformering van de kernbegrippen en modules van loonwaardebepaling* (AKC, oktober 2013). Vervolgens zijn er dertien realistische loonwaardecases beschreven, met filmpjes van arbeidssituaties. Het ging hierbij om cases van werknemers met arbeidsbeperkingen in het bedrijfsleven en de publieke sector, die al geruime tijd in die functie werken (d.w.z. een eventuele leercurve al hebben doorlopen). Deze dertien casussen zijn beschreven op basis van de basissystematiek.

Door experts van de betrokken systemen is, onder begeleiding van het AKC, voor elk van de dertien cases een juiste loonwaarde berekend, de zogenaamde gouden standaard. Dat gebeurde door voor elke casus de zes deelelementen van loonwaardebepaling gezamenlijk vast te stellen. Deze gouden standaarden dienden als ijkpunten voor het vaststellen van de mate waarin een methode meet wat het moet meten (validiteit). Drie van de dertien cases zijn gebruikt voor trainingsdoeleinden. Tien cases zijn in de experimentfase gebruikt.

Daarna zijn de volgende stappen van het loonwaardevalideringsonderzoek – in de vorm van een Randomized Controlled Trial (RCT)⁶ – doorlopen:

- **Meting T0:** loonwaarde-experts bepalen de loonwaarde van de tien cases in een gecontroleerde laboratoriumopzet. Duur: drie dagen.
- **Training interventiegroep:** een gedegen AKC-training van twee dagen in de basissystematiek. De interventiegroep wordt getraind, de controlegroep wordt niet getraind.
- **Meting T1:** loonwaarde-experts bepalen nogmaals (in een andere volgorde) de loonwaarde van de tien casussen in dezelfde gecontroleerde laboratoriumopzet. Duur: drie dagen.

⁵ *Convergente validiteit: hierbij vergelijkt men verschillende methoden (of verschillende meetlatten) die beogen hetzelfde begrip te meten met elkaar.*

⁶ *Randomized Controlled Trial (RCT): een onderzoek waarbij de werkzaamheid of effectiviteit van een instrument of interventie wordt onderzocht door de deelnemers aan het onderzoek op basis van toeval toe te wijzen aan een groep die het instrument toepast (of een nieuwe behandeling ondergaat) en een groep die het instrument niet toepast (of de normale behandeling ondergaat), 'usual care'. In een RCT wordt de situatie voorafgaand aan een interventie (T0) en na de interventie (T1 en eventueel T2 en zo verder) gemeten bij zowel de interventiegroep als de controlegroep, om het effect van het instrument of de nieuwe behandeling te meten.*

- **Beoordeling:** beoordeling van de loonwaardevaststellingen in T0 en T1 door koppels van experts van de verschillende systemen volgens standaardmethode van dubbele, onafhankelijke beoordeling met daarna een consensusmeeting en monitoring van de wijze van beoordeling.
- **Analyse:** analyse van de data en uitspraken over mate van inter-beoordelaarsbetrouwbaarheid en de mate van overeenkomst met de gouden standaard (juiste loonwaarde), de validiteit.
- **Rapportage en advisering:** rapportage en adviezen per systeem.

De oorspronkelijke opzet van de efficacy pilot was een twee-armige Randomized Controlled Trial (RCT)⁷ met tien casussen en tien deelnemers in de interventiegroep (IG) en tien deelnemers in de controlegroep (CG) (op basis van een 'power analyse').

De deelnemers van Dariuz, VTA en UWV (Sociaal Medische Zaken (SMZ) en WERKbedrijf) hebben de volgende stappen doorlopen:

- Random indelen in interventiegroep (IG) en controlegroep (CG).
- T0-meting (IG en CG) op basis van tien casussen.
- Training interventiegroep (IG) op basis van drie casussen.
- T1-meting (IG en CG) op basis van dezelfde tien casussen als T0 (in random volgorde).

Bij een twee-armige RCT en tien casussen zijn tien deelnemers nodig in IG en tien deelnemers in CG per loonwaardesystematiek (VTA, Dariuz, UWV) (op basis van een 'power analyse'). In de praktijk bleken er echter onvoldoende loonwaarde-experts beschikbaar om een volledige twee-armige RCT uit te voeren. De beschikbaarheid van 33 loonwaardedeskundigen (Dariuz: 10, VTA/DWI: 6, UWV: 17), leidde voor Dariuz en VTA/DWI tot een voor- en nameting (15 IG en 1 CG) en voor UWV tot een RCT (random toegedeeld: 8 IG en 9 CG, gestratificeerd naar SMZ/WERKbedrijf).

Met de zestien arbeidsdeskundigen van VTA en Dariuz is een voor- en nameting gehouden. Een meting voorafgaande aan de training (T0) en een meting na de training (T1). Voor UWV kon vanwege een voldoende aantal deelnemers een Randomized Controlled Trial (RCT) worden ingericht met een interventiegroep en een controlegroep: zeventien deelnemers zijn random toegewezen aan de experimentgroep (acht deelnemers) en aan de controlegroep (negen deelnemers). De experimentgroep kreeg de training wel, de controlegroep niet. Bij beide groepen werd een T0- en een T1-meting uitgevoerd. In totaal hebben 23 van de 33 deelnemers aan de pilot de training gevolgd. Eén deelnemer van VTA heeft de beoogde training niet gevolgd, maar heeft wel meegedaan aan de T0- en T1-metingen.

De twee metingen leverden $2 \times 10 \times 33 = 660$ loonwaardeberekeningen op. Die zijn geanalyseerd op drie van de zes in de basissystematiek vooraf gedefinieerde elementen van de loonwaardebepaling, te weten: Pi3 = het normloon, Pi4 = de prestatie t.o.v. de norm en Pi5 = de loonwaarde.

⁷ Twee-armige RCT: een Randomized Controlled Trial waarbij de deelnemers op basis van toeval ('random') zijn toegewezen aan twee groepen: een interventiegroep en een controlegroep. Er kunnen ook drie-armige of vier-armige RCT's zijn, waarbij meerdere, verschillende interventiegroepen worden vergeleken met een controlegroep.

De stappen van het loonwaardevalideringsonderzoek (Efficacy ⁸ onderzoek)		
Stap 1 (september 2013)	Samenstelling experimentgroep per methode die meedoet in de validatie.	Samenstelling controlegroep met dezelfde specificaties als de experiment groep.
Stap 2 = T0 (november 2013)	Experimentgroep maakt casuïstiek in een laboratoriumsituatie*.	Controlegroep maakt casuïstiek in een laboratoriumsituatie*.
Stap 3 (november 2013)	Training in de nieuwe basissystematiek (inhoud validatie-afspraken).	Geen training in de nieuwe basissystematiek (inhoud validatie-afspraken).
Stap 4 = T1 (november 2013)	In laboratoriumsituatie maken van dezelfde casuïstiek als bij stap 1, maar in een andere volgorde*.	In laboratoriumsituatie maken van dezelfde casuïstiek als bij stap 1, maar in een andere volgorde*.
Stap 5 = beoordeling (december 2013 – maart 2014)	Beoordeling van de loonwaardevaststellingen in T0 en T1 door koppels van experts van de verschillende systemen volgens de standaardmethode van dubbele consensusbeoordeling en monitoring.	
Stap 6 = analyse data (april - mei 2014)	Analyse van de data en uitspraken over de mate van bandbreedte ten opzichte van de gouden standaard.	
Stap 7 = advies (mei-juni 2014)	Advies en valideringsrapport. Adviezen en waardering van de methoden ten opzichte van de gouden standaard in termen van de mate van validiteit en inter-beoordelaarsbetrouwbaarheid.	

*voor experimentgroep en controlegroep dezelfde casuïstiek op hetzelfde moment

De uitgebreide beschrijving van de opzet van het loonwaardevalideringsproject van het AKC is te vinden in hoofdstuk 5 van AKC Cahier 13 A Validering loonwaardevaststelling. Deelrapport Inhoudsvalidatie methoden en opzet onderzoek loonwaarde validering. Standaardisering en uniformering van de kernbegrippen en modules van loonwaardebepaling. AKC, oktober 2013 (publicatie mei 2014).

⁸ Efficacy: de werkzaamheid van een instrument of interventie. Deze werkzaamheid wordt onderzocht in een laboratoriumsituatie, waarbij bepaalde belangrijke omstandigheden constant kunnen worden gehouden, dan wel kunnen worden 'gemanipuleerd' door de onderzoekers. Dit in tegenstelling tot effectiviteitsonderzoek, dat het effect van een instrument of interventie in de werkelijkheid onderzoekt.

3 Wetenschappelijke verantwoording

Door de gekozen opzet van het onderzoek kunnen voor het eerst, wetenschappelijk verantwoord, uitspraken over loonwaardebepaling worden gedaan.

De probleemstelling is drieledig:

- 1 Inzicht in de betrouwbaarheid en validiteit van de diverse methoden voor loonwaardebepaling, indien toegepast door loonwaardedeskundigen.
- 2 Inzicht in de betrouwbaarheid en validiteit van de overeengekomen basissystematiek voor loonwaardebepaling, indien toegepast door loonwaardedeskundigen.
- 3 Inzicht of de overeengekomen systematiek voor loonwaardebepaling een verbetering betekent wat betreft betrouwbaarheid en validiteit.

De probleemstelling van het onderzoek is vertaald in drie onderzoeksvragen:

- 1 Hoe groot is de inter-beoordelaarsbetrouwbaarheid (IBB) op T0 respectievelijk T1 op organisatieniveau (methode)? Neemt de IBB toe op T1 en is het verschil met T0 significant?
- 2 Wat is de mate van overeenstemming tussen de bepaalde loonwaardescores en de gouden standaard (juiste loonwaarde) op T0 en T1? Neemt deze overeenstemming toe op T1 en is het verschil met T0 significant?
- 3 Wat is de afstand tussen de bepaalde loonwaardescores en de gouden standaard (juiste loonwaarde) op T0 en T1? Neemt deze afstand af op T1 en is het verschil met T0 significant?

De nieuwe basissystematiek loonwaardebepaling kent, zoals eerder aangegeven, zes deelelementen. Daarvan vormen Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde) het hart. Daarom is deze eerste resultatenanalyse gebaseerd op Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde).

De uitkomstmaten met betrekking tot betrouwbaarheid en validiteit zijn voor Pi3, Pi4 en Pi5 afzonderlijk bepaald. Pi5 is namelijk samengesteld uit Pi3 x Pi4 ($Pi5 = Pi3 \times Pi4$ gecorrigeerd voor een eventuele deeltijdfactor).

Er is gewerkt met de volgende uitkomstmaten:

- inter-beoordelaarsbetrouwbaarheid (betrouwbaarheid);
- het aantal scores binnen 10%-bandbreedte ten opzichte van de gouden standaard (juiste loonwaarde) (validiteit);
- de absolute procentuele afstand tot de gouden standaard (juiste loonwaarde) (= $100\% \times \text{absolute verschil/gouden standaard}$) (validiteit).

Indien een +/- bandbreedte was afgesproken bij de gouden standaard, is uitgegaan

van het midden van die bandbreedte. Dit is een conservatieve benadering van de validiteit.

Sterke elementen van de opzet van de efficacy pilot zijn:

- een loonwaardesystematiek met 'inhoudsvaliditeit'⁹;
- door experts opgestelde 'real life' casussen;
- gouden standaarden voor casussen met 'inhoudsvaliditeit';
- drie duidelijke uitkomstmaten (inter-beoordelaarsbetrouwbaarheid, de mate van overeenstemming met en de afstand tot de gouden standaard;
- een uitgevoerd procesevaluatie;
- de deelname van drie loonwaardesystematieken.

De mogelijke zwaktes van het onderzoek zijn nader bekeken en konden veelal zonder de opzet van het onderzoek te verstoren worden gecompenseerd. Het ging daarbij om:

- alleen voor UWV een RCT;
- grotendeels alleen voor- en na meting;
- missings op T0 voor Pi3
- een klein aantal deelnemers per organisatie (met name VTA);
- mogelijk confounding¹⁰ tussen organisaties door verschil in kenmerken van de deelnemers;
- slechts gedeeltelijk rekening gehouden met clustering van geneste gegevens.

Voor een gedetailleerdere verantwoording wordt verwezen naar de *Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse*. PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014.

⁹ *Inhoudsvaliditeit ('content validity'): bij inhoudsvaliditeit wordt onderzocht of de empirische meetlat het gehele theoretische begrip dekt. Loonwaarde is een theoretisch begrip met een breed domein. Dat is door de experts uiteengegafeld in zes empirische indicatoren, die zijn gekoppeld aan de stappen voor de vaststelling van loonwaarde. Zo zijn afspraken gemaakt voor gegevensverzameling, het vaststellen van ijkpunten (normfunctie en normloon), het vaststellen van de prestatie (Tempo, Kwaliteit en Inzetbaarheid), rekenformules ($T \times K \times I$) en kosten (Additionele kosten) die niet in de loonwaarde verrekend worden. Een methode behoort alle aspecten van het te onderzoeken begrip te meten, indien het een representatieve weerspiegeling van het begrip wil geven. In dit project is zeer nauwgezet met de inhoudsvaliditeit omgegaan door eerst met alle experts een rapport op te stellen met alle te meten aspecten van het begrip 'loonwaarde'.*

¹⁰ *Confounding: er is sprake van confounding wanneer een derde factor, die gerelateerd is aan zowel de determinant als de uitkomst, het causale verband tussen die twee verstoort.*

4 Resultaten

4.1

Wetenschappelijke resultatenanalyse

4.1.1

Inter-beoordelaarsbetrouwbaarheid op T0 en T1 (> 0,7 is ruim voldoende tot goed¹¹)

De vraagstelling luidde: hoe groot is de inter-beoordelaarsbetrouwbaarheid (IBB) op T0 respectievelijk T1 op organisatieniveau? Neemt de IBB toe op T1 en is het verschil met T0 significant?

Uit de onderstaande tabel, waarbij ICC staat voor Inter Class Correlation, blijkt:

- De IBB is al hoog op T0 (> 0,7), maar wordt toch nog significant hoger op T1 voor de interventiegroep bij Pi4 (0,85) en Pi5 (> 0,9), niet bij Pi3.
- De IBB van de controlegroep blijft op T1 nagenoeg gelijk aan T0.
- Op T1 is de IBB van Pi3 lager dan die van Pi4 en Pi5.
- Op T1 heeft de UWV-interventiegroep de hoogste IBB voor Pi3 (0,821) en Pi5 (0,940), Dariuz voor Pi4 (0,942) met nagenoeg ex aequo voor de UWV-interventiegroep (0,914).

Resultaten IBB op T0 en T1 (uitkomstmaat ICC)							
T0	10 casussen T0			10 casussen T1			N
Totaal	Pi3	Pi4	Pi5	Pi3	Pi4	Pi5	
IG	0,809	0,728	0,781	0,737	0,859	0,901	23
CG	0,746	0,765	0,773	0,745	0,780	0,769	10
Dariuz	0,815	0,885	0,907	0,734	0,942	0,918	10
VTA	0,863	0,597	0,715	0,674	0,692	0,847	6/5
UWV	0,777	0,810	0,808	nvt	nvt	nvt	17
UWV-IG	0,815	0,853	0,842	0,821	0,914	0,940	8
UWV-CG	0,746	0,772	0,777	0,741	0,838	0,788	9
P*	0,168	0,000	0,000	0,036	0,000	0,000	
*significantie:	HOOGSTE	t.o.v.	LAAGSTE	ICC als testwaarde			

Legenda: Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde)

Zie voor verdere toelichting sheet 22 van de *Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse*. PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014.

¹¹ In onderzoek naar inter-beoordelaarsbetrouwbaarheid wordt een IBB van 0.7 of hoger bij professionals al als een goede score gezien, omdat professionals een bepaalde beleidsruimte hebben voor de wijze waarop zij hun werk uitvoeren. Immers, niet ieder aspect van een door een professional te beoordelen situatie kan worden gestandaardiseerd.

Conclusie

De scores van alle beoordelaars op de meetvariabelen Pi3, Pi4 en Pi5 komen goed overeen. Een score van 1,0 staat gelijk aan volledig gelijke scores van alle beoordelaars. De scores van de UWV-interventiegroep, Dariuz en VTA van boven de 0,8 en zelfs boven de 0,9 bij UWV en Dariuz, tonen aan dat op inter-beoordelaarsbetrouwbaarheid hoog gescoord wordt. De toename van de inter-beoordelaarsbetrouwbaarheid bij de tweede meting toont aan dat toepassing van de basissystematiek uniformerend werkt bij loonwaardebepaling.

4.1.2

Resultaten in termen van de mate van overeenstemming met de gouden standaard (juiste loonwaarde) (+/- 10%) (hoe hoger hoe beter)

De vraagstelling luidde: wat is de mate van overeenstemming tussen de scores en de gouden standaard (+/- 10%) op T0 en T1? Neemt deze overeenstemming toe op T1 en is het verschil met T0 significant?

Onderstaande tabel laat zien dat:

- de overeenstemming met de gouden standaard (+/- 10%) bij de interventiegroep op T1 significant hoger is dan op T0;
- de overeenstemming met de gouden standaard (+/- 10%) bij de controlegroep op T1 niet significant verschilt van T0;
- op T1 de grootste overeenstemming met de gouden standaard (+/- 10%) door Dariuz wordt bereikt voor Pi3 (74%) en nagenoeg ex aequo voor de UWV-interventiegroep (73,8%);
- op T1 de UWV-interventiegroep de grootste overeenstemming heeft bereikt met de gouden standaard (+/- 10%) voor Pi4 (68%) en Pi5 (53%).

Overeenstemming met gouden standaard binnen 10% bandbreedte over tien casussen							
	Pi3-T0	Pi4-T0	Pi5-T0	Pi3-T1	Pi4-T1	Pi5-T1	N (resp)
Dariuz	45,0%	26,0%	30,0%	74,0%	61,0%	45,0%	10
VTA	34,0%	32,0%	24,0%	50,0%	54,0%	36,0%	5
UWVCG	52,2%	56,7%	36,7%	56,7%	56,7%	34,4%	9
UWVIG	57,5%	60,0%	32,5%	73,8%	67,5%	52,5%	8
IG	47,0%	39,1%	29,6%	68,7%	61,7%	45,7%	23
CG	50,0%	53,0%	36,0%	56,0%	52,0%	33,0%	10
Percentage Pi significant verbeterd op T1 t.o.v. T0							

Legenda: Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde)

Zie voor verdere toelichting sheet 27 van de *Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse*. PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014.

Conclusie

De validiteit van de loonwaardebepaling neemt bij toepassing van de basissystematiek significant toe. De validiteit kan nog verder worden aangescherpt. Dariuz en UWV zijn dicht bij het doel dan VTA. Zij zitten dus op de goede weg.

4.1.3

Resultaten afstand t.o.v. gouden standaard (hoe lager, hoe beter)

Wat is de afstand tussen de scores en de gouden standaard op T0 en T1? Neemt deze afstand af op T1 en is het verschil met T0 significant?

- De afstand tussen de loonwaardescores en de gouden standaard is bij de interventiegroep op T1 lager dan op T0.
- De afstand tussen de loonwaardescores en de gouden standaard is bij de controlegroep op T1 niet significant verschillend van T0.
- Op T1 is de afstand tussen de loonwaardescores en de gouden standaard het laagst bij de UWV-interventiegroep voor Pi3 (6%), voor Pi4 (10%) en Pi5 (13,5%). Dariuz scoort op Pi4 nagenoeg gelijk als de UWV-interventiegroep.

Gemiddelde relatieve afstand tot de drie Pi's van de gouden standaard (juiste waarde), in termen van het procentueel absoluut rekenkundig gemiddeld verschil t.o.v. de tien gouden standaarden

	Pi3	Pi4	Pi5	Pi3	Pi4	Pi5
Dariuz	14,4%	24,0%	23,6%	8,7%	11,0%	18,6%
VTA	15,1%	28,6%	36,9%	12,3%	23,4%	26,0%
UWVCG	12,6%	14,6%	27%	11,2%	12,7%	22,5%
UWVIG	10,8%	13,3%	22,5%	6,0%	10,1%	13,4%
IG	13,2%	21,3%	26,1%	11,7%	13,4%	18,4%
CG	12,9%	14,7%	26,7%	11,6%	14,4%	23,6%
HOE LAGER HOE BETER						

Legenda: Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde)

Zie voor verdere toelichting sheet 29 van de *Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse*. PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014.

Conclusie

De prestatie in procenten in vergelijking tot de normprestatie (Pi4) wordt door UWV en Dariuz in vrijwel gelijke mate vastgesteld ten opzichte van de gouden standaard: beide organisaties zitten gemiddeld circa 10% van de gouden standaard (juiste normprestatie) af. Als de prestatie wordt doorberekend naar een geldbedrag (Pi5) zit UWV gemiddeld 13,5% van de gouden standaard (juiste loonwaarde) af, Dariuz 18,6% en VTA 26,0%.

4.2

Samenvattende conclusies

- Bij de toepassing van de basissystematiek neemt de inter-beoordelaarsbetrouwbaarheid toe tot een niveau dat ruim voldoende tot goed is.
- Bij de toepassing van de basissystematiek neemt de validiteit toe, maar die kan nog scherper wat betreft het percentage loonwaardebepalingen dat binnen een bandbreedte (+/- 10%) van de gouden standaard valt.
- De kleinste afstand van de door de deelnemers in het onderzoek bepaalde scores op de drie in de analyse betrokken Pi's ten opzichte van de gouden standaard, ligt onder of net boven de 10%. Dat is het geval voor de UWV-interventiegroep voor Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde).
- De afstand van scores van Dariuz op de drie in de analyse betrokken Pi's ten

opzichte van de gouden standaard ligt voor Pi3 (normloon), Pi4 (de prestatie t.o.v. de norm) onder of net boven de 10% en voor Pi5 (de loonwaarde) duidelijk boven de 10%, maar beneden de 20%.

- De afstand van scores van VTA/DWI ten opzichte van de gouden standaard ligt voor Pi3 (normloon) net boven de 10% en voor Pi4 (de prestatie t.o.v. de norm) en Pi5 (de loonwaarde) boven de 20%.
- De overeengekomen basissystematiek leidt tot een duidelijke verbetering van de inter-beoordelaarsbetrouwbaarheid en de mate van validiteit van de loonwaardevaststelling.
- De doelstelling van het loonwaardevalideringsonderzoek van het AKC was het verhogen van de inter-beoordelaarsbetrouwbaarheid en de mate van validiteit van loonwaardevaststelling. De resultaten geven aan dat dit doel gehaald is.

5 Aanbevelingen

De onderzoekresultaten laten zien dat de gezamenlijk afgesproken basissystematiek leidt tot loonwaarden die dicht bij de gouden standaard (juiste loonwaarde) liggen. Het is daarom aan te bevelen om nu over te gaan tot een officiële landelijke introductie van de overeengekomen basissystematiek voor loonwaardevaststelling. Dit is een eerste logische vervolgstap voor de borging van de kwaliteit van loonwaardevaststelling. Daarnaast wordt geadviseerd om via een meester-gezelconstructie de kwaliteit te monitoren en training op basis van ervaring te borgen.

Ten aanzien van de wenselijke sturing op betrouwbare en valide loonwaardebepalingen wordt op basis van de onderzoekresultaten voorgesteld om meerdere borgingslijnen te combineren. Uit de onderzoeksresultaten blijkt bijvoorbeeld dat sturen op de inter-beoordelaarsbetrouwbaarheid, die hoog is in het onderzoek, nog niet garandeert dat de loonwaardebepalingen gemiddeld binnen een bandbreedte van 10% van de gouden standaard blijven.

Het gaat om de volgende borgingslijnen:

- **Lijn 1:** invoering landelijke basissystematiek voor loonwaardebepaling op basis van het loonwaardevalideringsonderzoek van het AKC (inter-beoordelaarsbetrouwbaarheid). Alle loonwaardemethoden hanteren dan binnen hun methode dezelfde basissystematiek. Daarmee zijn ze vergelijkbaar en op eenzelfde manier te toetsen op inter-beoordelaarsbetrouwbaarheid en validiteit.
- **Lijn 2:** certificering van de uitvoerders door verplichte training en periodieke proeve van bekwaamheid (inter-beoordelaarsbetrouwbaarheid). Professionals die een loonwaardebepaling uitvoeren, moeten in termen van training en ervaring een periodieke proeve van bekwaamheid afleggen. Hierdoor wordt de noodzakelijke basisdeskundigheid voor loonwaardebepaling geborgd. Het onderzoek heeft nog eens duidelijk onderstreept dat loonwaardebepaling een specialisme is.
- **Lijn 3:** periodieke validering van methoden via de gouden standaard casuïstiek-methodiek, zoals uitgevoerd door het AKC.

Valideringsonderzoek biedt de meeste garantie op een betrouwbare en valide loonwaardevaststelling. Dat biedt de mogelijkheid de inter-beoordelaarsbetrouwbaarheid en de mate van overeenstemming respectievelijk de afstand tussen de scores en de gouden standaard vast te stellen. Het AKC heeft juist daarom het valideringstraject (met gouden standaardcasuïstiek) gestart.

Het AKC biedt andere loonwaardenmethoden ook de mogelijkheid aan om deel te nemen aan een dergelijke validering.

Bijlagen

- 1 Losse publicatie: AKC Onderzoekscahier 13 A Validering loonwaardevaststelling Deelrapport Inhoudsvalidatie methoden en opzet onderzoek loonwaardevalidering Standaardisering en uniformering van de kernbegrippen en modules van loonwaardebepaling. AKC, oktober 2013 (publicatie mei 2014).
- 2 Losse publicatie: Wetenschappelijke verantwoording uitvoering efficacy pilot en resultatenanalyse. PowerPointpresentatie AKC, Dr. Ton Schellart MBA, mei 2014.
- 3 De hoofdelementen van de AKC-basissystematiek voor loonwaardebepaling.

Bijlage 3 De hoofdelementen van de AKC-basissystematiek voor loonwaardebepaling

1 = aantal uitgevoerde taken

Exacte beschrijving van de taken die daadwerkelijk worden uitgevoerd in termen van een beknopte taakomschrijving en de omvang van de taak binnen de uitgevoerde arbeid om de juiste normfunctie te kunnen bepalen. De kleinste taak is minimaal een half uur per dag.

2 = Norm cao

Stappenplan voor het zoeken van de normfunctie binnen het bedrijf, de bedrijfs-cao, sector-cao, een andere cao of taakvergelijking van meerdere functies binnen cao('s). De normfunctie heeft minimaal 60% overlap qua taken en omvang van die taken met de daadwerkelijk uitgevoerde taken.

3 = Normloon

Vaststellen van normloon binnen de cao-schaal van de normfunctie op basis van leeftijd, opleiding en ervaring van de persoon waarvoor de loonwaarde wordt bepaald.

4 = Prestatiescore (per taak en/of totale functie) vastgesteld met de begrippen (Tempo x Kwaliteit x Inzetbaarheid) x % tijd.

Uniforme en transparante definities van de begrippen waarmee de arbeidsprestatie wordt bepaald: Tempo, Kwaliteit en Inzetbaarheid. Daarnaast is het begrip Additionele kosten eenduidig gedefinieerd om te voorkomen dat die meegewogen worden bij Tempo, Kwaliteit of Inzetbaarheid.

De verschillende methoden kunnen op hun eigen manier de gegevens verzamelen die bepalend zijn om een juiste inschatting te maken van Tempo, Kwaliteit, Inzetbaarheid en Additionele kosten. De methoden moeten wel transparant maken welke

gegevens per begrip gehanteerd worden. Ook moeten ze kunnen aantonen dat factoren die van invloed zijn op het bepalen van de prestatie op de juiste manier en maar op één plek verwerkt worden: óf bij Tempo óf bij Kwaliteit óf bij Inzetbaarheid óf bij Additionele kosten.

De gegevens worden onder meer verzameld op basis van een bezoek aan de werkplek, waarneming van de functie-uitoefening en in samenspraak met werkgever en werknemer. Het verzamelen van de gegevens kan per taak geschieden of over de functie als geheel in een keer. Per begrip Tempo, Kwaliteit en Inzetbaarheid wordt vastgesteld wat de waarde van de prestatie is ten opzichte van de prestatie die gevraagd wordt in de normfunctie.

Duidelijke rekenregels voor bepaling totaal loonwaarde: de totaalprestatie in de functie wordt berekend met de formule $T \cdot K \cdot I = TP$ (totale prestatie). Als er een analyse per taak heeft plaatsgevonden, vindt vermenigvuldigen plaats per taak $T \cdot K \cdot I \cdot (BT)^{12} = PT$. De optelsom van de prestaties per taak bepaalt dan de totale prestatie.

5 = Additionele kosten

De extra kosten die direct verbonden zijn met het functioneren van de werknemer worden wel vastgesteld, maar niet meegenomen in de loonwaardebepaling. Additionele kosten kunnen zowel eenmalige als periodiek terugkerende kosten zijn.

6 = Loonwaarde

De loonwaarde is dan $TP \cdot NL \cdot \text{deeltijdfactor} = LW$ (loonwaarde).

Hoe de loonwaarde wordt bepaald is dan duidelijk en transparant. Sturing op het goed toepassen van die afspraken binnen een methode en de juiste uitvoering door de uitvoerder kan alleen door de borgingslijnen 2 en 3 ook toe te passen.

¹² BT = Bijdrage Taak in de functie = % van de werktijd per taak.



Arbeidsdeskundig Kennis Centrum
Postbus 1058
3860 BB Nijkerk
(033) 247 34 57