

White Paper

Moderne Assessment Tools

Auteur:



Bas van de Haterd

professioneel bemoeial

Voorwoord

In deze white paper heb ik mijn ervaringen beschreven met alle vormen van moderne, digitale, assessment tools. De reis om meer te leren over deze technologieën is vele jaren geleden begonnen en ik heb het geluk gehad dat ik veel verschillende tools heb mogen testen gedurende deze jaren. In deze white paper zal ik mijn bevindingen over de verschillende tools en technologieën delen.

Ik heb besloten geen namen van leveranciers te noemen. Niet alle leveranciers hebben me de mogelijkheid geboden om hun tool te testen. Hoewel ik uiteraard teleurgesteld ben over hun gebrek aan transparantie, vind ik het niet gepast om hen te diskwalificeren op basis van de terughoudendheid van een enkele sales of marketing manager.

Ik heb besloten om te schrijven over mijn eigen ervaringen met de technologie in een breder perspectief. Deze white paper geeft een algemeen overzicht van de verschillende technologieën, hun geschiedenis, hun sterktes, zwaktes en potentieel, maar ook mijn eigen ervaringen met de technologie. In mijn belevenis laten ervaringen zich het best omschrijven met concrete voorbeelden en ik heb besloten mijzelf en mijn eigen ervaringen hiervoor te gebruiken.

Dus daar waar dit white paper is geschreven op een beschouwende manier, zijn er duidelijk persoonlijke ervaringen in verwerkt om het tastbaarder te maken. Aangezien ik zeer veel van deze assessments getest heb en veel van deze testen met beperkte afwijkingen hetzelfde uitwijzen, durf ik de stelling wel aan dat grote afwijkingen aan de test liggen. Je zou kunnen stellen dat al deze assessments mijn zelfkennis sterk hebben doen toenemen.

Introductie

De afgelopen jaren zijn er veel nieuwe assessment tools op de markt gekomen die een enorme impact zullen hebben op alle aspecten van het HR vakgebied. Of het nu gaat om recruitment, training, ontwikkeling en zelfs de (geestelijke) gezondheid van medewerkers, deze tools zullen op al deze gebieden van invloed zijn. Persoonlijk denk ik zelfs dat het onderwijs enorm kan profiteren van de mogelijkheden die deze tools bieden.

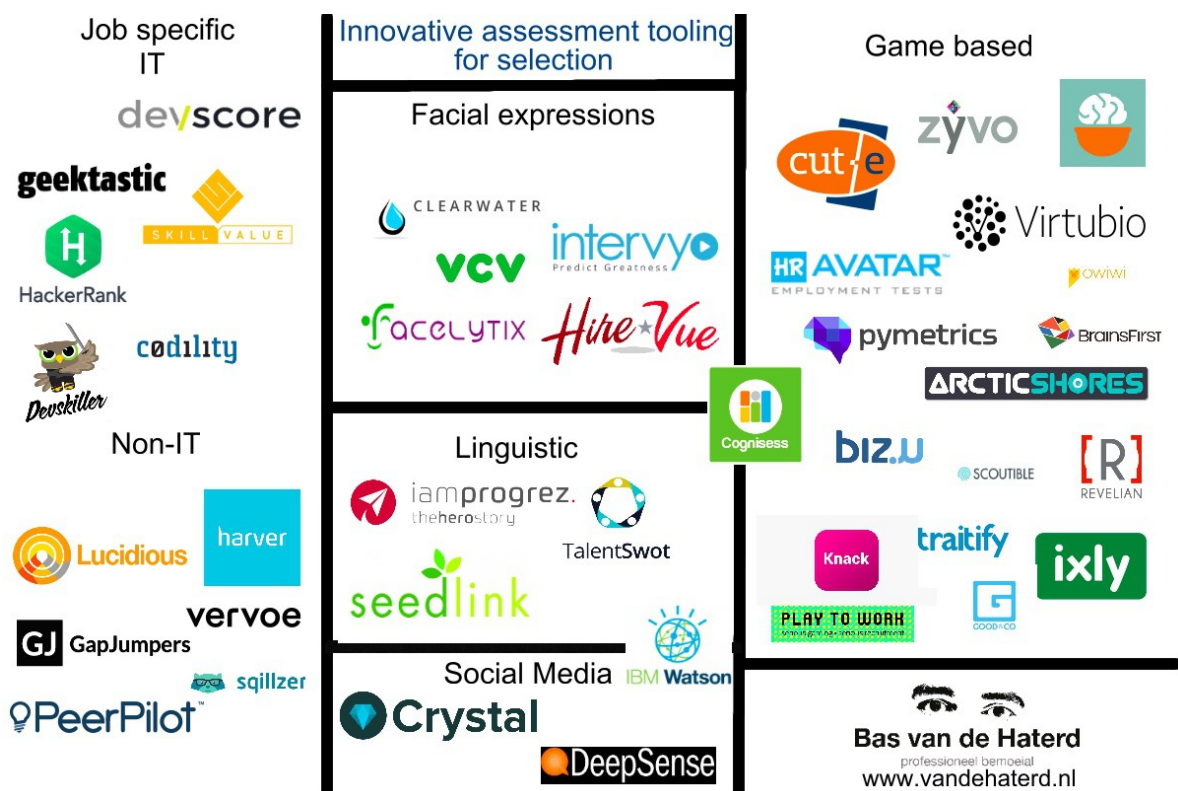
With great power comes great responsibility

Uncle Ben in Spiderman

Het probleem met zoveel kennis is dat we heel voorzichtig moeten zijn in hoe we met deze kennis omgaan. Sommige van deze tools zitten nog in hun kindertijd, andere zitten al in hun pubertijd, maar nog geen enkele is al geheel uitontwikkeld en volwassen. Meer onderzoek is nodig, maar het beste onderzoek is veldonderzoek. De technologieën die ik hieronder beschrijf zijn op de markt en doorgaans beter dan de huidige tools die we gebruiken om mensen te selecteren, namelijk het CV en het sollicitatiegesprek. Dat wil niet zeggen dat we deze tools blind moeten volgen. We moeten vragen stellen bij de uitkomsten, we moeten feedback geven aan de ontwikkelaars zodat ze de tools kunnen blijven verbeteren. De leveranciers, maar ook de maatschappij, moet blijven investeren in (fundamenteel) onderzoek in alle technologische gebieden die ik hieronder beschrijf. Het potentieel is gigantisch, als deze technologie goed wordt ontwikkeld en gebruikt kunnen we de meeste van onze vooroordelen voorkomen. Als we het verkeerd doen, worden onze vooroordelen in een algoritme vastgelegd dat nooit meer wordt aangepast.

Het onderdeel waar de research op dit moment het meest gebrekkig in is, zijn de kwaliteiten van de toppers in specifieke functies bij specifieke organisaties. Dit onderzoeken zal een van de grootste uitdagingen zijn in het optimaliseren van deze nieuwe technologieën.

Het landschap van innovatieve, niet vragenlijst gebaseerde, assessment tools zit er als volgt uit.



Verschillende technologieën zitten in verschillende stadia van ontwikkeling. En uiteraard zitten verschillende leveranciers binnen een categorie ook in verschillende stadia van ontwikkeling. Het is uiteraard niet zo dat elke leverancier even goed is.

In deze white paper ga ik niet in op de verschillende leveranciers en hun sterke en zwakke punten. Dit aangezien elke test die ik gedaan heb altijd maar een steekproef van N=1 heeft. Hoewel het ook mijn persoonlijke mening is, wil ik wel de technologische volwassenheid van de technologie als geheel rangschikken, waarbij ik kijk naar de meest ontwikkelde leverancier in combinatie met de wetenschappelijke basis waarop de technologie gebaseerd is.



Net geboren	Baby	Kindertijd	Pubertijd	Volwassen	Gepensioneerd	Voltooid leven
DNA	Social media	Linguïstiek	Games	SJG	Gamified vragenlijst	Vragenlijsten
Hersenscans		Gezichts-uitdrukkingen				

Vragenlijsten hebben wat mij betreft een mooi leven gehad, maar ze zijn inmiddels wel klaar om uit de handel genomen te worden als het gaat om het in kaart brengen van karaktereigenschappen. Ze zijn te makkelijk manipuleerbaar, en daar waar elke leverancier dit ontkent is er zeer veel bewijs dat dit wel mogelijk is. Nee, je kan jezelf niet de perfecte kandidaat maken. Ja, er zitten validatie vragen in. Echter weten kandidaten als geen ander hun zwakke plekken en met een vragenlijst gebaseerde assessment zijn die eenvoudig te verhullen. Ter illustratie een persoonlijk voorbeeld. Ik weet dat ik een confronterende stijl heb die soms problematisch kan zijn. Dat zal nooit uit een vragenlijst naar voren komen omdat ik weet hoe ik deze vragen kan manipuleren. Omdat iemand met een iets minder confronterende stijl geen reden heeft tot manipulatie zal die het wel eerlijk invullen, met als gevolg dat we gelijkwaardig uit de test komen op deze karaktereigenschap.

Uiteraard zijn vragenlijsten in bepaalde situaties wel nuttig, maar niet voor het bepalen van karaktereigenschappen. Er is bijvoorbeeld geen andere manier om iemands motivatie en ambitie te bepalen dan het te vragen. Ze kunnen ook handig zijn om iemands geprefereerde communicatiestijlen in kaart te brengen, hoewel op dit gebied andere tools ook kunnen werken.

Aan de andere kant van het spectrum liggen DNA gebaseerde assessments en hersenscans. Deze beide technologieën voelen waarschijnlijk als doodeng en dat is terecht. Beide worden nu een aantal jaren ontwikkeld en er is zelfs al één aanbieder van DNA gebaseerde assessments in Nederland actief. Dit is allemaal bijzonder nieuw en we weten nog maar heel erg weinig over wat DNA ons kan vertellen en hoe onze hersenen echt werken dat ik niet kan aanraden dat enige organisatie dit zou moeten gebruiken in de nabije toekomst. Dat wil niet zeggen dat deze technologie er niet is en dat de wetenschappelijke vooruitgang niet snel gaat. Het is goed om wel nu na te denken over hoe we als maatschappij hiermee om moeten gaan en duidelijke regels op moeten stellen m.b.t. het gebruik en zaken als privacy, veiligheid en het vernietigen van samples.

Inhoudsopgave

Voorwoord.....	2
Introductie.....	3
Waarom (digitale) assessments?	6
Game based assessments.....	8
Gamified vragenlijsten.....	8
Situational Judgement games.....	9
Psychometrische games.....	10
Eén spel of modulaair opgebouwd.....	11
Fit scores.....	11
Diversiteit.....	12
Gamers.....	13
Cognitieve games	14
Ontwikkeling & Sport.....	14
Samenvatting van alle game-based assessments.....	15
kandidaat-ervaring.....	15
Risico's van de technologie.....	15
Inzetbaarheid.....	15
Linguïstische Assessments.....	16
Kandidaat-ervaring.....	16
Risico's van de technologie.....	17
Meetbare eigenschappen	17
Inzetbaarheid	17
Social media assessments.....	18
Kandidaat-ervaring.....	18
Risico's van de technologie	19
Meetbare eigenschappen.....	19
Inzetbaarheid.....	19
Gezichtsuitdrukkingen.....	20
Kandidaat-ervaring.....	21
Risico's van de technologie.....	21
Meetbare eigenschappen.....	22
Inzetbaarheid.....	22
Hersenscans.....	23
Hoe werkt het?.....	23
De belofte.....	23
DNA.....	24
De belofte.....	24
Over de auteur.....	25

Waarom (digitale) assessments?

Voordat ik in ga op de verschillende nieuwe technologieën en al hun voor- en nadelen wil ik graag even het belang onderstrepen waarom we op basis van betere informatie mensen gaan selecteren.

- Meerdere onderzoeken hebben laten zien dat als vrouwen exact hetzelfde zeggen als mannen, ze gemiddeld als 20% minder competent worden gezien.
- Code van vrouwen op Github wordt 20% lager gewaardeerd als de screennaam duidelijk laat zien dat het een vrouw is.¹
- Nederlands onderzoek waarbij gesolliciteerd werd met een Nederlands klinkende naam en een moslim naam met dezelfde CV de Nederlandse naam in 32% van de gevallen werd uitgenodigd, de Moslim in 9%. Zelfs wanneer de Nederlandse naam in zijn brief vermeldde dat hij een veroordeling op zijn naam had staan, werd hij nog vaker uitgenodigd dan de Moslim.²
- In een onderzoek waarbij meer dan 400.000 CV's die op duizenden banen bij vele werkgevers hadden gesolliciteerd zijn geanalyseerd, bleek dat recruiters keken naar aantal jaren ervaring, meer dan het soort ervaring, maar alleen als een sollicitant onder een bepaalde leeftijd zat.³
- Amazon trainde een algoritme op basis van menselijk gedrag CV's te selecteren en het beoordeelde mannen structureel hoger dan vrouwen.⁴

Het probleem in al deze gevallen is tweeledig. Ten eerste gebruiken we iets voor onze eerste selectie waar geen enkel wetenschappelijk onderzoek ooit enige voorspellende waarde aan heeft kunnen toekennen: het CV. Een CV zegt namelijk wat je gedaan hebt, hoe lang en voor wie. De belangrijkste voorspeller van succes is echter hoe goed je dat werk gedaan hebt, gevolgd door de omstandigheden waaronder je die kwaliteit leverde. De tweede valkuil is onze menselijkheid. We zijn bevooroordeeld of we dat nu willen of niet.

Onderzoek⁵ geeft bijvoorbeeld aan dat we specifieke karaktereigenschappen toekennen aan uiterlijke kenmerken, terwijl er geen enkele correlatie is. Dit onderzoek kort samenvattend vinden we dikke mensen slordig en ongeorganiseerd, terwijl slanke mensen nieuwsgierig en vastberaden gevonden worden. Dit is de hoofdreden dat het vervangen van een geschreven CV door een video CV niet gaat helpen met het oplossen van vooroordelen. We vervangen enkel het ene vooroordeel, op basis van schrijfstijl en woorden, met een ander vooroordeel, op basis van uiterlijk.

Wanneer recruiters gevraagd worden wat een belangrijke reden is om een kandidaat meteen af te wijzen worden spelfouten vaak genoemd. Dat terwijl Duits wetenschappelijk⁶ onderzoek heeft aangetoond dat er geen enkele correlatie is tussen enige karaktereigenschap en spelfouten, niet eens nauwkeurigheid.

Tel hierbij op dat een sollicitatiegesprek behoorlijk nutteloos is,⁷ aangezien we alleen zoeken naar bevestiging wat we al denken gelezen te hebben in een cv, en het probleem van de selectie is duidelijk.

1 <https://www.theguardian.com/technology/2016/feb/12/women-considered-better-coders-hide-gender-github>

2 <https://www.nrc.nl/nieuws/2017/07/06/afkomst-weegt-bij-sollicitatie-zwaarder-dan-straftblad-11658392-a1565738>

3 <https://www.eur.nl/nieuws/big-data-voorspelt-wie-op-sollicitatiegesprek-mag-komen>

4 <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

5 <https://www.theatlantic.com/health/archive/2018/11/body-stereotypes-personality-debunked-eugenics/575041/>

6 https://www.ehochdreil-skillware.de/downloads/Zusammenfassung_Veranstaltung_Prof._Dr._Kanning.pdf

7 https://www.nytimes.com/2017/04/08/opinion/sunday/the-utter-uselessness-of-job-interviews.html?_r=0

Daarom moeten we naar betere pre selectie tools. Tools die ons een beter en onbevooroordeeld beeld geven van kandidaten. Alleen dan kunnen we de beste persoon voor de baan, zowel individueel als binnen het team, selecteren.

De reden om dit via digitale assessment tools te doen is een praktische: tijd en geld. Het is financieel onhaalbaar en verschrikkelijk qua kandidaat-ervaring om alle sollicitanten een dag te laten besteden bij een assessment bureau om een zo onbevooroordeeld mogelijk rapport van kwaliteiten en karaktereigenschappen te krijgen.

Game based assessments

De ontwikkeling van game based assessments is het verst van alle nieuwe technologieën. Dit soort assessments zijn gebaseerd op de grootste hoeveelheid wetenschappelijk onderzoek en er zijn al het meeste praktijk cases. Binnen de game based assessments zijn er verschillende subcategorieën mogelijk, waarbij de grenzen van elke categorie niet heel hard gedefinieerd zijn. Persoonlijk onderscheid ik vier categorieën:

- Gamified vragenlijsten
- Situational judgement games
- Psychometrische games
- Cognitieve games

Alle vier deze subcategorieën hebben hun eigen ontwikkelstadium. Vragenlijsten, ook gamified, hebben hun beste tijd gehad. Gegamificeerde vragenlijsten maken de ervaring leuker, maar de techniek niet beter. Situational Judgement Games (SJG) zijn een volwassen product. Game based assessments, zowel psychometrisch als cognitief, zitten beide in hun pubertijd. Behoorlijk ontwikkeld, goed te gebruiken, maar nog lang niet volledig ontwikkeld en stabiel.

Gamified vragenlijsten

Gamified vragenlijsten behoren wat mij betreft niet tot de echte game based assessments, maar aangezien ze als zodanig gezien worden door de markt heb ik ze in deze white paper opgenomen. Veelal zijn dit vragenlijsten met foto's, keuzes die je kan swipen en andere visuele aanpassingen. Soms hebben ze cognitieve spellen in de game opgenomen, zoals korte en lange termijn geheugen testen, multi tasking testen en meer kleine eenvoudige testen.

Vragenlijsten zijn de basis van veel assessments, hoewel hun wetenschappelijk validiteit meer en meer in twijfel getrokken is de afgelopen jaren. Bijna alle varianten, behalve de Big Five, zijn afgeschreven als niet wetenschappelijk en zelfs op de Big Five gebaseerde vragenlijsten hebben kritiek te verduren gekregen vanwege seksisme⁸.

Vragenlijsten hebben de mogelijkheid om gemanipuleerd te worden door kandidaten. Terwijl alle leveranciers claimen dat dit niet kan en ze hiervoor meten, blijkt uit onderzoek dat de gemiddelde kandidaat bijvoorbeeld veel analytischer en actie georiënteerder is dan de gemiddelde medewerker die zo'n test doet voor ontwikkeldoelinden. Om in een niet wetenschappelijk vorm te spreken, kandidaten zijn vaker rood en blauw.

Na ongeveer twee maanden, als een nieuwe medewerker zijn of haar proeftijd door is, komt ongeveer eenderde naar me toe en vragen ze me of ze de, vragenlijst gebaseerde, assessment over mogen doen. Aangezien we hun persoonlijke ontwikkelplannen baseren op de uitslagen en ze deze 'een beetje gestuurd hebben'.

- HR manager van een zorginstelling

Iets anders dat we niet moeten vergeten is dat vragenlijsten zelfassessments zijn. Zelfkennis is hiervoor van belang, maar tevens zal het beeld van een kandidaat afhankelijk zijn van zijn of haar omgeving.

Uiteraard zijn vragen wel geschikt om iemands ambitie, motivatie en drijfveren te bepalen. Belangrijke elementen voor toekomstig succes.

⁸ <https://qz.com/1201773/we-took-the-worlds-most-scientific-personality-test-and-discovered-unexpectedly-sexist-results/>

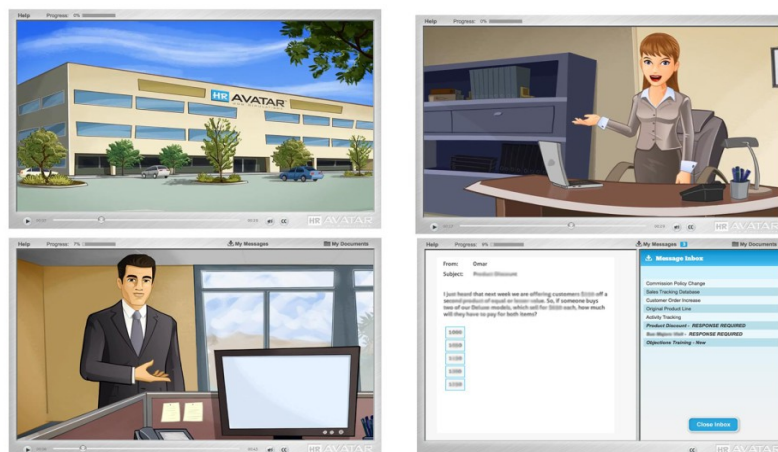
Situational Judgement games

Situational Judgement Games (SJG) zijn niet nieuw als assessment tool en hebben een lange geschiedenis als het om validiteit gaat. Hoewel technisch gezien een SJG lijkt op een vragenlijst is er een groot verschil. Bij een vragenlijst is het doorgaans heel duidelijk hoe vragen van invloed zijn op het karakter dat uit de test komt. Bij SJG's is dat veel minder duidelijk, omdat het echte situaties zijn waar beide antwoorden zowel goede als slechte elementen hebben. Hoewel het mogelijk is om algemene karaktereigenschappen uit een SJG te halen is de hoogste validiteit te behalen bij hele specifieke organisatie gerichte situaties. De beste antwoorden zijn dan afhankelijk van de bedrijfscultuur. Dit maakt SJG's vaak heel specifiek en niet te standaardiseren.

Verschillende Situational Judgement Games zijn gegamicifeerd, wat de ervaring van de kandidaat heeft verbeterd, net als de validiteit. Met een SJG zijn bijvoorbeeld kritieke incidenten na te bootsen die bij een organisatie plaats hebben gevonden om na te gaan hoe iemand in zo'n situatie zou handelen. Maakt die persoon een besluit of wacht hij of zij op de manager? Het juiste antwoord hierop is waarschijnlijk anders bij een start-up ten opzichte van een overheidsinstelling. Of wanneer je onder druk gezet wordt vanuit een manager, verandert een kandidaat zijn of haar mening. Het positieve van een SJG is het feit dat het nu mogelijk is om deze testen uit te voeren op een manier die je vroeger alleen in een fysieke omgeving van een assessment centrum kon doen. Je kan bijvoorbeeld kleine foutjes in informatie meegeven en kijken of een kandidaat die opvallend. Je kan een persoon heel boos laten kijken als die iets vraagt en kijken hoe een kandidaat daar op reageert.

Situational Judgement Games kunnen makkelijk cognitieve testen in de game inbouwen. Je kan eenvoudig het korte termijn geheugen meten, maar ook risico aversie. Dit kan bijvoorbeeld door de kandidaat informatie te geven, vijf minuten later hier iets over te vragen en te kijken of een kandidaat het goed onthouden heeft en meteen antwoord geeft of nog even terug kijkt naar de eerder verkregen informatie. Zoekt de kandidaat de informatie opnieuw op of wordt de vraag meteen beantwoord? En als de vraag meteen beantwoord wordt, was het antwoord goed of fout? SJG's zijn ook perfect voor het meten van planning en organisatie-vaardigheden, prioritering en werken onder druk.

Van alle game based assessments hebben SJG's de langste geschiedenis en de hoogste betrouwbaarheid als het gaat om historische accuratesse. Ze hebben een risico gemanipuleerd te worden aangezien je doorgaans wel door hebt wat er gemeten wordt, maar minder dan bij de standaard vragenlijst. Ze zijn vooral heel goed om te kijken hoe kandidaten om gaan met kritieke momenten die in een organisatie kunnen voorkomen in combinatie met het meten van verschillende cognitieve vaardigheden.



Job applicants complete an engaging simulation test, set in a work environment

Screenshots van HR Avatar

Psychometrische games

Psychometrische games meten je karakter gebaseerd op hoe je verschillende spellen speelt. Er is veel discussie over waar psychometrie eindigt en cognitief begint. De wetenschap heeft hier geen eenduidig antwoord op, dus hanteer een de volgende definitie:

Cognitieve kwaliteiten gaan over wat je kunt doen. Psychometrische kwaliteiten hebben betrekking op wat je daadwerkelijk doet. Als ik dus een hele hoge cognitieve score heb op nauwkeurigheid betekent dat dat ik heel precies en nauwkeurig kan zijn, maar niet dat ik dat altijd in alle omstandigheden ben. Persoonlijk ben ik bijvoorbeeld heel erg nauwkeurig als het gaat om usability en gebruikersgemak, maar ben ik verschrikkelijk slecht als het gaat om nauwkeurigheid in taal. Ik heb de cognitieve talenten om nauwkeurig te zijn, maar de omstandigheden bepalen of ik dat talent gebruik.

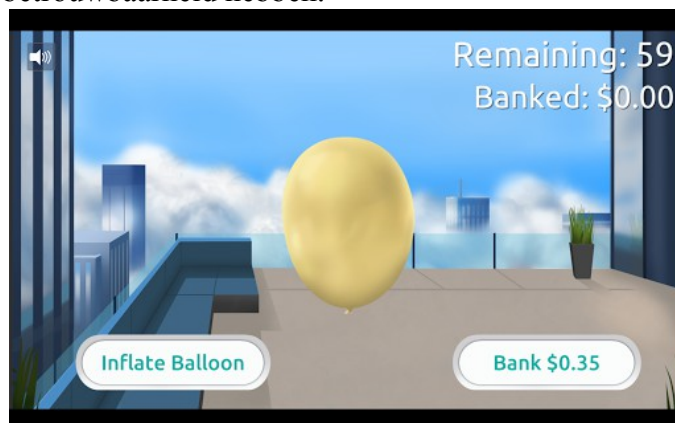
Een groot verschil tussen deze spellen is dat psychometrische spellen kijken naar je huidige staat en cognitieve assessments kijken naar het potentieel dat je hebt. Of een kandidaat ooit zijn volledige potentieel zal benutten is afhankelijk van veel factoren, van ambitie en motivatie tot de kansen die hij of zij krijgt de talenten te ontwikkelen.

Psychometrische games assessen ons karakter, doorgaans gebaseerd op de big five, maar er zijn mogelijkheden om één of zelfs twee stappen dieper te gaan. Psychometrische assessments kunnen skills beoordelen als GRIT (doorzettingsvermogen, omgaan met tegenslag), maar ook plannen en organiseren, risico aversie, zelfstandig denken, assertiviteit en veel, veel meer.

Er zijn veel games die veel verschillende talenten in kaart brengen en er worden er nog veel meer ontwikkeld. Een voorbeeld van het soort spellen dat je tegenkomt bij psychometrische games:

- Het opblazen van ballonnen die elk moment kunnen knappen;
- Gerechten serveren aan klanten met verschillende gezichtsuitdrukkingen;
- Angry birds-achtige schietspellen waar je je steeds moet aanpassen aan nieuwe elementen;
- Schieten op vallende elementen;
- Het openen van een kluis met de mogelijkheid opnieuw te beginnen zo vaak je wilt;
- Het vinden van een X in een veld van 20 bij 20 letters;
- Het beantwoorden van vragen of de letter dezelfde was als de vorige;
- Het draaien zodat een waterleiding water van A naar B brengt;
- Het links of rechts klikken afhankelijk van de stand of de kleur van de pijl.

Al deze spellen meten specifieke skills. Vaak kan een spel meerdere zaken meten, waardoor ze in combinatie een hogere betrouwbaarheid hebben.



Screenshot van Arctic Shores

Eén spel of modulair opgebouwd

Verschillende leveranciers hebben verschillende manieren om een psychometrische game-based assessment aan te vliegen. Sommige bieden modulaire spellen aan waar je kan aangeven welke zaken je wilt meten en de kandidaat krijgt alleen die games. Dit scheelt veel tijd voor de kandidaat en biedt een betere kandidaat-ervaring. Andere aanbieders hebben één spel die altijd alles meet en dus is de tijdsinvestering van de kandidaat altijd gelijk. Beide varianten hebben voor- en nadelen. Bij het vaste spel is het mogelijk om elke eigenschap meervoudig te meten. Elk level in een spel meet andere zaken, maar het levert ook data ter controle van de eerder gemeten eigenschappen. Dit maakt deze meetmethode robuuster en de uitkomsten betrouwbaarder. De negatieve kant is de mindere kandidaat-ervaring aangezien de tijdsinvestering voor elke sollicitant veel langer is. Voor modulaire spellen is het omgekeerde waar. Ze hebben een betere kandidaat-ervaring, maar zijn minder betrouwbaar. Minder is nadrukkelijk niet onbetrouwbaar. Het is het beste te vergelijken met een auto voorbeeld. Een Volvo V60 is volgens recente crash tests veel betrouwbaarder dan een Ford Tourneo, maar er is geen reden om je daar niet veilig in te voelen.

Fit scores

Een van de mogelijkheden die elke game-based assessment kan bieden is een fit of match score. Door je huidige medewerkers ook de games te laten spelen en aan te geven wie goed en wie minder presteren is het mogelijk om na te gaan welke kwaliteiten daadwerkelijk het verschil maken tussen goed en matig presteren in die bepaalde functie bij jouw specifieke organisatie. Er zijn hier echter heel veel valkuilen die ontweken moeten worden.

Het eerste wat we ons moeten realiseren is dat er nog maar heel weinig onderzoek is gedaan naar de vraag welke talenten nu echt het verschil maken in welke banen, laat staan deze banen binnen jouw specifieke organisatie. Als we kijken naar de kenmerken die we momenteel vragen in vacatures en die door traditionele assessments bekeken worden komen deze vaak voort uit theoretische exercities. Een groep wijze mannen (of vrouwen), vaak met wat input van een manager, hebben een competitie-model bedacht vol met karaktereigenschappen. Dat is doorgaans niet de beste manier om zo'n model te ontwikkelen.

We weten ook dat bedrijfscultuur van grote invloed is op iemands functioneren. We kennen allemaal de verhalen van de geweldige verkoper die in zijn nieuwe baan niets verkocht of de geniale ontwikkelaar die in een nieuwe functie niets voor elkaar kreeg. Een baan is meer dan een combinatie van taken, banen verschillen van bedrijf tot bedrijf.

De beste manier om een fit score vast te leggen is door je huidige werknemers de assessment te laten doen en deze werknemers in te delen in bijvoorbeeld drie categorieën: geweldig, acceptabel en matig en een algoritme te laten bepalen welke eigenschappen het verschil maakt tussen deze drie. Het probleem hiermee is dat je voldoende mensen nodig hebt die dezelfde baan hebben, minimaal 50. Aangezien er weinig organisaties zijn waar 50 mensen precies dezelfde baan hebben is er vaak de neiging om te gaan groeperen en dat is een enorme fout. Laten we verkoop als voorbeeld nemen. Je hoeft geen sales expert te zijn om te begrijpen dat een telefonische verkoopondersteuner hele andere kwaliteiten moet hebben dan een account manager die in de auto zit en klanten bezoekt. Het groeperen van mensen om voldoende steekproef te krijgen is de meest gemaakte fout.

Sommige organisaties bekijken alleen de toppers, de mensen die ze willen klonen. Hoewel dit niet altijd fout is, geeft dit wel een heel hoog risico op een zogenaamde 'false positive'. Om dit wederom met een voorbeeld uit de verkoop te omschrijven. De kans is groot dat alle toppers hoog scoren op extraversie. Aangezien de kans groot is dat alle verkopers hierop geselecteerd zijn met als gevolg dat de slechte verkopers hier waarschijnlijk ook hoog op scoren. Door alleen toppers te testen weet je niet of en hoeveel invloed dit heeft.

Een hele belangrijke stap die nooit overgeslagen mag worden wanneer er fit scores gemaakt worden is de menselijke beoordeling van de criteria gebruikt door het algoritme. Voor elke functie moeten mensen beoordelen of er enige logica dat een bepaalde karaktereigenschap meegewogen wordt bij de bepaling van een fit score.

Ondanks de risico's die verbonden zijn aan het ontwikkelen van een fit score is dit wel wat deze tools zo ongelooflijk krachtig maken. En laten we niet vergeten, ook al is de score misschien niet beter, de kans dat hij beter is dan de huidige selectie op basis van een cv is wel heel groot.

Diversiteit

Er zijn heel veel zorgen over diversiteit als het gaat om het gebruik van assessments en algoritmes. Door onze beste medewerkers te klonen zou die omlaag gaan. Toch geven alle organisaties aan waar met dit soort tooling gewerkt wordt dat er een verhoging is van diversiteit. Hierbij wordt diversiteit gedefinieerd als etniciteit, geslacht en soms leeftijd. De reden hiervoor is eenvoudig, in plaats van dat onze onbewuste vooroordelen karaktereigenschappen toekennen aan mensen op basis van hun geslacht, etniciteit en leeftijd gaan we het nu daadwerkelijk meten. We zullen niet meer onbewust een 60 jarige niet uitnodigen voor een product development rol omdat deze niet innovatief is, we zien het innovatie potentieel uit de test. We denken niet meer dat een vrouw niet goed kan onderhandelen en daarmee ongeschikt is als verkoper, we meten haar onderhandelvaardigheden. We leggen niet meer het CV van de moslim op de afwijs stapel voor een administratieve baan omdat we bang zijn dat hij misschien gaat frauderen, we meten zijn betrouwbaarheid.

Door het meten van de daadwerkelijke eigenschappen die van belang zijn, in plaats van de attributie van eigenschappen op basis van iemands werkervaring of achtergrond kunnen organisaties digitale assessment tools inzetten om hun diversiteit te vergroten.

Er is echter nog één belangrijk vraag die nog beantwoord moet worden, aangezien deze technologie nog niet zo lang op de markt is. Wat betekent dit op de lange termijn voor 'diversity of thought'. De diversiteit in gedachten en eigenschappen? Meestal neemt deze ook toe als de diversiteit in achtergronden toeneemt, maar nu selecteren we misschien andere achtergronden, maar wel met dezelfde karaktereigenschappen. Een belangrijke vraag om in de gaten te houden.

Een andere onbeantwoorde vraag heeft betrekking op de verhouding tussen team en individuele prestaties. Geen enkel voetbalteam zal met 11 aanvallers een wedstrijd winnen, hoe goed ze ook zijn. Daarom is het zo belangrijk om op de juiste, specifieke talenten te selecteren per functie.



Gamers

Een veel gestelde vraag is doorgaans of gamers een voordeel hebben boven niet gamers bij game based assessments. Het antwoord is waarschijnlijk wel, maar niet omdat het beide betrekking heeft op games. De wetenschap heeft aangetoond dat gaming je hersenen traint. Zo wordt gaming ingezet om de degeneratie van de hersenen te vertragen bij bijvoorbeeld Parkinson patiënten. Onderzoek heeft ook aangetoond dat gamende chirurgen minder fouten maken door betere oog-hand coördinatie. Gamen is een training voor je hersenen zoals sit-ups voor je buikspieren en hardlopen voor je uithoudingsvermogen. Hierbij is het soort games dus ook van invloed op het soort cognitieve kwaliteiten die je traint. Dus als gamers een voordeel hebben is het omdat hun hersenen ook net iets sneller of beter werken op die specifieke taken die ze met hun favoriete spellen trainen.



Cognitieve games

Zoals ik al eerder schreef is er een hele dunne scheidslijn tussen cognitieve en psychometrische games. Daarom zal ik in dit hoofdstuk niet herhalen wat ik al heb geschreven over fit scores, diversiteit en het mogelijke voordeel dat gamers hebben. Dat geldt allemaal net zo goed voor cognitieve games als voor psychometrische games.

Veel psychometrische assessments komen voort uit de cognitieve wetenschap. Ze bouwen door op dezelfde onderzoeken die neurowetenschappers hebben gedaan over de werking van de hersenen. Het belangrijkste verschil dat ik zie tussen psychometrische testen en cognitieve testen is volgens mijn definitie het verschil tussen huidig gedrag en potentieel gedrag. Cognitieve assessments testen je potentieel, je sterke en zwakke punten en je natuurlijke voorkeuren. Terwijl psychometrische testen meer gaan over je huidige gedrag. Een cognitieve test kijkt naar de bouwstenen van je hersenen. Dat je bepaalde bouwstenen hebt, betekent dat je onder de juiste omstandigheden bepaald gedrag zal vertonen. Cognitieve testen meten het potentieel van een persoon en hun gedragingen onder de juiste omstandigheden. Aangezien de huidige omstandigheden mogelijk niet optimaal zijn zegt het niets over hun huidige gedrag. Aangezien mensen geconditioneerd zijn door hun huidige omstandigheden kan dit tot gevolg hebben dat het enige tijd duurt voordat iemand het gedrag zal vertonen dat op basis van de cognitieve testen logisch lijkt. Het afleren van bestaand gedrag is immers niet eenvoudig.

Betekent dit dat cognitieve assessments beter zijn dan psychometrische? Dat ligt er maar helemaal aan wat je wilt meten. Als je doel is om te kijken naar een fit van de persoon met de functie zal een psychometrische test waarschijnlijk een beter resultaat geven. Het is ook makkelijker om te matchen met de huidige populatie. Als je doel is om te kijken naar de mogelijkheden voor ontwikkeling in de toekomst is een cognitieve test waarschijnlijk iets beter, wat niet wil zeggen dat er geen mogelijkheden zijn om ook met een psychometrische game hier iets over te zeggen. Open staan voor nieuwe ervaringen, één van de Big Five karaktereigenschappen die met psychometrische testen kan worden gemeten, heeft duidelijke invloed op het lerend potentieel. Ook betekent het niet dat je geen onmiddellijke fit kan meten met cognitieve games. Het betekent alleen dat als je ze naast elkaar zet, de kracht van een psychometrische game is om te matchen met de huidige baan en de kracht van een cognitieve game op toekomstige ontwikkelingen ligt.

Ontwikkeling & Sport

Een van de gebieden waar cognitieve games momenteel veel worden toegepast is in de sport, met name in het voetbal. Door de cognitieve skills van jeugdspelers te bepalen is men in staat om de toekomstige 'slimme spelers' te selecteren en op de juiste positie op het veld te zetten. Een speler met bijvoorbeeld een lage inhibitie (zelfbeheersing) en een lage score op het loslaten van aandacht is een verschrikkelijke verdediger, aangezien hij bij balverlies niet snel kan omschakelen in een andere rol en vervolgens de tegenstander onderuit schopt. Een speler met een anticipatie gecombineerd met een hoge handelingssnelheid is een goede spelverdeler, aangezien hij in kleine ruimtes een paar stappen vooruit kan denken. Dit is natuurlijk puur de mentale kwaliteit van een speler, de fysieke kwaliteiten worden met deze spellen niet meegenomen.

De scores zijn uitmuntende voorspellers van toekomstig succes, sinds vooral op hele jonge leeftijd de fysieke mogelijkheden van kinderen nogal uiteen kunnen lopen, zelfs bij kinderen van dezelfde leeftijd. Daarom is enkel prestatie op het veld vaak een hele matige voorspeller. Vroeg rijpe kinderen zijn sterker, maar kunnen makkelijk worden ingelopen door laat rijpe kinderen, die vaak op jonge leeftijd al geen kansen meer krijgen.

Deze testen bieden dus ook veel mogelijkheden op het gebied van persoonlijke ontwikkeling. Cognitieve game-based assessments kunnen gebruikt worden om te kijken of iemand die mentale bouwstenen heeft om echt goed te worden in de baan waar deze persoon voor getraind wil worden. Je moet uiteraard wel weten welke kwaliteiten belangrijk zijn om goed te zijn in een baan. Laten we het voorbeeld nemen van een onderhandelaar. Die moet snel veel informatie kunnen verwerken en besluiten kunnen nemen, net als de spelverdelers bijvoorbeeld.

Deze testen werken ook erg goed voor mensen met weinig zelfkennis of zelfwaarde, zoals mensen die al jaren in de bijstand zitten. Deze testen bieden de mogelijkheid om hun potentieel in kaart te brengen, waarbij ze getraind kunnen worden op hun sterke punten. Deze testen kunnen tevens gebruikt worden om deze doelgroep zelf alsmede potentiële werkgevers te overtuigen dat ze wel over bepaalde talenten beschikken.

Cognitieve game-based assessments geven inzicht in iemands potentiële talent. Met een nog beter assessment die ook kijkt naar iemands huidige gedrag kun je ze bijvoorbeeld goed combineren met Situational Judgement Games.

Samenvatting van alle game-based assessments

In aanvulling op alle eerder genoemde informatie kan ik niet genoeg benadrukken dat er grote verschillen zijn tussen de verschillende leveranciers binnen de verschillende deelgebieden van game-based assessments. Zowel in kwaliteit als qua kosten. Hoewel er weinig partijen zijn die ik persoonlijk zou afraden heeft elke leverancier zijn eigen voor- en nadelen.

kandidaat-ervaring

De kandidaat-ervaring van game-based technologie is in grote mate afhankelijk van de leverancier. Sommige zijn beter dan andere. De tijdsinvestering die gevraagd wordt van de kandidaat is doorgaans hoger dan bij andere vormen van assessments en gelijk aan die van vragenlijsten. Een van de redenen hiervoor is mij verteld door wetenschappers die deze games ontwikkelen is dat deze tijdsinvestering nodig is om zaken als uithoudingsvermogen en doorzettingsvermogen te testen. Spellen korter dan 35 tot 45 minuten zijn minder betrouwbaar in het testen van doorzettingsvermogen en gedurende langere tijd gefocussed blijven.

Ondanks de relatief grote tijdsinvestering die een kandidaat in een game-based assessment moet stoppen is de ervaring doorgaans een stuk beter dan het invullen van vragenlijsten. Of dit zo blijft naarmate deze technologie meer gemeengoed wordt, is natuurlijk de vraag.

Risico's van de technologie

Van alle nieuwe innovatieve assessment tooling zijn game-based assessments het meest volwassen. Het grootste risico zit in de 'one-sizes-fits-all' aanpak, waarbij teveel functies op een hoop gegooid worden. Het grootste gevaar zit in het bepalen van de talenten die iemand nodig heeft om goed te zijn in een specifieke functie bij een specifieke organisatie, waarbij het aantal mensen dat nodig is voor een valide steekproef hoger is dan het aantal medewerkers dat die organisatie in de functie heeft. Het risico bij deze technologie zit niet in de technologie zelf, maar in de toepassing ervan.

Inzetbaarheid

Deze technologie kan ingezet worden voor alle soorten banen op elk niveau. Het is bijzonder belangrijk om te bepalen welke kwaliteiten je wilt assessen en welke spellen daar het beste bij passen. Er is echter geen (soort) functie waar het met deze technologie onmogelijk voor is om een goed beeld van te voren.

Linguïstische Assessments

De wetenschappelijke basis voor linguïstische assessments is ongeveer 70 jaar geleden gelegd toen in Amerika soldaten die posttraumatische stress stoornissen hadden geholpen werden via schrijfworkshops. Hoewel 70 jaar misschien lang lijkt, zijn de grote doorbraken pas de afgelopen jaren gemaakt. De eerste linguïstische assessment technologie ging uit van een vaste woordenlijst. Bepaalde woorden hadden bepaalde scores in bepaalde categorieën. Hoewel het systeem werkte werden er wel serieuze vragen gesteld bij de validiteit van het gebruikte vocabulaire. Immers was dit vocabulaire met de scores vastgesteld door een aantal professoren en promovendi op basis van hun eigen meningen en gevoel. De huidige gebruikte methodologie behelst zogenaamde open vocabulaires, waarbij algoritmes vaststellen welke woorden meer en minder gebruikt worden door mensen die via (zelf) assessments bepaalde karaktereigenschappen hebben.

De werking van linguïstische assessments is redelijk eenvoudig. Een sollicitant schrijft een stuk tekst. Sommige aanbieders vragen om generieke teksten zoals 'schrijf een stuk over jezelf'. Andere partijen stellen specifieke vragen zoals 'geeft eens een voorbeeld waarom je een doorzetter bent' of 'wat is je beste verkoopprestatie'. Er zijn onderzoekers die verklaren dat vragen naar specifieke eigenschappen je karakter rondom die eigenschap beter in kaart zou brengen, dit is wetenschappelijk (nog) niet algemeen geaccepteerd als de waarheid. De technologie kijkt niet inhoudelijk naar het antwoord, dat kunnen recruiters eventueel wel natuurlijk, maar analyseren het op basis van de woorden die je gebruikt, de lengte van de zinnen, de interpunctie en nog een aantal zaken. Vanuit die informatie kan een persoonlijkheidsprofiel gebouwd worden.

Kandidaat-ervaring

De kandidaat-ervaring met linguïstische assessments is uitstekend. Het enige dat kandidaten hoeven te doen is een klein stukje tekst te schrijven of een aantal vragen te beantwoorden. Vragen die relevant zijn voor de baan. Het kan zelfs ingezet worden om social media content te analyseren, maar meer hierover in het hoofdstuk social media assessments. In vergelijking met game-based assessments en vragenlijsten is de tijd die benodigd is voor een linguïstische assessment zeer beperkt.

De kandidaat-ervaring wordt enkel beter wanneer linguïstische assessments worden gecombineerd met andere, reeds geaccepteerde, technologie zoals chatbots of video-interviews. Kandidaten zijn inmiddels bekend met een kort video-interview, waarbij de gesproken tekst omgezet kan worden in geschreven tekst en geanalyseerd kan worden via linguïstische analyses.

Ook chatbots zijn steeds normaler in het wervingsproces en bieden ook een goede sollicitatie-ervaring. Een chatbot kan bijvoorbeeld dienen als een vervangen van de sollicitatiebrief. Bij bepaalde functies, variërend van stages tot verpleegkundigen en van vakkenvullers tot monteurs, speelt het CV een steeds minder belangrijke rol. Het gaat vaak op een paar functionele vragen zoals het wel of niet hebben van een diploma of certificaat gecombineerd met motivatie en karakter. Deze vragen kunnen uitstekend via een chatbot gesteld worden, waarbij de tekst geanalyseerd kan worden en een karakterprofiel kan geven. Op deze manier kan voor grote groepen met een goede kandidaat-ervaring bijvoorbeeld ook mobiel solliciteren eenvoudig mogelijk gemaakt worden.

Een grote vraag rondom deze technologie is wel of de kandidaat de uitkomsten zal accepteren, wanneer deze weet dat niet alleen het antwoord inhoudelijk wordt geëvalueerd, maar ook de linguïstiek over hoe het is geschreven. De onbekendheid met deze methodologie staat brede acceptatie op dit moment nog in de weg.

Risico's van de technologie

Linguïstische assessments hebben belangrijke risico's die moeten worden opgelost. Mijn persoonlijke ervaring is dat deze technologie problemen heeft met professionele schrijvers. Ik heb vijf boeken, vele white papers en duizenden artikelen geschreven en daarmee een schrijfstijl ontwikkeld die niet geheel meer past bij mijn karakter. Ik heb geleerd gestructureerder te schrijven, iets dat mij niet natuurlijk aankomt waaien, wat ook terug te lezen is in mijn eerste artikelen. Ook heb ik geleerd mezelf minder centraal te zetten. Linguïstische assessments waarderen mijn 'organisatietalent' systematisch te hoog en maken me introverter dan ik ben. Hierbij kijk ik uiteraard ook naar de vergelijking met alle andere assessmentvormen. Onder professionele schrijvers vallen niet alleen auteurs of redacteurs; wetenschappers en beleidsmedewerkers bij de overheid hebben bijvoorbeeld ook een schrijfstijl aangenomen die doorgaans niet natuurlijk is voor hen.

Mogelijk dat deze zwakte is op te lossen door social media posts te gebruiken, met een chatbot of een video-interview. Immers gebruiken mensen in conversaties, zoals met een chatbot of via een video-interview, doorgaans wel hun meest natuurlijke communicatievorm. Mijn persoonlijke ervaringen is dat linguïstische analyse via een chatbot en ook via video interviews een beter beeld geeft van mijn karakter dan via de analyse van een artikel of enkele vragen die in essay vorm beantwoord moeten worden.

Een ander belangrijk risico van de technologie is dat deze gebruikt wordt met te weinig woorden om op te analyseren. Het minimum voor een betrouwbare analyse wordt gesteld op 500 woorden.

Meetbare eigenschappen

De meeste linguïstische assessments rapporteren op de Big Five karaktereigenschappen. Dit betekent dat je een zeer goed gedocumenteerde, maar heel generiek, beeld van een persoon krijgt.

Andere tools gaan wel een aantal stappen verder in de eigenschappen waarop ze rapporteren, met gemengde resultaten. Eigenschappen als voorstellingsvermogen, autoriteit-uitdagend, assertiviteit, doelgerichtheid, zelfbewustheid, altruïsme en bescheidenheid worden door bepaalde tools ook op gerapporteerd.

Er zijn leveranciers die beweren dat de betrouwbaarheid toeneemt door mensen specifieke vragen over specifieke eigenschappen te stellen. Een vraag als 'omschrijf hoe je functioneert in een team' zou dus een betrouwbaardere analyse geven van je karaktereigenschappen rondom teamwork. De wetenschappelijke onderbouwing hiervoor is flinterdun, maar het is niet onmogelijk.

Een bijzonder prettige bijkomstigheid van linguïstische assessments is dat je ook de antwoorden op de vragen hebt. Dus naast een analyse van de karaktereigenschappen op basis van de gebruikte woorden kun je ook de eigenlijke antwoorden gebruiken en kijken of dit de antwoorden zijn die bij de organisatie passen.

Inzetbaarheid

In de huidige fase van de ontwikkeling zou ik terughoudend zijn om linguïstische assessments in te zetten voor iets anders dan volume recruitment. Er is eenvoudigweg nog niet genoeg onderzoek naar de vraag hoe verschillende groepen die veel schrijven, van wetenschappers tot beleidsmedewerkers, journalisten, redacteurs en bloggers hun schrijfstijl aanpassen vanwege hun baan. Voor volume recruitment, zoals bijvoorbeeld verplegers, caissières, vakkenvullers en heftruckchauffeurs en is deze tool waarschijnlijk uitermate geschikt. Het geeft een inzicht in iemands karaktereigenschappen en mogelijke fit met de organisatie. De tool zou waarschijnlijk in de zorg en het onderwijs van hele grote waarde kunnen zijn om te bepalen of zij-instromers passen bij het werk dat ze ambiëren.

Social media assessments

Social media assessments zijn voornamelijk gebaseerd op de eerder beschreven linguïstische assessments, maar met aanpassingen voor gedrag op social media. In veel gevallen kijken de tools naar Facebook Likes of accounts die je volgt op Twitter. Op basis van deze informatie word je profiel aangepast.

Zoals aangegeven in de paragraaf over linguïstische assessments is het taalgebruik op social media natuurlijker en zou het in theorie een deel van de problemen die linguïstische analyses hebben met professionele schrijvers moeten kunnen ondervangen. In mijn ervaring is dit helaas vooral theorie. De meeste tools waar ik het genoeg van heb gehad te mogen testen faalden faliekant. De tools waar likes en interacties werden gebruikt om linguïstische analyses te verbeteren gaven juist veel slechtere resultaten. De enige tool die ik heb getest die een behoorlijk goed resultaat gaf, kijkt voor zover ik heb kunnen ontleden primair naar iemands LinkedIn profiel. Hoewel ik niet met zekerheid kan zeggen dat ze geen linguïstische analyse maken van e-mails bij mensen die de e-mail plugin gebruiken. Deze tool is een black box en geeft geen openheid van zaken waar ze naar kijken, dus ik kan alleen gokken hoe het werkt op basis van mijn ervaring en de resultaten. Ondanks, of misschien dankzij, hun geslotenheid is het wel de enige tool in deze categorie die goed werkt.



Interessant bij het testen van deze tools was dat analyses van mijn Twitter profiel hele andere karaktereigenschappen weergeven dan die van mijn Facebook profiel. Een analyse op basis van Twitter maakt mij veel competitiever, daar waar ik op basis van mijn Facebook karakter een rustig en overdenkend persoon ben. Misschien zijn beide overigens wel gelijk, afhankelijk van de situatie.

Kandidaat-ervaring

De kandidaat-ervaring is zowel fantastisch als verschrikkelijk. Het is fantastisch omdat er geen enkele andere tool is die je kan analyseren zonder dat je er iets voor hoeft te doen. Eén klik geeft de tool toegang tot je profiel en klaar. Het is de minste tijdsinvestering van alle tools die je een kandidaat kunt bieden.

Het probleem is dat je de tool toegang geeft tot je social media, je privé social media. Als je een Facebook profiel hebt, zoals ik, dat echt persoonlijk is, is dat een hele stap. Daarom communiceer ik ook verschillend op Facebook en Twitter, het is is echt privé, het ander meer zakelijk. Dat geeft mogelijk ook de reden voor de verschillende karakters. Een organisatie, zelfs al is het een onafhankelijk algoritme, toegang geven tot mijn privé profiel voelt als een inbreuk op mijn privacy. Dus in theorie is de kandidaat-ervaring geweldig, in de praktijk voelt het voor velen heel vervelend.

Risico's van de technologie

Het belangrijkste risico in dit geval is de mens. Een paar tools gaven inzicht in het 'waarom' van bepaalde karakterbepalingen. Het was duidelijk dat mensen hier bepaalde eigenschappen hebben gekoppeld aan specifieke Facebook likes. De assumptie dat Linkin Park fans minder gelukkig zijn met hun leven bleek één van de redenen om mij verkeerd te classificeren.

Het grote probleem met social media analyses is de 'one size fits all' analyses van gedrag. Ik heb bijvoorbeeld de pagina's van Shell en een paar autofabrikanten ooit geliked vanwege hun recruitment acties die ik als recruitment expert graag wilde zien. Op basis hiervan classificeren social media assessments me als een autoliefhebber. Ik ben alles behalve dat.

De technologie heeft zeker veel potentieel, maar is duidelijk nog in een bijzonder vroeg ontwikkelstadium. Het grootste risico is dat onze menselijke assumpties en vooroordelen in deze technologie worden gebouwd.

Meetbare eigenschappen

De beloftes zijn enorm. De huidige resultaten ondermaats. De huidige technologie is redelijk in mensen op basis van de Big Five, of DISC, in te delen. Sommige tools zijn goed in het toevoegen van specifieke karaktereigenschappen of communicatiestijlen. Bijvoorbeeld wie je het beste kunt bellen en wie je beter kunt mailen. En waar je de nadruk op moet leggen om iemands aandacht te trekken. Moet je gemeenschappelijke connecties benadrukken of de reden waarom je contact zoekt? Is iemand toekomst georiënteerd of zoekt die persoon wetenschappelijke onderbouwing?

Inzetbaarheid

In de huidige fase van ontwikkeling zou ik deze technologie nog niet inzetten. Er zitten nog teveel risico's aan en de uitkomsten zijn nog te onbetrouwbaar. Gezien het potentieel van deze technologie is het wel enorm belangrijk om de ontwikkelingen in de gaten te houden. Met de snelheid van de technologische ontwikkelingen kan binnen afzienbare tijd deze technologie wel klaar voor de markt zijn.

Gezichtsuitdrukkingen

Gezichtsuitdrukkingen zijn een zeer controversieel onderwerp binnen de assessment industrie. Elk jaar komen er wel enkele wetenschappers die artikelen publiceren⁹ met waarschuwingen tegen het gebruik van deze technologie, met name de micro-expressies. Er is geen peer reviewed artikel te vinden over dit onderwerp is hun standpunt en dat klopt. In elke waarschuwing wordt de relatie gelegd met de Nazi-wetenschap van fysionomie, waarmee de Duitsers dachten Joden te kunnen identificeren aan o.a. de omvang van hun hoofd. Met de waarschuwing dat we nooit moeten toestaan deze pseudowetenschap terug te laten komen. Dat is ook de reden dat elk artikel bij voorbaat geweigerd wordt door alle gerenommeerde wetenschappelijke tijdschriften. Elke redactie is bang voor de risico's van associatie met nazi-wetenschap, ondanks dat de psychometrische assessments op basis van micro-expressies niets te maken hebben met de nazi-wetenschap van weleer.

Uiteraard moeten we ons heel erg bewust zijn van de risico's die deze technieken met zich meebrengen. De technologie kan gebruikt worden om karaktertrekken te bepalen, niet de fouten in iemands karakter. Er is een groot verschil tussen iemand die status gedreven is, weinig respect heeft voor hiërarchie en een kort lontje heeft en iemand die een crimineel is omdat dergelijke karaktereigenschappen veel voorkomen bij criminelen. Daar waar karaktereigenschappen in kaart gebracht kunnen worden is wat een persoon met die eigenschappen doet nog altijd een vrije keuze. Persoonlijk zou ik er meer moeite mee hebben om seksuele voorkeuren te bepalen op basis van gezichtsuitdrukkingen op foto's, maar daar had de academische wereld geen moeite mee gezien dit onderzoek¹⁰.

We moeten ook heel voorzichtig zijn om te kunnen bepalen of iemand liegt. Dit is bijzonder moeilijk vast te stellen omdat je hiervoor een 'baseline' nodig hebt. En hoe weet je zeker dat iemand eerlijk is als je deze persoon digitaal interviewt? Hoewel zelfs zonder baseline is een algoritme waarschijnlijk beter in het bepalen dan bijna elk mens of een persoon liegt of gewoon heel erg nerveus is, is het nog geen exacte wetenschap en derhalve moeten we dit nog niet proberen.

Het gebruik van micro-expressies om karakters te bepalen is al best ver ontwikkeld, ondanks het feit dat de wetenschappelijke basis bijzonder beperkt is. Mijn persoonlijke ervaringen met de technologie zijn bijzonder goed. Het is erg kandidaatvriendelijk en de technologie heeft mogelijkheden om eigenschappen vast te stellen die veel andere assessment technologieën niet hebben. Uiteraard heeft het ook minder positieve kanten, bepaalde zaken die bijvoorbeeld game based assessments kunnen vaststellen zijn weer niet met micro-expressies vast te stellen.

Dit hoofdstuk heet echter niet micro-expressies, maar gezichtsuitdrukkingen. De reden hiervoor is omdat de ontwikkeling verder gaat dan alleen micro-expressies. In 2018 is er een wetenschappelijke studie uitgekomen¹¹ die vast heeft gesteld dat oogbewegingen tijdens normale taken in staat zijn om vier van de Big Five karaktereigenschappen vast te stellen. Aangezien de huidige micro-expressies juist iets minder zijn op de Big Five zou dit een welkome aanvulling kunnen zijn op deze technologie.

Persoonlijk denk ik dat gezichtsuitdrukkingen nog in de kinderschoenen staan qua ontwikkeling, net als linguïstische analyses. In bepaalde gevallen zijn ze te gebruiken en van grote waarde, maar ik zou ze nog niet adviseren om het hele selectieproces op te baseren.

9 <https://theintercept.com/2018/12/06/artificial-intelligence-experts-issue-urgent-warning-against-facial-scanning-with-a-dangerous-history/>

10 <https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph>

11 <https://www.frontiersin.org/articles/10.3389/fnhum.2018.00105/full>

Kandidaat-ervaring

De kandidaat-ervaring is geweldig. Afhankelijk van de leverancier moet een kandidaat of tussen de 3 en 6 specifieke vragen beantwoorden in 30 seconde per vraag, of meer generieke vragen van een minuut per vraag of gewoon drie minuten over zichzelf praten. De tijdsinvestering is dus erg kort en in alle gevallen is de video op te nemen met je mobiele telefoon of webcam.

Het enige nadeel m.b.t. de kandidaat-ervaring ervaring is dat met video zaken als etniciteit, geslacht, leeftijd en fysiek voorkomen niet te verbergen zijn. Aangezien dit allemaal zaken zijn waar recruiters onbewuste vooroordelen over kunnen hebben zullen sommige mensen het solliciteren via video als een negatieve kandidaat-ervaring zien.

Risico's van de technologie

De technologie heeft drie zwakke plekken. De technologie is onbetrouwbaar als de kandidaat:

- alcohol gedronken heeft;
- drugs gebruikt heeft;
- een geestesziekte heeft.

Hoewel de assessment technologie geen betrekking heeft op leugendetectie, maar karakterbepalingen, is deze laatste het eenvoudigste uit te leggen aan de hand van het voorbeeld van een pathologische leugenaar. Ik bedoel hiermee iemand met de echte geestesziekte, niet iemand die gewoon heel veel liegt. Het probleem met een pathologische leugenaar is dat deze persoon niet weet dat hij of zij liegt, met als gevolg dat het onbewuste geen signalen geeft dat deze persoon liegt, dus is het niet te meten. Hetzelfde geldt bijvoorbeeld voor iemand met een narcistische persoonlijkheidsstoornis. Ik bedoel dan dus niet de gemiddelde narcist, maar iemand met de aandoening. Aangezien het aantal mensen met dergelijke aandoeningen heel beperkt is, kan dit risico genegeerd worden.

Een ander risico bij deze technologie heeft betrekking op huidskleur. Alle gezichtsherkenningstechnologie heeft meer moeite met het correct herkennen van donkere mensen. Zowel bij identificatie, maar ook bij assessment technologie. De oorzaak hiervoor is eenvoudig, hoe donkerder een oppervlak, des te minder contrast er te zien is, aangezien op een zwarte achtergrond geen schaduw te zien is. De resultaten van mensen met een donkere huidskleur zijn derhalve minder betrouwbaar.

Het grootste risico van deze technologie is echter dat wetenschappers erg terughoudend zijn om te overwegen dat ons onbewuste een hele goede bron kan zijn van karaktereigenschappen. Daarom worden alle onderzoeken die wel gedaan worden onmiddellijk gedemoniseerd in plaats serieus geanalyseerd met als gevolg dat er het risico van twijfelachtige leveranciers sterk toeneemt.

Meetbare eigenschappen

Het unieke van micro-expressies assessments zijn de eigenschappen die je ermee in kaart kan brengen. Een van de testen die ik heb mogen ervaren schaaft mensen in op drie types: actie georiënteerd, mens georiënteerd en analytisch. Dit zijn drie zeer nuttige algemene classificaties aangezien je alle drie de types nodig hebt voor een goed functionerend team. Iemand moet de 'waarom' vraag stellen, dit is de mens georiënteerde persoon. Iemand moet de rationale, de business case, de meetbare succesfactoren willen vaststellen, de analytische persoon. En dan moet er ook nog iemand de stap zetten om daadwerkelijk iets te gaan doen, de actie-gerichte persoon. Het is uiteraard niet zo dat iemand slechts één van deze drie eigenschappen heeft, wel is er bijna altijd één dominant. Vanuit diversiteitsoogpunt is het bijzonder handig om in een team altijd per eigenschap minimaal één persoon te hebben bij wie die eigenschap dominant is.

Echter kan je met micro-expressies een stap dieper gaan nog op de communicatiestijlen van personen. Mijn assessment gaf bijvoorbeeld aan dat ik onvoorspelbaar kan overkomen op sommige mensen en heel direct ben in mijn communicatiestijl. Mijn potentiële zwakte, gezien mijn hoge score op actie-gerichtheid, is te snel handelen. De kracht van micro-expressies is het in kaart brengen van deze bijzonder handige eigenschappen, met name gericht op communicatiestijlen en -voorkeuren.

De huidige technologie is nog niet heel erg goed om mensen in te delen op de Big Five, maar met de ontwikkelingen op het gebied van oogbewegingen die eerder besproken zijn zou dat de komende jaren wel eens opgelost kunnen worden. Ook kan de gesproken tekst middels linguïstische assessments worden geanalyseerd, wat tot een betrouwbare Big Five analyse kan leiden.

Inzetbaarheid

Ik zou de inzet van micro-expressie gebaseerde assessments met name inzetten voor team optimalisatie. De selectie van complementaire persoonlijkheden binnen een afdeling in plaats het selecteren van mensen die op ons lijken. Dit kan op elk niveau, van raden van bestuur tot net afgestudeerden.

Voor trainee programma's zou ik adviseren om te kijken welke karakter eigenschappen het beste presteren in het programma, maar ook hier om de diversiteit in karakters in de gaten te houden. Als trainees de toekomstige managers van je organisatie zijn is het belangrijk om mentale diversiteit bij de instroom goed te managen.

Tot slot zou ik micro-expressie assessments inzetten bij onverklaarbare twijfel na een gesprek. Er zijn gevallen waar een kandidaat gewoon ongeschikt is, maar in ander gevallen heeft een kandidaat 'een allergie' geraakt bij de eerste communicatie met de recruiter of hiring manager. Wanneer zowel de recruiter of de hiring manager als de kandidaat een micro-expressie assessment doet, kun je bepalen welke mogelijke communicatieve allergieën tussen beide zouden kunnen zitten. Als de onverklaarbare twijfel voortkomt uit een dergelijke communicatieve allergie zijn deze vaak op te lossen door ze bespreekbaar te maken.

Hersenscans

Hersenscans klinken waarschijnlijk erg futuristisch en dat zijn ze ook. Hersenscans voor assessments zijn nog niet op de markt. De wet verbiedt het ook om ernaar te vragen. Dat wil niet zeggen dat we dit ook in de toekomst moeten uitsluiten.



Hoe werkt het?

Wanneer iemand onder een hersenscan machine, doorgaans een MRI, ligt krijgt deze persoon verschillende stimuli te zien. Afhankelijk van het deel van de hersenen dat activiteit vertoont en oplicht op de scan kunnen karaktereigenschappen heel specifiek en nauwkeurig worden vastgelegd. Er is geen enkele manier om vals te spelen. Dat gezegd hebbende wil een bepaalde reactie van de hersenen op een bepaalde stimulans niet zeggen dat die persoon ook altijd dat gedrag zal vertonen.

Een voorbeeld. De wetenschap heeft aangetoond dat pedoseksualiteit een aangeboren seksuele voorkeur is, net als homoseksualiteit en heteroseksualiteit. Het feit echter dat iemand is aangetrokken tot hele jongen kinderen wil niet zeggen dat deze persoon er ook naar zal handelen. Sterker, de meeste doen dat niet omdat hun morele gevoelens dat niet toestaan.

Een tweede manier waarop hersenscans kunnen werken is door de hersenen te testen tijdens bijvoorbeeld een game-based assessment. Het is mogelijk om de activiteit in de hersenen te testen bij het behalen van bepaalde scores, waardoor het groeipotentieel kan worden vastgesteld. Zoals al beschreven in het hoofdstuk over game-based assessments, gamers trainen hun hersenen met games, net als je spieren traint met fysieke training. Hersenen zijn, net als spieren, trainbaar. Beide hebben echter ook genetische limieten die van persoon tot persoon verschillen.

De belofte

De belofte van hersenscans is enorm. Om door te gaan om het hierboven gegeven voorbeeld, veel pedoseksuelen geven niet toe aan hun verlangens. Het kan echter verstandig zijn om ze echter niet teveel in de verleiding te brengen en ze bijvoorbeeld geen baan in de kinderopvang te geven.

Een ander voorbeeld, één die al eens onderzocht is met behulp van hersenscans, heeft te maken met fraude. Alle grote en aanhoudende fraudeurs, niet de gevallen van de eenmalige gelegenheidsfraudeurs, zagen met bepaalde stimuli bepaalde hersendelen oplichten. Aangezien deze mensen blijkbaar eenvoudig verleid kunnen worden tot fraude is het misschien mensen met deze eigenschappen geen CFO te maken.

Voor topverkopers is er ook een bepaald deel van de hersenen dat oplicht onder bepaalde stimuli, het zou verstandig kunnen zijn deze mensen dus juist aan te nemen als verkoper.

DNA

DNA zijn de bouwstenen van ons leven en langzaam kraken we de code van het leven. Net als bij hersenscans verbiedt de wet ons om DNA te vragen aan kandidaten of medewerkers. Dat wil niet zeggen dat er niet al aanbieders zijn die DNA gebaseerde ontwikkel assessments aanbieden. Sterker, in Nederland zijn er al aanbieders die met een beetje DNA van je baby je al indicaties geven welke eigenschappen deze gaat krijgen. Zolang een individu het vraagt en de resultaten krijgt mag het van de wet.



Hoewel we langzaam meer en meer leren over onze DNA, DNA markers en hoe we DNA kunnen manipuleren en aanpassen, is er nog veel meer dat we niet weten dan dat we wel weten. Het feit dat mijn DNA erfelijkheidstest recentelijk mijn voorgeslacht aanpaste van 8% Brits naar 50% Brits vanwege 'een verbeterde steekproef' geeft wel aan hoe betrouwbaar deze testen op dit moment zijn.

De belofte

De belofte is enorm. Alles dat we zijn, of kunnen zijn, zit in ons DNA. Er is geen wetenschappelijke consensus hoeveel van wat we zijn nature (DNA) is en hoeveel nurture (opvoeding). Sommige wetenschappers claimen dat 80% van wie we zijn genetisch bepaald is, andere claimen dat onze genen slechts 20% van ons uiteindelijke succes bepalen. Ook zijn er theorieën die stellen dat je genen voor een deel ook je opvoeding bepalen. De nurture of nature noemt Robert Plomin¹² dat. Als kind zoek je de omgeving op die bij je natuur past. Aangezien er geen wetenschappelijke consensus is houd ik het persoonlijk op de veilige 50%-50%.

Als we het over DNA hebben moeten we het ook over CRISPR hebben. Met deze technologie kan je je DNA aanpassen. Ik heb het dan niet over het implementeren van nieuwe genen, maar we kunnen slapende genen actief maken of actieve genen uitzetten. Het gaat dus niet over het implementeren van een stukje DNA van een reptiel of een beer in de mens, maar om de aan/uit schakelaar van een gen om te zetten. Dit kan bijvoorbeeld heel handig zijn bij een nierpatiënt die niet voldoende EPO aanmaakt omdat het gen, dat daar voor moet zorgen, nooit aan is gegaan bij die persoon. Dezelfde techniek kan echter ook bij een sporter worden toegepast die hetzelfde gen in de hoogste versnelling zet. Deze techniek leidde in november 2018 tot grote opschudding toen een Chinese wetenschapper bekend maakte middels CRISPR techniek de eerste HIV resistente baby's geboren waren¹³. Als middels CRISPR DNA aangepast kan worden en we kennen de genen die bijvoorbeeld lees en leerprestaties beïnvloeden, liggen hier grote kansen en risico's voor talentontwikkeling.

DNA analyses zijn de eerste stap in het genetisch potentieel weten. De tweede is er naar handelen door bijvoorbeeld bij kinderen al op vroege leeftijd te stimuleren op zaken waar ze goed in kunnen zijn. De derde stap is het aanpassen en een betere versie van jezelf maken. Deze laatste twee stappen zijn in mijn optiek doodeng.

Hoewel DNA assessments al op de markt zijn, staat het nog aan het hele beginstadium van de ontwikkeling. Ze zullen ongetwijfeld de komende decennia enorm verbeteren en iets serieus worden om rekening mee te houden.

¹² Robert Plomin – Blueprint, how DNA makes us who we are

¹³ <https://www.volkskrant.nl/wetenschap/claim-chinese-hoogleraar-eerste-genetisch-gemanipuleerde-baby-s-ter-wereld-bf669f83/>

Over de auteur

Bas van de Haterd is professioneel bemoeial. Hij is een HR technologie adept en gelooft dat we met moderne technologie de wereld veel mooier, gelijkjer en eerlijker kunnen maken. Dat moderne technologie ons kan helpen om mensen hun talenten optimaal te benutten en discriminatie op de arbeidsmarkt bijvoorbeeld op te lossen.

Echter, als we als maatschappij geen hele duidelijke afspraken maken en een goede visie ontwikkelen zullen films als 1984 en V for Vendetta paradijselijk lijken.



Bas heeft de afgelopen jaren vele tientallen moderne assessment tools mogen testen en uiteraard enkele traditionele assessments gedaan. Vanuit die positie is hij in staat op de kwaliteit van de assessments goed in te schatten. Hij heeft in deze paper zijn bevindingen van de tools beschreven.

Bas is een veel gevraagd (internationaal) spreker over het verbeteren van recruitment en selectie.



Matěj Matolín

Head of HR, Trainer,
Speaker, Blogger

31 oktober 2018, Matěj heeft
met Bas samengewerkt, maar
voor verschillende bedrijven

I had a chance to see Bas's presentation at TA Live about the essentials of a career site. It was an amazing talk full of specific hints, figures, methods, and examples. It was one of the most useful presentations I've attended so far. Later on, I met him and Social Recruiting Days and his talk about modern assessment systems was again top class. I wish there are more speakers like him.



Hung Lee

Co-founder & CEO at
Workshape.io - Talent
Matching for Tech ♦
@HungLee

31 oktober 2018, Hung heeft
met Bas samengewerkt, maar
voor verschillende bedrijven

Bas gave a superb talk at Social Recruiting Days in Berlin
#SRD18.

He has a unique style and method of delivery - challenging, engaging, entertaining and insightful. Would recommend Bas if you want someone to deliver a high value presentation on tech / work / people.

Hij organiseert ook twee events per jaar over innovatie in recruitment.

- Het internationale event Talent Acquisition Live (www.ta-live.com)
- Het Nederlandse event Digitaal-Werven (<http://www.digitaal-werven.nl/event>).

Contact kan via de website: <http://www.vandehaterd.nl> of e-mail: bas@vandehaterd.nl