

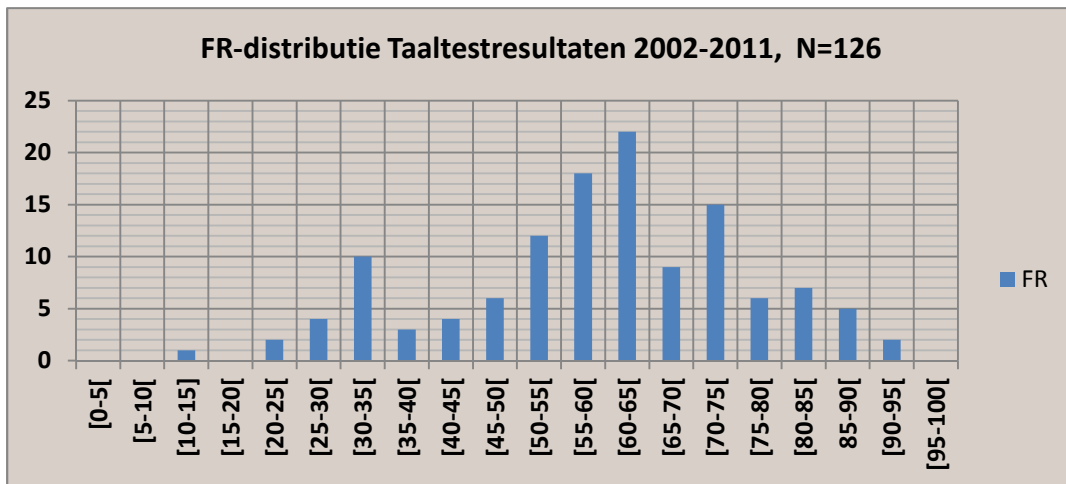
### 1. Inleiding: Er bestaat een positieve correlatie tussen de schrijftaalttest en schoolresultaten

In het verleden werden steeds opvallende positieve correlaties vastgesteld tussen schoolresultaten en schrijftaalttestresultaten. Soms waren deze correlaties significant en soms niet. Bij vroeger onderzoek naar spreekvaardigheid en leesvaardigheid vonden we zelden eenduidige positieve correlaties tussen schoolresultaten. Meestal waren deze correlaties onbeduidend laag of zelfs negatief en zelden positief.

***De schrijftaalttest blijkt een redelijke betrouwbare indicator te zijn voor potentieel schoolsucces, mits we rekening houden met de grote individuele verschillen die we vaststellen onder onze jongeren qua inzet en andere competenties.***

Op het eerste gezicht lijkt dat misschien verwonderlijk. De schrijftaalttest meet immers eerder grammaticale taalbeheersing dat tot uiting komt in een dictee, spelling, vervoegingen van werkwoorden in de tegenwoordige en verleden tijd, meervoud, zinnen bouwen, alfabetisch rangschikken en een bondige beschrijving geven van een strip. De taalttest is een standaard schrijftest die al vanaf 2002 tot nu 2011 wordt gebruikt in het Keerpunt en waar onderzoek naar de spreiding en het gemiddelde + bijhorend schoolniveau goed gekend is. We kunnen daarom zeer goed de taalscores vergelijken en beoordelen met de schoolresultaten.

klassemidden	% '02- '11	FR	RFR	niveau
2	[0-5[	0	0,00%	
7	[5-10[	0	0,00%	
12	[10-15[	1	0,79%	5,55%
17	[15-20[	0	0,00%	
22	[20-25[	2	1,59%	
27	[25-30[	4	3,17%	
32	[30-35[	10	7,94%	
37	[35-40[	3	2,38%	10,31%
42	[40-45[	4	3,17%	
47	[45-50[	6	4,76%	
52	[50-55[	12	9,52%	
57	[55-60[	18	14,29%	48,41%
62	[60-65[	22	17,46%	
67	[65-70[	9	7,14%	
72	[70-75[	15	11,91%	
77	[75-80[	6	4,76%	15,88%
82	[80-85[	7	5,56%	
87	[85-90[	5	3,97%	
92	[90-95[	2	1,59%	
97	[95-100[	0	0,00%	
	totaal	126	100,00%	
gemiddelde + 95% T-interv		58,75	(55,7 - 61,7)	
	min	22		
	Q1	52		
	Med	62		
	Q3	72		
	max	92		



*Men zou eerder verwachten dat leesvaardigheid en spreekvaardigheid een grotere impact heeft op schoolresultaten, zoals dat vaak wordt beweerd bij PISA-onderzoek. Maar indien we de schrijftaaltest gaan koppelen aan eindtermen of competenties zoals: de essentie kunnen weergeven van een tekst, zich correct kunnen uitdrukken, een verslag of brief kunnen maken, kunnen rapporteren op het werk, enz..., dan blijkt snel waarom er toch eerder een positieve correlatie bestaat tussen de taaltestresultaten en de competenties die onze jongeren van verzorging voldoende moeten kunnen beheersen om met succes hun beroep te kunnen uitoefenen.*

**Opmerking!**

*Een positieve correlatie vaststellen betekent niet altijd dat er een doorslaggevend en significante samenhang moet bestaan. Deze samenhang wordt immers ook beïnvloed door de taalvaardigheidverschillen in een klas. In een klas waar de spreiding tussen taalvaardige leerlingen gemiddeld groter is, zal sneller een samenhang worden vastgesteld tussen taalverschillen en verschillen in schoolresultaten. Een klas die uitsluitend uit sterke taalvaardige leerlingen bestaat of uitsluitend uit zwakkere leerlingen, zal minder significante correlaties opleveren om de eenvoudige redenen dat de taalverschillen klein zijn en dan minder relevant doorwegen op schoolresultaten. In zulke klassen zullen verschillen tussen jongeren qua inzet, motivatie, doorzetting, inzicht eerder de schoolresultaten beïnvloeden dan wat de taaltest meet.*

*Voor het onderzoek werden de gemiddelde testresultaten voor PAV vergeleken met de taaltestresultaten.*

**2. Wat stellen de gemiddelde PAV testcores voor in dit onderzoek?**

De gemiddelde PAV testresultaten omvatten het gemiddelde schoolresultaat bekomen bij zowel open boektesten, gesloten boektesten, moduletesten LBC (periode september tot en met april) De vragen in deze testen omvatten kennisvragen, inzichtvragen, een eigen mening geven, verbanden leggen, PC-vaardigheden toepassen, het kunnen opzoeken en verwerken van informatie, het kunnen toepassen van kennis en theorie met voorbeelden. De vragen bestaan uit zowel open vragen, gesloten vragen als meerkeuzevragen.

Het is immers de bedoeling in PAV om elke jongere, ongeacht verschillen in vaardigheden, de kans te geven op verscheidene manieren te tonen in hoever ze de modules en de thema's hebben verwerkt en te testen in hoever de jongeren de eindtermen in voldoende mate beheersen. Omdat een klas zowel bestaat uit zwakkere jongeren als sterke jongeren en we niet de bedoeling hebben jongeren te selecteren maar aan te moedigen en vooruitgang te maken, kunnen zelfs zwakkere jongeren die zich voldoende hebben ingezet en de leerstof en vaardigheden afdoend beheersen, nog behoorlijk scoren. Tegelijk kunnen sterke leerlingen tonen dat ze meer kunnen en zijn er voldoende moeilijke en uitdagende vragen voor hen voorzien. Op die manier worden jongeren die zich sterk inzetten beloond met een sterk resultaat, maar worden jongeren die zwakker zijn niet geconfronteerd met deprimerende scores. Jongeren die echter onvoldoende de leerstof en vaardigheden beheersen zullen niet slagen op de testen. Om te voorkomen dat jongeren weinig kans hebben op slagen voor PAV-testen of moduletesten, krijgen alle jongeren de kans in de klas deze testen voor te bereiden. Wie de testen

goed voorbereid heeft goede kansen op slagen. Wie echter weinig moeite doet, zal weinig kans hebben op slagen of het risico lopen magere scores te behalen.

### 3. Welke twee klassen werden onderzocht en vergeleken?

In dit onderzoek hebben we twee verschillende klassen van verzorging gekozen die een vergelijkbaar PAV themapakket hebben gekregen en vrijwel identieke PAV -testen. **Er zijn enkel verschillen tussen beide klassen wat de moduletesten betreft. Er zijn ook verschillen tussen beide klassen qua instapniveau.**

VZ2A bestaat uit eerstejaars die de richting verzorging volgen. In deze klas zitten zowel jongeren op het niveau tweede graad als 2<sup>de</sup> jaar derde graad.

VZ3B bestaat uit tweedejaars verzorging. In deze klas komt 1 meisje in aanmerking voor het middelbaar diploma. Alle andere meisjes komen in aanmerking om een getuigschrift 2<sup>de</sup> jaar derde graad te behalen. **De verwachting is dan ook dat de meisjes van VZ3B gemiddeld hoger zullen scoren voor PAV en taal dan VZ2A.** Onderaan in de tabel zijn de resultaten weergegeven van de jongeren van VZ2A/ VZ3B die aan alle PAV -testen hebben deelgenomen tussen september 2010 en april 2011 + de taalttest. **De gemiddelde PAV- score bedraagt voor VZ2A: 70,5% en voor de taalttest: 60,4%. De gemiddelde PAV- score bedraagt voor VZ3B: 81,3% en voor de taalttest: 70,8%. De meisjes van VZ3B scoren dus zoals verwacht ook hoger op de beide testen.** Zij hebben dan ook duidelijk al vooruitgang geboekt tegenover de eerstejaars.

Naam	klas	PAV % gemiddeld	Taal %	instapniveau <sup>1</sup>
Ayan	VZ2A	68,8	57	BUSO /BSO/DBSO – 2 <sup>de</sup> graad
Hanan	VZ2A	71,5	74	5 BSO + DBSO
Louise	VZ2A	75	62	Dislexie / 5 BSO + DBSO
Jasmina	VZ2A	55	53	BUSO + DBSO – 2 <sup>de</sup> graad
Kimberley	VZ2A	82	56	5 BUSO + DBSO
Emel	VZ3B	93	82	4 ASO / 5 TSO + DBSO
Donika	VZ3B	72	56	OKAN + DBSO
Pauline	VZ3B	80,5	82,5	4 TSO + DBSO – diploma MO
Engedi	VZ3B	75,5	59	4 BSO + DBSO
Stefanie	VZ3B	85	75,5	5 BSO + DBSO
Yasmina	VZ3B	82	70	4 BSO + DBSO
<b>testperiode</b>	<b>sept'10-apr'11</b>			
	<b>gemiddelde</b>	<b>76,4</b>	<b>66,1</b>	

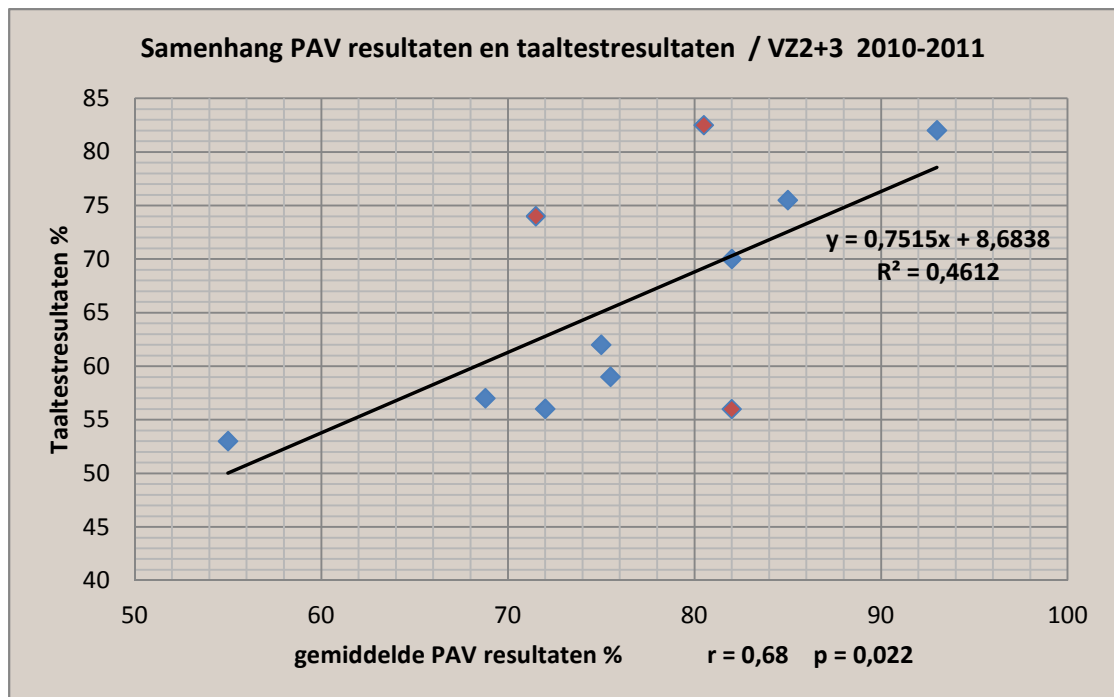
<b>ANOVA</b>	<b>F = 5,29</b>	<b>p = 0,0323</b>
	<b>R<sup>2</sup> = 0,2093</b>	

<sup>1</sup> In de tabel vind je ook het instapniveau van de jongere voor ze in het CL&W Het keerpunt kwamen. **De meeste jongeren komen administratief in aanmerking voor een getuigschrift 2<sup>de</sup> jaar 3<sup>de</sup> graad mits voldaan aan de eindtermen op het einde van het schooljaar. Twee jongeren kunnen een getuigschrift 2<sup>de</sup> graad behalen en 1 jongere komt in aanmerking voor het diploma middelbaar onderwijs, mits de voorwaarde voldaan te hebben aan de eindtermen voor het behalen van het diploma op het einde van het schooljaar.**

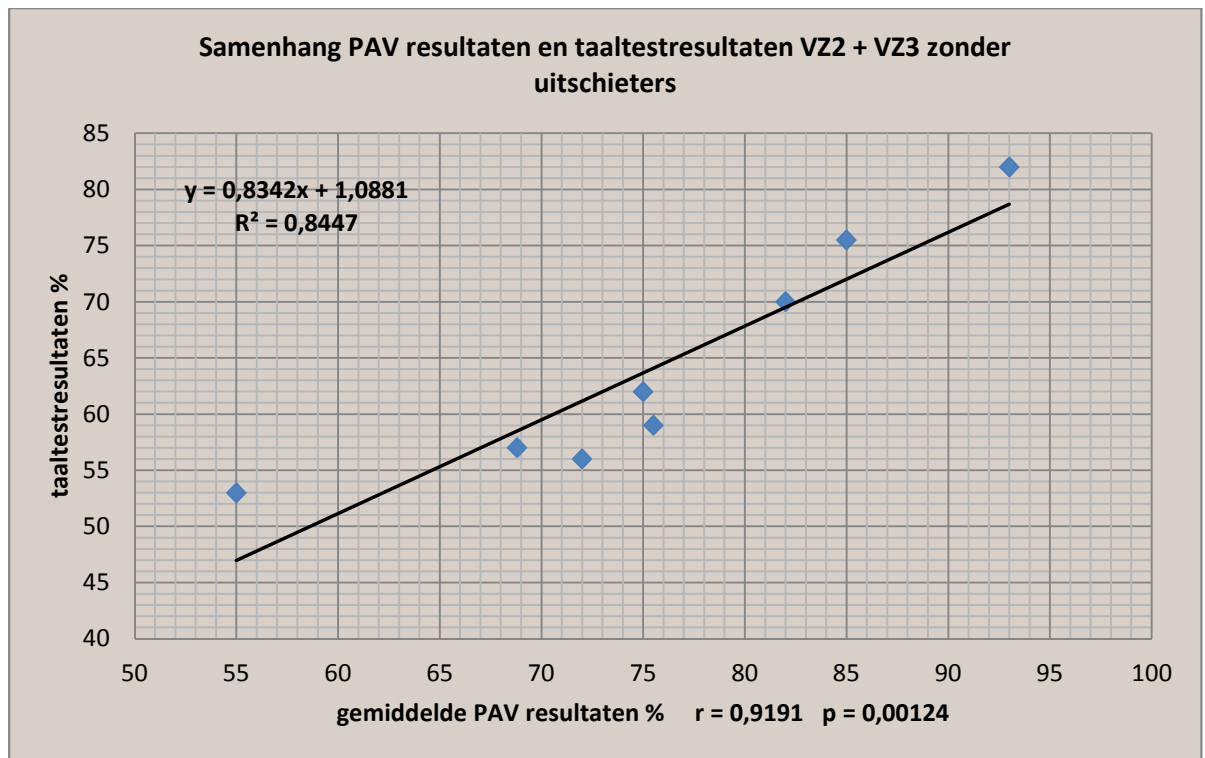
#### 4. Wat kunnen we te weten komen via een ANOVA analyse en correlatieonderzoek tussen de beide testscores?

Uit de ANOVA- analyse blijkt dat de gemiddelde scores van beide testresultaten significant van elkaar verschillen. Via de determinatiecoëfficiënt  $R^2$  blijkt dat circa 21 % van de variatie kan toegeschreven worden aan verschillen in testscores tussen de taaltest en het PAV- testgemiddelde. Concreet betekent dat dan circa 79 % van de variatie van de scores kunnen toegeschreven worden aan verschillen tussen de jongeren wat betreft hun individuele taalvaardigheid, inzetverschillen, motivatieverschillen en het beheersen van de geteste vaardigheden in de PAV -testen.

Er is geen sprake van een doorslaggevende overlapping tussen de taaltestscores en PAV -scores, zodat we mogen besluiten dat de beide testen wel degelijk voldoende verschillen van elkaar zowel qua opzet en doel als de score. Via correlatieonderzoek kunnen we nu onderzoeken in hoever er een samenhang bestaat tussen de gemiddelde PAV- testscore en de taaltestscore. Indien we een significante samenhang vaststellen, dan betekent dat taalvaardigheden tot een bepaald niveau samenhangen met de PAV- schoolresultaten en met elkaar verband houden. Dat verband kan wederzijds zijn. Een jongere die gemiddeld hoger scoort op de taaltest, zal dan meer kans hebben ook hoger te scoren op de PAV -testen. Omgekeerd geldt dat jongeren die doorgaans goede PAV -resultaten behalen, ook meer kans zal hebben hoger te scoren op de taaltest.



**Uit het spreidingsdiagram en correlatieonderzoek blijkt dat er wel degelijk een significante samenhang bestaat tussen de taaltestscores en de PAV -testscores.** De samenhang is van die aard dat circa 46 % van de PAV -resultaten een verband houden met de taaltestresultaten en omgekeerd. 54 % van de resultaten op de testen hangen eerder samen met andere factoren of toeval. **Interessant bij correlatieonderzoek zijn de uitschieters die de correlatiecoëfficiënt sterk doen dalen. Deze zijn rood gemarkeerd op het spreidingsdiagram en werden groen gemarkeerd in de tabel onder 3.** Op de volgende bladzijde vind je een spreidingsdiagram waarbij we de rode uitschieters hebben weggelaten om beter te beoordelen in hoever ze de samenhang tussen taaltestscores en PAV -scores beïnvloeden. Hanan, Kimberley en Pauline wijken sterk af van de verwachte samenhang. We kunnen dan ook de afwijking van de regressielijn verklaren door te onderzoeken waarin deze drie leerlingen verschillen tegenover de overige acht onderzochte leerlingen.



*Indien we de drie afwijkende leerlingen weglaten, dan valt op hoe sterk het verband is tussen de schoolresultaten en de taalttest. Opnieuw stellen we vast dat de taalttest, mits we rekening houden met onverwachte verschillen en afwijkingen onder sommige jongeren, een voldoende betrouwbaar meetinstrument blijkt te zijn om potentieel schoolsucces te voorspellen. Er bestaat nu een samenhang voor 85% tussen taaltestscores en PAV -testscores in plaats van 46%. Slechts 15 % van de samenhang hangt samen met toeval of andere factoren die zorgen voor afwijking van de regressielijn.*

##### 5. Waarin verschillen de 3 uitschieters van de overige onderzochte jongeren?

Het zou nu onverstandig zijn te besluiten dat er voldoende argumenten bestaan om te beweren dat we via de taalttest adequaat, mogelijk schoolsucces kunnen voorspellen. Daarom onderzoeken we nu waarom de resultaten van de drie jongeren: Hanan, Kimberley en Pauline, zo sterk afwijken van de overige onderzochte jongeren. **De beschrijving is gebaseerd op notities tijdens de TBO's in de loop van het schooljaar.**

**Hanan** is een potentieel sterke leerling die de beste taalttestresultaten behaalde van haar klas (VZ2A). Toch scoort ze op haar PAV -testen gemiddeld lager dan verwacht. Hanan heeft in voltijds onderwijs tot een 5<sup>de</sup> jaar BSO met succes gevolgd en nadien het DBSO gedurende een jaar. Haar taalttestresultaten komen daarmee mooi overeen met het verwachte taalniveau van 70% voor een leerling BSO 3<sup>de</sup> graad (zie tabel pagina 1 → donkergroene balk). Wie doorgaans hoger scoort dan 75% op de taalttest, heeft meestal een niveau 3<sup>de</sup> graad BSO, TSO. Wie hoger scoort dan 80% of 85% op de taalttest heeft doorgaans een niveau 3<sup>de</sup> graad TSO, ASO. Hanan kan vlot met de PC teksten verwerken, maar is wat luier in het voorbereiden van haar testen. Ze weet dat ze goed kan volgen en is daarom wat nonchalanter. Toch doet ze doorgaans haar best, maar ze is niet zo prestatiegericht dat ze het uiterste wil behalen. Ze is tevreden met een voldoende tot behoorlijk resultaat. Deze ingesteldheid kan voor een deel verklaren waarom ze wat lager scoort dan verwacht op de PAV -testen. Hanan is daarentegen wel de beste en de snelste om een opdracht zelfstandig af te werken op PC. Ze heeft oog voor structuur, kan goede zinnen formuleren, goed de essentie weergeven en ze zal daarom voor zulke opdrachten beter scoren dan de overige

leerlingen van haar klas. Bovendien is ze ook bereid andere te helpen die minder goed overweg kunnen met de PC.

**Kimberley** is een bijzonder voorbeeld van een uitzonderlijke sterke leerling die uit het BUSO komt. Zoals verwacht scoort ze lager op de taaltest, maar tegenover de gemiddelde leerling DBSO bevindt ze zich nog net binnen de ondergrens van de gemiddelde marge. Uit ervaring weten we dat zulke leerlingen nog sterk kunnen groeien en vooruitgang maken, indien ze gemotiveerd doorzetten. En Kimberley is dan ook sterk gemotiveerd. Hoewel ze trager doorzicht heeft dan de gemiddelde leerling DBSO, oefent ze erg gedreven en is ze erg leergierig. Ze bereidt zich dan ook zeer goed voor op de testen, studeert en oefent ook thuis. Kimberley is een mooi voorbeeld van een sterke BUSO leerling die erin slaagt door inzet en discipline hoger te scoren dan menig andere leerling op de PAV -testen. De goede resultaten die ze behaalt heeft bovendien een versterkend effect. Ook al heeft Kimberley in tegenstelling tot Hanan meer tijd en moeite nodig voor PC –opdrachten tot een goed einde te brengen, vraagt ze gerichte hulp en leert snel bij. Zo slaagt ze er in toch zeer goede resultaten te behalen. Bovendien heeft ze een sterk geheugen dat haar ook helpt met het instuderen van leerstof voor gesloten boektesten. Ze wil echt weten en kunnen en compenseert daarmee ruim haar gebrek dat ze meer moeite heeft met taal om teksten te verwerken. Uit ervaring weten we dat zulke leerlingen na 1 of twee jaar er kunnen in slagen hun taaltestscores te verbeteren met 10 tot 20%. Het uiterste plafond voor zeer sterke BUSO leerlingen bevindt zich op de taaltest rond 70%. Hoger raken ze niet. Maar voor de opleiding volstaat dit, indien ze ook sociaal vaardig blijken te zijn op de werkvloer en goed kunnen samenwerken met collega's. Hun inzet en doorzettingsvermogen maakt veel goed. Er zijn evenwel ook voorbeelden van BUSO -leerlingen die opvallend zwakke resultaten behalen. Jasmina van VZ2A is een zwakke BUSO- leerling die veel ondersteuning en opvolging nodig heeft.

**Pauline** is een "slimme leerling" maar bij haar heerst de wet van de minste krachtinspanning. Omdat ze makkelijk goede resultaten kan behalen, is ze snel tevreden op haar PAV- testen. In principe komt Pauline in aanmerking om haar diploma middelbaar onderwijs te behalen. Ze heeft zeker de capaciteiten hiervoor, maar haar werkhouding maakt dat ze maar gematigde resultaten behaalt. Ze begint laat met opdrachten en het is zeer de vraag of ze dit schooljaar er in zal slagen deze opdrachten met succes af te werken. Wat Kimberley heeft, ontbreekt bij Pauline. Ze riskeert dan ook hierdoor dat ze pas volgend schooljaar de nodige eindtermen zal bereiken om het diploma middelbaar onderwijs te behalen. Nochtans zou Pauline met wat extra moeite dit diploma zonder teveel problemen kunnen behalen. Pauline heeft TSO gevolgd met succes tot het vierde jaar en heeft getoond in het DBSO getoond ze meer capaciteiten heeft. Ze behaalt dan ook een vergelijkbare score met Emel die tot het 4<sup>de</sup> jaar ASO en het 5<sup>de</sup> jaar TSO les heeft gevolgd in het voltijds onderwijs. Ook in BGV is de inzet van Pauline gematigd en zou ze veel hoger scoren en betere resultaten behalen.

## 6. De waarde van klassieke testcores om jongeren objectiever te kunnen evalueren

*Tegenwoordig is het de mode om in het onderwijs meer en meer leerlingen te beoordelen via ellenlange competentielijsten met een zwart-wit beoordeling: voldoet of voldoet niet. Deze 'evaluitisgril' afkomstig vanuit de wereld van ondernemingen en bedrijven, verbergen echter een zwaar gebrek. Zij lijken objectiever omdat ze zeer exhaustief vele competenties en vaardigheden opsommen waaraan moet voldaan worden. De snelle manier van beoordelen in een zwart-wit mededeling zorgt echter voor een administratieve last die eerder een totale beoordeling in de weg staat en weinig nuances toelaat. Uit competentiebeoordelingen kan men moeilijker statistisch waardevolle genuanceerde informatie filteren zoals men dat met klassiekere testen wel kan doen.*

Bovendien zijn ze omwille van de exhaustieve opsomming vaak hemeltergend ineffectief en tijdrovend omdat de beoordelingen berusten op momentopnames en de quoteringen weinig oog heeft voor nuances. Indien voor PAV elke les of thema naar eindtermen toe afzonderlijk zou beoordeeld worden, zal de ene keer een jongere voldoende scoren en een andere keer onvoldoende. Om 1 willekeurige eindterm of leerplandoel op deze wijze te beoordelen is het risico groot dat jongeren negatiever worden beoordeeld of te positief worden beoordeeld afhankelijk van de frequentie en context waarin de eindterm of leerplandoel wordt geëvalueerd. Het overzicht ontbreekt en vaak moeten leerkrachten dan vaststellen dat de zogenaamde geobjectiveerde competentie en vaardigheidsbeoordelingen niet altijd overeenstemmen met hun algemene oordeel of de jongere nu al of niet de eindtermen of het leerplandoelen voldoende heeft bereikt. Het risico is daarom groot dat zulke

competentielijsten en leerplandoelen worden aangepast of gecorrigeerd in de loop van het schooljaar zodat ze toch min of meer in overeenstemming komen met het algemene oordeel van de leerkrachten. Dat is ook de reden waarom een leerplandoel rapport overzichtelijker en objectiever is. Een leerplandoel rapport geeft een graduele score (dat in essentie een puntenscore is!!!) weer die meer overeenstemt met het algemene oordeel die leerkrachten vormen over leerlingen en een overzicht geeft over een langere periode van meerdere clusters van eindtermen. Vaak verdenkt men dan leraren van subjectieve invulling, maar zoals blijkt uit de PAV- test scores en taalttest scores, blijken test scores meer toetsbare en objectievere informatie te bevatten dan men denkt en die bovendien nog meer mogelijkheden geven om deze statistisch te onderzoeken in hoeverre de cijfers of scores overeenstemmen met het oordeel dat men heeft over de jongere.

Hanan, Pauline en Kimberley zijn mooie voorbeelden om dat te staven. Op het eerste gezicht lijkt Pauline een prima leerling te zijn als je naakt naar haar scores kijkt zonder ze te plaatsen in de schoolcontext. Maar wanneer je ze koppelt aan onderzoek en plaatst in hun klassikale context, stellen we vast dat Pauline inderdaad uitschiet op een aantal vlakken en dat de commentaren van Pauline in de TBO's niet uit de lucht zijn gegrepen. Uit de analyse van de test scores van Pauline komt duidelijk naar voren dat ze ook overeenstemmen met het algemene oordeel van leerkrachten over Pauline en dat deze inderdaad niet zo subjectief lijken te zijn. Dezelfde vaststellingen gelden voor Hanan en Kimberley, maar ook voor andere leerlingen zoals Donika. Uit de laatste TBO blijkt dat Donika nog beter zou kunnen presteren op de werkvloer en in de BGV lessen voor haar opdrachten als ze de Nederlandse taal nog beter zou beheersen. Uit de PAV- testen en de taalttest stellen we vast dat inderdaad Donika het laagst presteert van alle in haar klas en dat er sprake is van een duidelijke samenhang met taal.

**Leerkrachten evalueren vaak sneller en effectiever zonder daarom ellenlange competentielijsten te hoeven raadplegen om jongeren goed te evalueren. Onze hersenen zien sneller patronen dan deze zichtbaar worden in ingewikkelde en exhaustieve onoverzichtelijke evaluatielijsten, die teveel dienen om inspectie tevreden te stellen en hun zucht naar overdreven bewijslast, maar naar jongeren weinig meerwaarde hebben en door leerkrachten terecht vaak als overdreven tijdrovend en weinig efficiënt worden beschouwd.**

#### 7. Is de opvolging van competenties, eindtermen en leerplandoelen dan een nutteloze bezigheid?

Neen, het is de zware, exhaustieve, tijdrovende, inefficiënte en onoverzichtelijke boekhoudkundige wijze van evalueren die leerkrachten terecht afwijzen.

**Het nut en de zin van eindtermen, competenties en leerplandoelen bestaat uit het gericht oog hebben voor de facetten die aandacht verdienen om jongeren te evalueren. Het vergroot de invalshoek en beoordelingscriteria om jongeren beter te kunnen evalueren. Maar indien het doel een middel wordt om het algemene oordeel van leerkrachten uit te schakelen of te herleiden tot het opsommen en invullen van boekhoudkundige formulieren, dan vervalt elk nut en zin ervan. Teveel diverse evaluatietechnieken en de exhaustieve opsomming van tientallen competenties hebben een averechts effect: ze worden ingevuld om in orde te zijn en niet meer om werkelijk te evalueren. De leerkracht wordt geïnfantiliseerd tot het boekhoudkundig weergeven en invullen van evaluatieformulieren die een persoonlijke beoordeling zoveel mogelijk moeten uitsluiten onder het mom van streven naar meer objectiviteit die helemaal niet wordt aangetoond of bewezen!!! Men wil zoveel mogelijk het persoonlijke oordeel van leerkrachten vervangen door objectieve lijsten waarvan de objectiviteit meer schijn is en daarom bureaucratisch wordt aangepast om in orde te zijn met de regelgeving. Op deze manier wordt de leerkracht herleid tot een domme incompetent evaluator die moet geholpen worden door experts die weinig rekening houden met de klassikale en schoolse context. De huidige dwang om te screenen en te evalueren tot het lachwekkende toe, doet denken aan het verhaal van dokter Knock. Zijn er niet genoeg patiënten, vertel dan dat ze ziek zijn en je hebt opnieuw vele cliënten. Aan placebo- evaluatie hebben leerkrachten geen boodschap!**

Bovendien (en dat is niet enkel een vaststelling in het onderwijs, maar ook bij coaches op de werkvloer in bedrijven die mee werknemers opleiden en opvolgen!!) worden de meeste competenties zonder veel nuances beoordeeld: geslaagd of niet geslaagd/ bereikt of niet bereikt. Zulke tweedelingen zorgen al voor problemen om dat later zinvol te hanteren als opsomming belangrijker wordt dan het overzicht. Omdat bovendien elke nuance wordt uitgesloten zijn ze per definitie minder objectief. daarenboven bevordert deze manier van werken en evalueren 'reglementitis'!! En werknemers die daarmee geconfronteerd reageren dan zeer spontaan op het

getrouw klakkeloos invullen van formulieren die nadien weinig meerwaarde geven tot analyse dan de loutere beschrijving ervan. Het is net zoals het klimaat beoordelen op grond van het weer. Indien men plaatselijke en tijdelijke meteorologische metingen (hoe gedetailleerd ook) laat primeren tegenover een overzicht op langere termijn en een groter oppervlak. Men zal dan nooit een totaalbeeld en algemeen patroon kunnen ontdekken en beschrijven van het klimaat dat heerst in een bepaald gebied gedurende een lange periode. Omdat we met jongeren werken zijn weersvoorspellingen minder relevant dan klimaatvoorspellingen. En leerkrachten hebben daarom ook een hekel aan het 'meteorologische' evalueren van jongeren. Leerkrachten wensen voornamelijk een totaalbeeld te kunnen vormen van jongeren en op basis daarvan te beoordelen of ze slagen of niet. Tijdelijke 'onweersbuien, stormen' zijn dan maar enkel relevant wanneer ze te frequent voorkomen dan normaal verwacht wordt en bijdragen tot klimaatwijziging. Leerkrachten zijn daarom minder geïnteresseerd om dagelijks het weer te noteren, tenzij bepaalde afwijkende weerspatronen uitzonderlijk meer voorkomen dan normaal verwacht. Het komt er dus op neer op een verstandige wijze steekproefsgewijs te evalueren en de meest relevante informatie te laten doorwegen bij een eindbeoordeling. Deze eindbeoordeling mag dan wel gebaseerd zijn op competenties en eindtermen, maar daarom moet de beoordeling ervan overzichtelijk, snel en efficiënt zijn en overeenstemmen met het globale oordeel van de leerkracht. Leerkrachten hebben totaal geen behoefte aan de vaak ergerlijke, infantiliserende en het exhaustieve invullen van evaluatiemappen en formulieren met diverse criteria, wanneer ze weinig tot niets bijdragen tot betere en overzichtelijke informatie over de jongere.

**Leerkrachten hebben geen behoefte aan 'evaluïtis', ze hebben echter wel de behoefte aan klare overzichtelijke informatieve evaluatie zonder teveel reglementitis en Kafkaïaanse opgelegde bewijslastcriteria. En zoals ook blijkt uit het onderzoek: Punten zijn vaak informatiever en objectiever dan menige exhaustieve competentielijst of eindtermenlijst die wordt ingevuld met de tweedelige informatie: bereikt / niet bereikt. Zij laten meer ruimte open voor nuance en we kunnen ze zelfs toetsen en onderwerpen aan statistisch onderzoek om betere conclusies te trekken en oordelen van vooroordelen te onderscheiden.**