VOORWAARDEN VOOR EFFECTIEVE FORMATIEVE TOETSING

Een praktische review

Reviewstudie uitgevoerd in opdracht van en gesubsidieerd door NRO-PPO: Dossiernummer 405-14-534, eindrapport 30-09-2014

Kim Schildkamp[a], Maaike Heitink[a], Fabienne van der Kleij[b], Inge Hoogland[a], Anne Dijkstra[a], Wilma Kippers[a] en Bernard Veldkamp[a]

[a] Universiteit Twente (contact: k.schildkamp@utwente.nl)
[b] Australian Catholic University

UNIVERSITEIT TWENTE.

INHOUDSOPGAVE

1 Samenvatting 3

2 Voorwoord en aanleiding 6

3 Inleiding en onderzoeksvraag 7

4 Theoretische Achtergrond 10

4.1 VOORWAARDEN TOETS 10

4.2 VOORWAARDEN DOCENT 10

4.3 VOORWAARDEN LEERLING 11

4.4 VOORWAARDEN CONTEXT 11

5 Resultaten Literatuur review 12

5.1 VOORWAARDEN TOETS 12

6.1.1 Voorwaarden toets: Kenmerken van formatieve toetsen 12

6.1.2 Voorwaarden toets: Data systemen, data en feedback 13

6.1.3 Voorwaarden toets: Digitale toetssystemen 14

6.1.4 Voorwaarden toets: Het proces van (digitale) formatieve toetsing 15

6.1.5 Voorwaarden toets: (Geautomatiseerde) Feedback 16

6.1.6 Opvallende verschillen: Voorwaarden toets 16

5.2 VOORWAARDEN DOCENT 17

6.2.1 Voorwaarden docent: Kennis, vaardigheden en attitude 17

6.2.2 Voorwaarden docent: Samenwerken en het stellen van doelen 18

6.2.3 Voorwaarden docent: Het betrekken van leerlingen 19

6.2.4 Opvallende verschillen: voorwaarden docent 19

5.3 VOORWAARDEN LEERLING 19

6.3.1 Voorwaarden leerling: Kennis, vaardigheden en houding 20

6.3.2 Voorwaarden leerling: Rol bij het proces 20

6.3.3 Opvallende verschillen: voorwaarde leerling 20

5.4 VOORWAARDEN CONTEXT 21

6.4.1 Context binnen de school: Kenmerken van de school 21

6.4.2 Context binnen de school: Benodigde tijd en middelen 22

6.4.3 Context binnen de school: Het mogelijk maken van docentprofessionalisering 22

6.4.4 Context binnen de school: Leiderschap 23

6.4.5 Context buiten de school 24

6.4.6 Opvallende verschillen: Voorwaarden context 24

6 Resultaten focusgroepen 25

6.1 FOCUSGROEPEN PO 25

6.1.1 Voorwaarden toets: PO 25

6.1.2 Voorwaarden docent: PO 26

6.1.3 Voorwaarden leerling: PO 27

6.1.4 Voorwaarden context: PO 27

6.1.5 Voorwaarden ouders: PO 28

6.2 FOCUSGROEPEN VO 29

6.2.1 Voorwaarden toets: VO 29

6.2.2 Voorwaarden docent: VO 30

6.2.3 Voorwaarden leerling: VO 31
6.2.4 Voorwaarden context: VO 31
6.3 FOCUSGROEP MBO 32
6.3.1 Voorwaarden toets: MBO 33
6.3.2 Voorwaarden docent: MBO 33
6.3.3 Voorwaarden leerling: MBO 34
6.3.4 Voorwaarden context: MBO 34
6.4 FOCUSGROEP EXPERTS 35
6.4.1 Voorwaarden toets: Experts 35
6.4.2 Voorwaarden docent: Experts 36
6.4.3 Voorwaarden leerling: Experts 36
6.4.4 Voorwaarden context: Experts 36

7 Conclusie 38
7.1 LITERATUUR REVIEW 38
7.2 FOCUSGROEPEN 39
7.3 RICHTLIJNEN VOOR DE PRAKTIJK 40

8 Discussie 42

9 Referenties 44
9.1 REFERENTIES GEBRUIKT VOOR THEORIE IN HET RAPPORT 44
9.2 REFERENTIES LITERATUUR REVIEW 45

Bijlage 1: Methode/proces 49

1 SAMENVATTING

Summatief en formatief toetsen

Toetsing is een essentieel onderdeel van onderwijs. Toetsen kunnen niet alleen helpen de voortgang van leerlingen en effectiviteit van lesgeven in kaart te brengen, maar kunnen ook als instrument dienen om onderwijs op maat aan te bieden. Traditioneel wordt in het onderwijs onderscheid gemaakt tussen summatieve en formatieve toetsen. Terwijl summatieve toetsen gericht zijn op het beoordelen van leerprestaties, hebben formatieve toetsen als doel informatie te geven aan docenten en leerlingen over de mate waarin de leerstof beheerst wordt. Deze informatie kan worden gebruikt om het leerproces waar nodig bij te sturen en zo lesgeven en leren te verbeteren. Onder toetsen verstaan we niet alleen formele toetsen, maar alle manieren waarop bewijs wordt verzameld over de voortgang van het leren van leerlingen. Het kan bijvoorbeeld gaan om een presentatie, portfolio, observaties in de les, discussies en praktische opdrachten.

De vele onderzoeken die tot nu toe zijn uitgevoerd op het gebied van formatief toetsen laten zien dat het niet eenvoudig is om formatief toetsen effectief in te zetten in de onderwijspraktijk. Het onderzoek beschreven in dit rapport heeft daarom als doel docenten inzicht te geven in factoren die van belang zijn bij het inzetten van formatief toetsen. Door factoren te belichten die een rol spelen bij de toets, de docent, de leerling en de context wordt een totaalbeeld geschetst van voorwaarden die nodig zijn om formatief toetsen op een succesvolle manier te implementeren in de praktijk. Dit onderzoek richt zich specifiek op het primair-, voortgezet- en middelbaar beroeps –onderwijs (PO, VO en MBO).

Formatief toetsen: opbrengstgericht werken, assessment for learning en diagnostisch toetsen

In dit onderzoek is toegespitst op de volgende drie verschillende benaderingen van formatief toetsen: opbrengstgericht werken (OGW), assessment for learning (AfL) en diagnostisch toetsen (DT). Opbrengstgericht werken omvat het systematisch verzamelen en analyseren van gegevens, om op basis hiervan beslissingen te kunnen nemen met betrekking tot het verbeteren van het onderwijs. Assessment for learning is onderdeel van de dagelijkse praktijk van docenten en leerlingen en is gericht op het sturen van het leerproces. Diagnostisch toetsen heeft als doel leerbehoeften van individuele leerlingen te identificeren en zo de ontwikkelingsfase waarin een leerling zich bevindt in kaart te brengen.

Voorwaarden voor de inzet van formatief toetsen in de klas

Aan de hand van een reviewstudie zijn voorwaarden achterhaald voor het gebruik van deze drie benaderingen van formatief toetsen in de praktijk. Daarnaast zijn de gevonden voorwaarden geverifieerd, aangevuld en vertaald naar de praktijk met behulp van focusgroepen (discussiegroepen) waarin zowel experts (4) en docenten uit het PO (4), VO (12) en MBO (4) deelnamen. De zoek- en selectieprocedure voor deze review resulteerde in 72 studies. Het valt op dat van de 72 gevonden studies maar drie studies in de context van het MBO uitgevoerd zijn. Ook zijn er maar weinig studies gericht op diagnostisch toetsen gevonden. Dit wil niet zeggen dat de voorwaarden gevonden in deze studie minder van belang zijn voor diagnostisch toetsen of het MBO, maar dat op deze gebieden nog niet veel wetenschappelijk bewijs verzameld is. Hieronder zullen de voorwaarden m.b.t. de toets, de docent, de leerling en de context kort besproken worden

Voorwaarden m.b.t. de toets

De toets is een middel om formatief toetsen uit te voeren. Het is daarom belangrijk dat toetsen betekenisvol zijn en deze zowel inhoudelijk als didactisch aansluiten bij het curriculum, de leerdoelen en het niveau van de leerlingen. Aansluitend hierop werd specifiek in de focusgroep met docenten uit het middelbaar beroepsonderwijs genoemd dat toetsen opdrachten en taken moeten omvatten die

gerelateerd zijn aan de relevante beroepspraktijk. Daarnaast is het belangrijk dat het doel van de toets en de beoordelingscriteria van te voren duidelijk zijn voor zowel de docent als de leerlingen. Verder wordt aanbevolen dat formatieve toetsen regelmatig afgenomen worden (minimaal om de zes weken), zodat leerlingen regelmatig feedback ontvangen, dat leerlingen genoeg tijd krijgen om toetsvragen te beantwoorden en dat toetsen (klassikaal) besproken worden. Tijdens de focusgroepen met docenten werd tevens aangestipt dat frequent toetsen noodzakelijk is, maar wel met mate moet gebeuren. Verder is een belangrijke voorwaarde bij formatief toetsen dat toetstaken ook in feedback moeten voorzien. Feedback heeft een effect op de motivatie om te leren.

Voorwaarden m.b.t. de docent

De docent speelt een essentiële rol bij formatief toetsen in de klas. Veel studies, maar ook deelnemers in alle focusgroepen, gingen in op competenties die docenten nodig hebben om formatief toetsen te kunnen integreren in het onderwijs. Verder kwamen onderwerpen als het geven van feedback, samenwerking en het betrekken van leerlingen in het toetsproces aan bod. Docenten hebben kennis en vaardigheden nodig om toetsgegevens te kunnen verzamelen, analyseren en interpreteren. Vervolgens moeten docenten in staat zijn op basis van deze resultaten leren en lesgeven te verbeteren. Dit laatste betekent dat ook een goede didactische en vakinhoudelijke basis van de docent verlangd wordt. Daarnaast bleek uit de resultaten van een beperkt aantal studies dat docenten ook in staat moeten zijn zelf toetsen te ontwikkelen. In de focusgroepen waren de meningen hierover verdeeld. Niet alleen kennis en vaardigheden maar ook een positieve houding ten opzichte van formatief toetsen bleek belangrijk voor het daadwerkelijk gebruik van deze toetsen. Tot slot is een onmisbare competentie bij het implementeren van formatief toetsen het vermogen van docenten om goede feedback te geven. Feedback moet een kort maar krachtig advies zijn dat gebaseerd is op heldere criteria en gericht is op verbeterpunten.

Naast kennis en vaardigheden die docenten nodig hebben is ook de samenwerking tussen docenten (en schoolleiding) van belang. Docenten kunnen veel van elkaar leren en het uitwisselen van ervaringen kan stimulerend en motiverend werken. Daarbij blijkt een effectieve strategie het gezamenlijk doelen stellen omtrent het gebruiken van (toets)gegevens in de school. Daarnaast is het niet alleen bevorderlijk als docenten samenwerken met collega's omtrent formatief toetsen, maar ook met leerlingen. Het betrekken van leerlingen in het toetsproces vergt een coachende functie van de docent waarin hij/zij leerlingen begeleidt bij bijvoorbeeld het beoordelen van eigen werk of het beoordelen van medeleerlingen.

Voorwaarden m.b.t. de leerling

De voorwaarden die betrekking hebben op de leerling gaan in op de rol van leerlingen in het formatieve toetsproces en de competenties die leerlingen nodig hebben om te kunnen profiteren van formatief toetsen in de klas. Veel studies benadrukken dat het belangrijk is om leerlingen een actieve rol te laten spelen in het toetsproces. Dit kan leiden tot hogere motivatie en betere leerprestaties. Daarnaast hebben leerlingen kennis en vaardigheden nodig om zowel eigen werk als het werk van medeleerlingen van nuttige feedback te kunnen voorzien. Aan de hand van toetscriteria moeten leerlingen in staat zijn op hun eigen werk te reflecteren, om zo verbeterpunten en vervolgstappen in hun eigen leerproces te kunnen identificeren. Het verkrijgen van dergelijke kennis en vaardigheden vergt oefening voor leerlingen. Ook noemen docenten in de focusgroepen dat rekening gehouden moet worden met een bewustwordingsproces, omdat de manier van toetsen anders is dan veel leerlingen gewend zijn (leerlingen zijn alleen aan summatief toetsen gewend).

Voorwaarden m.b.t. de context

Om aan bovengenoemde voorwaarden te kunnen voldoen is facilitering vanuit de context cruciaal. Dit betreft factoren van de context binnen de school en factoren van de context buiten de school. Binnen de school gingen de meeste studies in op de cultuur van de school, leiderschap, het faciliteren van tijd en

middelen en het faciliteren van professionele ontwikkeling. Buiten de school richtten voorwaarden zich vooral op overheidsregelingen en beleid. Een belangrijke voorwaarde binnen de school is de aanwezigheid van een toetscultuur waarin het verbeteren van het onderwijs centraal staat en ruimte is om te experimenteren. Een gedeelde visie, normen, doelen en verwachtingen zijn in een dergelijke cultuur van belang. Zowel de literatuur als alle deelnemers van de focusgroepen wijzen erop dat leiderschap hierin een essentiële rol speelt. Schoolleiders zijn niet alleen als cultuurdrager geïdentificeerd, maar moeten ook zorgen voor de nodige ondersteuning en infrastructuur. Ook het investeren van tijd en middelen zijn hierin van belang en dit wordt nog eens benadrukt door de docenten die deelnamen aan de focusgroepen. Docenten hebben tijd nodig om samen te werken aan formatieve toetsen en budget moet uitgetrokken worden voor professionele ontwikkeling. Professionele ontwikkeling van docenten en schoolleiders is nodig voor zowel het gebruiken, analyseren en interpreteren van data, als het vertalen van de resultaten naar interventies in de klas. Ondersteuning door specialisten wordt daarbij als wenselijk ervaren.

Conclusie

Uit de literatuurreview en de focusgroepen komt duidelijk naar voren dat er behoefte is aan professionalsering en ondersteuning op het gebied van formatief toetsen. Er is een grote mate van overeenstemming gevonden tussen de wetenschap en de praktijk. Het belang van formatief toetsen is duidelijk, maar de toepassing hiervan in de klas vraagt nog veel van scholen, schoolleiders, docenten en leerlingen. Vervolgonderzoek moet duidelijk maken op welke wijze formatief toetsen geïmplementeerd kan worden zodat het leidt tot de beoogde resultaten; het leren van leerlingen.

2 VOORWOORD EN AANLEIDING

Formatief toetsen kan leiden tot betere leeropbrengsten van leerlingen, mits het op de juiste manier wordt ingezet. Er is veel onderzoek gedaan naar formatieve toetsing in het onderwijs en de effecten hiervan. Onderzoeksresultaten blijken echter moeilijk te vertalen naar de praktijk (Sluijsmans, Joosten-ten Brinke, & Van Der Vleuten, 2013). Dit onderzoek richt zich daarom op de voorwaarden waaraan moet worden voldaan om formatief toetsen zo in te zetten binnen in de praktijk, dat het tot de gewenste resultaten leidt: betere leerprestaties van leerlingen. In dit rapport wordt onderscheid gemaakt tussen (1) voorwaarden met betrekking tot de toets zelf, het toetssysteem en het proces (inclusief feedback), (2) voorwaarden met betrekking tot de docent, (3) voorwaarden met betrekking tot de leerling en (4) voorwaarden met betrekking tot de context.

Dit onderzoek richt zich op het primair onderwijs (PO), voortgezet onderwijs (VO) en middelbaar beroepsonderwijs (MBO). Binnen deze onderwijssectoren verschilt de gangbare terminologie. Voor een eenduidige presentatie van de resultaten zal consistent gebruik worden gemaakt van de termen *docent* en *leerling*, deze termen kunnen echter ook worden gelezen als *leerkracht* en *student*. De resultaten worden apart gepresenteerd voor het PO, VO en MBO. Gevonden verschillen en overeenkomsten worden bediscussieerd in de conclusie en discussie van dit rapport.

Dit rapport heeft als doel inzicht te geven in factoren die van belang zijn bij het inzetten van formatief toetsen. Door middel van een review van wetenschappelijke literatuur is een samenvatting gemaakt van factoren die van belang zijn. Door middel van focusgroepen met docenten uit het PO, VO en MBO en onderwijsexperts zijn de gevonden voorwaarden geverifieerd, aangevuld en vertaald naar de praktijk.

De resultaten van het onderzoek geven richting aan de onderwijspraktijk op het gebied van formatief toetsen. De vele onderzoeken die tot nu toe zijn uitgevoerd naar dit onderwerp wijzen uit dat het niet eenvoudig is om formatief toetsen effectief in te zetten. Door factoren te belichten die een rol spelen met betrekking tot de toets, de docent, de leerling en de context wordt een totaalbeeld geschetst van bevorderende factoren. De resultaten van dit onderzoek geven ook input voor de ontwikkeling van een interventie rondom docentprofessionalisering op het gebied van formatief toetsen, recent onderzoek heeft aangetoond dat hier behoefte aan is (Sluijsmans et al., 2013). Veel docenten hebben namelijk niet of nauwelijks de kans gehad om te leren over formatief toetsen in hun formele opleiding.

3 INLEIDING EN ONDERZOEKSVRAAG

In het onderwijs wordt onderscheid gemaakt tussen formatieve en summatieve doelen van toetsing. Summatieve toetsen hebben als doel de leerprestaties van een leerling te beoordelen en vervolgens een beslissing te nemen met betrekking tot selectie, classificatie, plaatsing of certificering (Sanders, 2011). Formatieve toetsen hebben als doel informatie te geven aan de docent en leerlingen over de mate waarin de leerstof al wel of niet beheerst wordt. Deze informatie kan worden gebruikt om het leerproces waar nodig bij te sturen. Onder toetsen verstaan we niet alleen het formele toetsen, maar alle manieren waarop bewijs wordt verzameld over de voortgang van het leren van leerlingen. Dit wordt ook wel aangeduid met de term *assessment*. Het kan bijvoorbeeld gaan om een presentatie, portfolio, observaties in de les, discussies en praktische opdrachten. In dit rapport wordt de term 'toetsen' uitwisselbaar gebruikt met de term assessment.

Beweringen van onderzoekers over een positief effect van formatief toetsen op het leren van leerlingen worden slechts in beperkte mate gestaafd door degelijk wetenschappelijk onderzoek. Bestaand onderzoek naar deze effecten is vaak gebaseerd op onzuivere, ongepubliceerde of verouderde studies (Bennett, 2011). Daarnaast bestaat er in de wetenschap geen eenduidige definitie van formatief toetsen, wat het lastig maakt een goede inschatting te maken van de effecten van formatief toetsen. Desondanks zijn veel onderzoekers overtuigd van de mogelijk positieve effecten die formatief toetsen kan hebben op leren. Door de breedheid van de definitie van formatief toetsen geeft wetenschappelijk onderzoek echter slechts in beperkte mate richting aan de praktijk. Er zijn namelijk heel veel verschillende manieren om toetsresultaten in te zetten om leren te bevorderen.

Om duidelijkheid te scheppen voor zowel de wetenschap als de praktijk hebben Van Der Kleij, Vermeulen, Schildkamp en Eggen (2013) een literatuurstudie uitgevoerd. Zij definiëren formatief toetsen als elke toets die feedback geeft welke bedoeld is om leren te ondersteunen en gebruikt kan worden door docenten en/of leerlingen. Verder maken zij onderscheid tussen drie benaderingen van formatief toetsen: opbrengstgericht werken (OGW), assessment for learning (AfL) en diagnostisch toetsen (DT). De betekenis van deze drie benaderingen wordt in onderstaande paragrafen verder toegelicht.

OGW omvat het systematisch verzamelen en analyseren van data om op basis hiervan beslissingen te kunnen nemen met betrekking tot het verbeteren van het onderwijs (Schildkamp & Kuiper, 2010; Schildkamp, Lai & Earl, 2013). AfL is onderdeel van de dagelijkse praktijk van docenten en leerlingen waarbij het gaat om het zoeken van informatie en het reflecteren en het reageren op informatie uit dialogen, demonstraties en observaties op een manier die bijdraagt aan het leerproces (Klenowski, 2009). DT betreft het identificeren van specifieke leerbehoeften van leerlingen (Wiliam, 2011). Bij DT gaat het erom te achterhalen hoe een bepaalde taak wordt opgelost als indicatie voor de ontwikkeling van de betreffende leerling (Crisp, 2012; Keeley & Tobey, 2011). In Tabel 1 zijn de belangrijkste verschillen en overeenkomsten tussen de drie benaderingen van formatief toetsen samengevat.

TABEL 1 DE BELANGRIJKSTE VERSCHILLEN EN OVEREENKOMSTEN TUSSEN OGW, AFL EN DT (GEBASEERD OP VAN DER KLEIJ ET AL., 2013)

	OGW	AfL	DT
Doelen	De kwaliteit van onderwijs en instructie verbeteren door middel van het gebruik van data bij het monitoren en sturen van de onderwijspraktijk om opgestelde doelen te bereiken (bijv. verhoogde leerlingprestaties)	De kwaliteit van het leerproces verbeteren door leerlingen actief te betrekken bij het evalueren van en reflecteren op hun eigen leerproces en leerprocessen sturen door middel van continue feedback.	Gedetailleerde gegevens verzamelen over de zone van naaste ontwikkeling, voorkennis en redeneerstijlen van leerlingen om beslissingen te nemen over aanpassingen in de leeromgeving naar aanleiding van behoeften van de leerling.
Niveau	▪ School ▪ Klas ▪ Leerling	▪ Klas ▪ Leerling	▪ Leerling
Toetsmethoden	▪ Gestandaardiseerde toetsen ▪ Door docenten geconstrueerde toetsen ▪ Gestructureerde observaties in de les	▪ Informele dialogen in de klas ▪ Door docenten geconstrueerde toetsen ▪ Praktische demonstraties ▪ Portfolio's ▪ Peer en self-assessment	▪ (Adaptieve) toetsen met toetsvragen die gedetailleerde informatie aan het licht brengen over de redenering van de leerling
Feedback loops	▪ Directe en uitgestelde feedback ▪ Retrospectief; naar aanleiding van een formeel toetsmoment remediëren	▪ Directe feedback ▪ Interactief, continue feedback tijdens de les. Soms retrospectief (remediërend) of proactief (leren van voorgaande leerlingcohorten)	▪ Directe en uitgestelde feedback ▪ voornamelijk retrospectief (remediërend), mogelijk interactief of proactief

De drie benaderingen van formatief toetsen overlappen gedeeltelijk, maar er zijn ook duidelijke verschillen. Alle benaderingen hebben als doel informatie te verzamelen over het leren van leerlingen en dit in te zetten om leren te bevorderen. Verschillen zijn bijvoorbeeld te vinden in het niveau, de gebruikte toetsmethodes en de verschillende feedbackloops (zie Tabel 1) (Van der Kleij et al., 2013). OGW heeft bijvoorbeeld zowel betrekking op het niveau van de leerling, de klas en de school, terwijl AfL gericht is op het niveau van de klas en de leerling. DT is alleen gericht op het niveau van de leerling en gaat dieper in op specifieke onderwijsbehoeften van leerlingen dan AfL (zie Figuur 1).

FIGUUR 1 NIVEAUS VAN OGW, AFL EN DT (GEBASEERD OP VAN DER KLEIJ ET AL., 2013)

Doordat deze drie benaderingen van formatief toetsen (deels) verschillende doelen hebben en zich op verschillende niveaus in het onderwijs afspelen zijn hier andere voorwaarden voor de praktijk aan verbonden. Daarom is het zinvol om bij het identificeren van factoren die van belang zijn bij het inzetten van formatief toetsen onderscheid te maken tussen OGW, AfL en DT. In dit onderzoek wordt daarom onderscheid gemaakt tussen OGW, AfL en DT, of een combinatie van deze benaderingen.

Ondanks dat er al veel (review)studies zijn uitgevoerd naar formatieve toetsing blijkt de bruikbaarheid hiervan voor de onderwijspraktijk beperkt te zijn. Het is daarom belangrijk om samen met de praktijk te onderzoeken aan welke voorwaarden voldaan moet worden om ervoor te zorgen dat formatief toetsen wel tot de gewenste effecten leidt: betere leerprestaties van leerlingen. Om deze reden is gedurende het onderzoek gefocust op praktische studies naar formatief toetsen in de klas. Tevens hebben we docenten uit het PO, VO, MBO en experts bevraagd met betrekking tot de voorwaarden waar aan voldaan moet worden voor het effectief gebruiken van formatieve toetsing in de klas. De hoofdvraag van deze studie luidt:

> *Aan welke voorwaarden moet worden voldaan m.b.t (1) de toets zelf, toetsinstrumenten en het toetsproces inclusief de vorm van feedback die het oplevert, (2) de rol van de docent, (3) de rol van de leerling en (4) de context waarin de formatieve toetsing plaats vindt, zodat formatieve toetsing leidt tot betere leerprestaties van leerlingen?*

4 THEORETISCHE ACHTERGROND

De wetenschappelijke literatuur op het gebied van formatief toetsen suggereert dat het nuttig is onderscheid te maken tussen voorwaarden met betrekking tot de toets, de docent, de leerling en de context. In dit hoofdstuk wordt voor elk van deze factoren een bondige theoretische onderbouwing en toelichting gegeven.

4.1 VOORWAARDEN TOETS

De kenmerken van de toets zelf en de vorm van feedback die een toets oplevert kunnen beïnvloeden hoe en of de inzet van formatieve toetsen (OGW, AfL en/of DT) leidt tot het leren van leerlingen. Een belangrijk uitgangspunt bij de kwaliteit van toetsing is of de doelen van de toets eenduidig vastgesteld zijn en of de toets aansluit bij deze doelen. Bij formatief toetsen is het specifiek van belang dat de toets informatie oplevert die inzicht geeft in het leren van leerlingen om sturing te kunnen geven aan het leren. De kenmerken van de toets bepalen voor een groot deel welke en hoe gedetailleerde feedback de toets kan geven. Welk type informatie van belang is, hangt af van de benadering van formatief toetsen. In OGW ligt de nadruk op objectieve informatie; dit type informatie wordt vaak verzameld met formele toetsen, meestal ontwikkeld door toetsspecialisten. Bij AfL speelt de docent een grote rol in het toetsontwerp: het toetsen vindt plaats als een geïntegreerd aspect van de les, en er wordt constant ingespeeld op de behoeften van leerlingen. Bij DT moet de toets kunnen identificeren wat de leerling doet, hoe die dat doet en of er misconcepties zijn binnen een bepaald gedefinieerd leerdomein. Het soort informatie dat men probeert te verkrijgen met een toets bepaalt voor een groot deel de kenmerken van de toets, zoals de aard van het toetsproces en de lengte van de toets. De resultaten van meerdere toetsen geven een beter beeld van de voortgang van een leerling dan de resultaten op een enkel toetsmoment (Van der Kleij et al., 2013).

Binnen alle benaderingen speelt feedback een belangrijke rol. Feedback kan worden omschreven vanuit twee perspectieven (Sadler, 1989): 1) de informatie in de vorm van toetsresultaten die docenten en andere belanghebbenden inzicht geeft in het leren van leerlingen en 2) feedback die direct gegeven wordt aan leerlingen op basis van hun antwoorden op een toets of toetsvraag. Technologie wordt tegenwoordig veelvuldig ingezet voor het geven van beide vormen van feedback. Zo worden computerprogramma's ingezet om docenten feedback te geven in de vorm van automatisch gegenereerde toetsrapportages uit leerlingvolgsystemen. Ook wordt technologie vaker ingezet om de resultaten van leerlingen op toetsen te analyseren en monitoren. Een ander voorbeeld is het gebruik van digitale formatieve toetsing waarbij directe feedback wordt gegeven aan leerlingen op basis van hun antwoord op een toetsvraag (zie http://toetswijzer.kennisnet.nl/html/feedback/default.shtm).

Verder wordt het leerlingen laten beoordelen van hun eigen werk of dat van hun medeleerlingen, ook wel aangeduid met *peer- en self-assessment*, in de literatuur erkend als een belangrijk onderdeel van formatief toetsen. Leerlingen kunnen hierdoor leren hun leerproces te reguleren doordat zij inzichten krijgen in kwaliteitscriteria. Deze inzichten kunnen leerlingen gebruiken om toekomstig leren te sturen. Tevens bevordert het gebruik van peer- en self-assessment de ontwikkeling van metacognitieve vaardigheden, oftewel *leren hoe te leren* (Stobart, 2008).

4.2 VOORWAARDEN DOCENT

De docent is van groot belang voor het effectief inzetten van formatieve toetsing. Docenten moeten namelijk dagelijks beslissingen te nemen op basis van gegevens over hoe het leerproces van leerlingen te sturen. Een algemene term voor toetsvaardigheden die wordt gebruikt in de literatuur is *assessment*

literacy. Deze term verwijst naar de capaciteit van docenten om (zowel individueel als gezamenlijk) toetsresultaten te analyseren en begrijpen, en om op basis hiervan plannen te ontwikkelen op het niveau van de klas en de school om betere resultaten te behalen (Fullan & Watson, 2000). Een correcte interpretatie van toetsgegevens is noodzakelijk voor het op een adequate manier aanpassen van de instructie aan de behoeften van de leerling. Daarnaast is vakdidactische kennis cruciaal, omdat de docent in moet springen op de behoeften van leerlingen. Tot slot is het van belang om te kijken naar de samenwerking tussen docenten (en schoolleiders). Effectieve samenwerking tussen docenten kan de inzet van formatieve toetsen bevorderen. Hierbij gaat het om het samen bespreken van toetsresultaten, hoe deze tot stand zijn gekomen en hoe deze verbeterd zouden kunnen worden (Schildkamp et al., 2013).

4.3 VOORWAARDEN LEERLING

De rol van de leerling in formatief toetsen wordt niet veelvuldig bediscussieerd in de wetenschappelijke literatuur. In het proces van (effectief) formatief toetsen speelt de leerling echter een actieve rol. De rol van leerlingen in het proces van peer- en self-assessment is evident. Maar om effectief formatief toetsen te faciliteren, moeten leerlingen ook actief meewerken aan verschillende vormen van toetsing en actief gebruik maken van feedback wil dit een effect kunnen hebben op hun leerproces. Volgens Stobart (2008) moet aan drie condities worden voldaan, wil feedback effectief kunnen zijn:

- de leerling heeft de feedback nodig;
- de leerling ontvangt de feedback en heeft tijd om het te gebruiken;
- de leerling wil en is in staat om de feedback te gebruiken.

De eerste voorwaarde hangt samen met de kenmerken van de toets en de huidige kennis en vaardigheden van leerlingen in relatie tot het doel. Als de leerling het leerdoel al bereikt heeft, is er geen behoefte aan feedback en kan het ook weinig effect hebben op het behalen van het gestelde doel. De tweede voorwaarde betekent dat leerlingen de kans moeten krijgen om hun leerproces bij te sturen en om vervolgens in nieuwe toetsmomenten aan te tonen dat zij voortgang maken. Aan de laatste conditie kan voldaan worden door een hoge motivatie van leerlingen om te leren. Dit kan verder gefaciliteerd worden door het ter beschikking stellen van benodigde hulpmiddelen, zoals remediërend lesmateriaal (Stobart, 2008).

4.4 VOORWAARDEN CONTEXT

Formatief toetsen vindt niet plaats in isolatie, maar is idealiter een geïntegreerd onderdeel van de lespraktijk. Contextfactoren kunnen van invloed zijn op de manier waarop formatieve toetsing ingezet wordt binnen de school en of dit al dan niet leidt tot betere leerprestaties. De schoolleider heeft bijvoorbeeld invloed op de manier waarop formatieve toetsing ingezet wordt. Een belangrijke rol voor schoolleiders is het aanmoedigen, faciliteren (bijvoorbeeld door het beschikbaar stellen van tijd en middelen) en motiveren van docenten voor het formatief gebruiken van toetsen. Ook moeten er duidelijke doelen opgesteld worden (zie ook voorwaarden m.b.t. de toets). Verder blijkt ondersteuning bij het gebruik van formatief toetsen essentieel, zowel intern (ondersteuning in de school bij het analyseren en gebruiken van toetsen) als extern (bijvoorbeeld training in toetsgebruik) (Mandinach & Honey, 2008). Ook contextfactoren buiten een school zijn van invloed op de toepassing van formatieve toetsing. Landelijk beleid op het gebied van toetsgebruik en het afleggen van verantwoording kunnen bijvoorbeeld van grote invloed zijn op de toetscultuur in de school. Ook is de mate waarin het onderwerp formatief toetsen behandeld wordt op lerarenopleidingen bepalend voor de kennis en vaardigheden van docenten (Popham, 2009).

5 RESULTATEN LITERATUUR REVIEW

In dit hoofdstuk worden de resultaten van de literatuur review gepresenteerd. Factoren waarvoor binnen tenminste twee studies aanwijzingen zijn gevonden worden besproken. De overige factoren (waarvoor slechts één studie bewijs levert) worden weergegeven in tabellen maar niet besproken in de tekst. Het is belangrijk om te benadrukken dat als voor een bepaalde factor geen bewijs is gevonden in deze review dit niet hoeft te betekenen dat deze factor geen rol speelt. Naar sommige factoren is simpelweg geen of onvoldoende onderzoek gedaan. Factoren die van belang zijn gevonden binnen een bepaalde onderwijssector zijn mogelijk ook relevant voor andere onderwijssectoren.

In de tabellen wordt onderscheid gemaakt tussen PO, VO en MBO en de drie benaderingen naar formatief toetsen; OGW, AfL en DT in. De nummers in de tabellen verwijzen naar de betreffende referenties in de referentielijst. Aan het eind van de paragraaf worden de meest opvallende verschillen tussen de verschillende benaderingen en sectoren belicht. Er zijn relatief weinig relevante artikelen gevonden over DT. Ook zijn er weinig artikelen van voldoende kwaliteit gevonden over formatief toetsen in het MBO. Meer informatie over het zoekproces is te vinden in Bijlage 1.

5.1 VOORWAARDEN TOETS

De kenmerken van toetsen, systemen en processen zijn cruciaal voor formatief toetsen. In de review zijn 46 studies gevonden die inzicht geven in de voorwaarden van de toets. De resultaten van de zoektocht zijn weergegeven in Tabel 2. Een aantal van deze studies gingen over zowel het PO als het VO en/of over zowel OGW als Afl. De gevonden voorwaarden kunnen ingedeeld worden in vijf categorieën; (1) de kenmerken van formatieve toetsen, (2) de kenmerken van data systemen, data en feedback, (3) de kenmerken van digitale toetssystemen, (4) de kenmerken van (digitale) formatieve toetsing en (5) de kenmerken van (geautomatiseerde) feedback aan leerlingen.

TABEL 2 ZOEKRESULTATEN ARTIKELEN VOORWAARDEN TOETS

	Totaal	PO	VO	MBO	OGW	AfL	DT
Aantal studies	46	20	29	3	25	29	4

6.1.1 Voorwaarden toets: Kenmerken van formatieve toetsen

De kenmerken van de toetsen en daarmee de informatie die met de toetsen over de leerprestaties van de leerlingen wordt verkregen is van belang bij alle drie de vormen van formatief toetsen. Veel gevonden studies gingen over een mix van OGW en AFL. Deze studies benadrukten een aantal algemene kwaliteitskenmerken van toetsen, zoals het aansluiten van de inhoud bij het curriculum [16, 17], het gebruik van specifieke items voor diagnose van misvattingen [11, 17, 32, 69], en het gebruik van zinvolle items (toetsvragen) [13, 41] die passen bij het niveau van de leerlingen [13, 16, 71] en bij nationaal vastgestelde standaarden [13].

In drie studies over OGW en Afl wordt ingegaan op de noodzaak van de afstemming van de inhoud van de toets en de data die de toets oplevert over de leerresultaten op de (nationaal) vastgestelde standaarden [13, 30, 38]. In studies over AfL wordt ingegaan op de noodzaak om de inhoud van de toets af te stemmen op de doelen van de school [41], om ijkpunten vast te leggen voor de klassen [38]. Verder wordt benadrukt dat toetsen te analyseren moeten zijn en feedback moeten geven op het niveau van de individuele leerling [38]. Ook de aansluiting tussen de toetstaak en de beoogde leeropbrengst wordt genoemd als voorwaarde [17,71].

TABEL 3 KENMERKEN VAN FORMATIEVE TOETSEN

	PO	VO	MBO	OGW	AfL	DT
Afstemming van de toets en de items op:						
Het curriculum [16, 17]	x	x		x	x	x
Het niveau van de leerlingen [13, 16, 71]	x		x	x	x	
(Nationale) standaarden [13, 38] voor alle jaren [30]	x	x		x	x	
Leerdoelen (van scholen) [41] en summatieve toetsen [72]		x	x		x	
De beoogde leeropbrengsten [17, 71]		x	x		x	x
Toetsen en items moeten:						
Items gebruiken uit itembanken aangepast aan de context van scholen [11]		x		x	x	
Misvattingen opsporen [11, 17, 32, 69]	x	x		x	x	x
Betekenisvol [13, 41], interessant en motiverend [34] zijn	x	x		x	x	
Creativiteit aanmoedigen en boeiend zijn [13]	x			x	x	
Meerdere cognitieve conflicten bevatten [69]	x				x	x
Wat taalgebruik betreft afgestemd zijn op de leerlingen [11]		x		x	x	
Afgestemd zijn op de ontwikkelingsbehoefte [11]		x		x	x	
Passen bij de leerbehoefte [32]		x				x
Logisch in de lessen, het coachen en de beroepspraktijk passen [34, 71]		x	x	x	x	
Gericht zijn op hogere orde vaardigheden [34]		x		x	x	
Intuïtieve antwoorden van leerlingen ontlokken [20]		x			x	
Eenvoudig te analyseren zijn, directe feedback geven en informatie verschaffen op het niveau van de individuele leerling [38]		x			x	
Toetsen die gebruik maken van rubrics kunnen het leren bevorderen [44]		x			x	
Gericht zijn op leren en lesgeven in plaats van op beoordelen [72]			x		x	
De vraagstelling:						
Kan gebruikt worden om de leerling op het goede niveau te bevragen [17]		x			x	x
Vraagstellingen waarbij leerlingen verklaringen moeten opschrijven geven informatie over het niveau waarop ze de stof beheersen [17]		x			x	x
Open vraagstellingen geven ruimte voor leerlingen om te laten zien wat ze denken, wat een rijke basis geeft voor feedback [17]		x			x	x
Het gebruik van verschillende soorten vraagstellingen geeft de leerlingen meerdere mogelijkheden om te laten zien wat ze denken en kunnen [17]		x			x	x
Diagnostische toetsen:						
Diagnostische toetsen zijn nodig om diagnostische gevolgtrekkingen te doen over het leerniveau (klassikale toetsen voldoen niet) [35]		x				x

6.1.2 Voorwaarden toets: Data systemen, data en feedback

Tabel 4 geeft een overzicht van de kenmerken van data systemen (bijvoorbeeld informatie management systemen, leerlingvolgsystemen), analyse en gebruik. Deze studies gaan over OGW, AFL of een combinatie van beiden. Het blijkt dat de toegankelijkheid van de data een belangrijk kenmerk is [25, 29, 30, 31, 58, 60, 61], waarbij de data afgestemd moeten zijn op wat de gebruiker nodig heeft [55, 56, 64] en van hoge kwaliteit is (de data moeten bijvoorbeeld valide en betrouwbaar zijn) [29, 30, 31, 61]. Een ander belangrijk kenmerk is dat het data systeem goed moet functioneren [34, 66], dat de voortgang van de leerling gemonitord kan worden aan de hand van de behaalde toetsresultaten [10, 61]. Deze data moeten zowel op klas- als op individueel niveau geanalyseerd kunnen worden [30, 31] en gecombineerd kunnen worden [31] terwijl ze uit verschillende systemen komen [61]. Bovendien wordt in deze studies de noodzaak benadrukt om meerdere informatiebronnen te gebruiken om de leervoortgang te beoordelen, waaronder informele bronnen [7, 30, 38, 61].

TABEL 4 KENMERKEN VAN DATA SYSTEMEN, DATA EN FEEDBACK

	PO	VO	MBO	OGW	AfL	DT
Data systemen en software moeten:						
Goed toegankelijke zijn (voor de individuele leraar) [25, 29, 30, 31, 58, 60, 61]	x	x		x		
Goed functioneren [34, 66]	x	x		x	x	
Flexibel zijn, bijv. ten aanzien van de input [30, 34]		x		x	x	
De mogelijkheid bieden om de gegevens uit te splitsen op klas en op individueel niveau [30, 31], en ook op individueel niveau kunnen rapporteren [25]	x	x		x		
De mogelijkheid bieden om verschillende soorten data te combineren [31], zelfs als deze data uit verschillende systemen komen[61]	x	x		x		
De mogelijkheid bieden om op eenvoudige wijze data in te voeren [56]	x			x		
De output uit een datasysteem moet uitgebreider zijn dan enkel een grafiek [21] en begrijpelijk zijn [11]	x			x	x	
De mogelijkheid bieden om afwijkende leerprestaties in een vroeg stadium te signaleren [30]		x		x		
Het uitvoeren van een item analyse ondersteunen [30]		x		x		
De voortgang van leerlingen kunnen monitoren op basis van toetsresultaten [10, 61]	x	x		x		
Data:						
Moeten van hoge kwaliteit zijn (bijv. (ervaren) validiteit, accuraatheid, en tijdigheid) [29, 30, 31, 61]	x	x		x		
Moeten aansluiten bij de behoeften van de gebruiker [55, 56, 64]	x			x		
Overdraagbaar zijn van het PO naar het VO [31]		x		x		
Als er te veel data beschikbaar zijn weten docenten niet waar ze moeten beginnen [58]		x		x		
Feedback:						
Rapportages moeten duidelijk te interpreteren en op tijd voor handen zijn [4]	x			x		
Rapportages moeten informatie geven op het niveau van de hele school, voor de verschillende vakken en over de verschillende vaardigheden die leerlingen moeten beheersen [4]	x			x		
De resultaten van tussentoetsen kunnen op de volgende manier gebruikt worden: breng de gegevens in kaart, interpreteer de gegevens, identificeer problemen, probeer oplossingen uit in de klas en pas de interventies aan op basis van nieuwe gegevens [3]	x			x		
Rapporten genereren die makkelijk en eenduidig te interpreteren zijn, een complete overzicht geven en bruikbare informatie geven voor de docenten [64]	x			x		

6.1.3 Voorwaarden toets: Digitale toetssystemen

Tabel 5 geeft een overzicht van kenmerken die van belang zijn bij digitale toetssystemen. In tegenstelling tot de kenmerken in Tabel 3, betreft deze tabel kenmerken van digitale toets omgevingen/hulpmiddelen die gebruikt wordt bij het toetsen van leerlingen. Zulke tools worden op het moment hoofdzakelijk voor AfL gebruikt, omdat ze korte feedback loops mogelijk maken en geautomatiseerde feedback kunnen geven. Binnen DT bieden toetssystemen mogelijkheden om preciezer en efficiënter te kunnen meten. Echter, voor geen van de kenmerken van digitale toetssystemen is bewijs gevonden in meer dan één studie.

TABEL 5 KENMERKEN VAN DIGITALE TOETSSYSTEMEN

	PO	VO	MBO	OGW	AfL	DT
Integreer technologie in het onderwijs [11,34]		x		x	x	
Geregelde meerkeuze toetsen met onmiddellijke feedback [49]			x	x	x	
Een systeem moet uitleg bevatten over hoe je het systeem en de feedback kunt gebruiken, en voorbeelden van gebruik laten zien [50]	x	x		x	x	
Toetssystemen moeten docenten de mogelijkheid geven om de inhoud van de item bank aan te vullen of aan te passen [32]		x				x
Het gebruik van multimedia kan de authenticiteit van items verhogen [32]		x				X
De prestaties van de leerling kunnen verbeteren als de inhoud van de toetsing aangepast wordt op het niveau van de leerling [51]		x				x
CRSs geven feedback aan de docent over wat de leerling begrijpt en denkt [12]		x			x	
Bij het gebruik van CRSs, moet rekening gehouden worden met een aantal beperkingen ten aanzien van validiteit [12]		x			x	
Het systeem moet mogelijkheden voor remediering bieden [32]		x				x
Een systeem moet geautomatiseerd rapporteren opstellen voor leerlingen, ouders en docenten [32]		x				x

6.1.4 Voorwaarden toets: Het proces van (digitale) formatieve toetsing

Tabel 6 geeft een overzicht van de relevante kenmerken van het proces van (digitale) formatieve toetsing. Geen van de gevonden kenmerken kwam naar voren in meer dan één studie. Specifiek voor OGW wordt aanbevolen om ongeveer eens in de zes weken te toetsen [30]. Dit gaat nadrukkelijk niet om summatieve maar om formatieve toetsing [7]. Specifiek voor AfL wordt aanbevolen om expliciete leerdoelen te formuleren, om het doel van de toetsen uit te leggen en om de behoeften van individuele leerlingen in kaart te brengen en daarop in te spelen [38]. Bovendien wordt het samenwerken van leerlingen aangemoedigd, en wordt aanbevolen hen eigenaar te laten zijn van hun toetsproces [38].

TABEL 6 KENMERKEN VAN HET PROCES VAN (DIGITALE) FORMATIEVE TOETSING

	PO	VO	MBO	OGW	AfL	DT
Gebruik niet te vaak formatieve toetsen [7]		x		x		
Regelmatig (e.g. om de zes weken) toetsen om te kijken wat er geleerd is [30]		x		x		
Beloon leerlingen formeel voor tussentijds behaalde resultaten (bv. certificaat) [71]			x		x	
Regelmatig (wekelijks) korte formatieve toetsen (10-15 meerkeuze items) [49]			x	x	x	
Gebruik formatief toetsen als een proces van constante verbetering gebaseerd op eerdere prestaties en uitgevoerde opdrachten [71]			x		x	
Het vaak afnemen van formatieve toetsen verhoogt de leeropbrengst [49]			x	x	x	
Formatief toetsen moet goed opgebouwd zijn en leerprestaties monitoren [46]		x			x	
AfL bij schrijven vergt aanpassingen in zowel het toets- als instructieproces [33]		x			x	
Op basis van toetsen gedifferentieerde instructie geven [25, 47]	x	x		x	x	
Leg het doel van de oefeningen uit, expliciteer verwachtingen en beoogde leeropbrengsten [38]		x			x	
Leerlingen moeten genoeg tijd hebben om de vragen te beantwoorden [11]		x		x	x	
Gebruik discussies in de klas om te reflecteren op formatieve toetsen [34]		x		x	x	
Houd CRS sessies kort om te voorkomen dat de aandacht afdwaalt [28]		x		x	x	
Test de CRS apparatuur van te voren [28]		x		x	x	
Gebruik CRSs alleen voor formatieve doeleinden [28]		x		x	x	

	PO	VO	MBO	OGW	AfL	DT
Gebruik meerkeuzetoetsen in combinatie met directe feedback [49]			x	x	x	
Observeer en luister om te begrijpen wat de leerbehoeften van de leerling zijn en reageer daar op [38]		x			x	
Faciliteer meerdere vormen en methoden van leren [38]		x			x	

6.1.5 Voorwaarden toets: (Geautomatiseerde) Feedback

Tabel 7 geeft een overzicht van kenmerken van zinvolle (geautomatiseerde) feedback. Meerdere studies wijzen op het belang van het geven van feedback en de invloed die dit heeft op de motivatie om te leren [19, 22, 51]. Het wordt aangeraden om feedback te geven door de prestaties van de leerlingen te koppelen aan (van te voren bekende) criteria [72] om er voor te zorgen dat leerlingen zich op de juiste manier focussen op hun werk [19, 24, 41]. De feedback moet eerlijk, gedetailleerd en beknopt zijn, moet aangeven wat er fout gaat en hoe de leerling zich kan verbeteren [1, 19, 22, 48, 71]. De resultaten met betrekking tot feedback zijn beperkt in dit onderzoek, omdat er niet specifiek gezocht is naar literatuur over feedback. Voor een overzicht van effectieve feedback in computergestuurde toetsen wordt de geïnteresseerde lezer doorverwezen naar http://www.toetswijzer.nl/html/feedback/default.shtm.

TABEL 7 KENMERKEN VAN (GEAUTOMATISEERDE) FEEDBACK

	PO	VO	MBO	OGW	AfL	DT
Feedback moet een eerlijk, gedetailleerd en beknopt advies zijn en een duidelijke vergelijking bevatten van te voren bekende criteria [1, 19, 22, 24, 41, 48, 71, 72]	x	x	x	x	x	
Feedback voor zwakkere leerlingen moet duidelijke aanwijzingen voor verbetering bevatten [21]	x			x	x	
Geef ook cijfers als onderdeel van feedback zodat leerlingen weten waar ze staan [48]		x		x	x	
Onnodige feedback of feedback op het verkeerde niveau werkt demotiverend/frustrerend voor leerlingen [22]	x				x	
Leerlingen zijn van mening dat docenten duidelijk aan moeten geven dat de rol van feedback is dat leerlingen ervan leren, in plaats van dat het gegeven wordt om het leren te beoordelen [24]	x	x			x	
Geef leerlingen de tijd om feedback te verwerken [19]		x			x	
Geef feedback aan leerlingen gericht op het verbeteren van hun vaardigheden om problemen op te lossen [16]	x			x	x	
Geef leerlingen feedback op hun leerstrategieën [16]	x			x	x	
Feedback moet meer gericht zijn op het proces dan op de prestatie [44]		x			x	
Als rubrics gebruikt worden en de feedback is gericht op het proces, dan verhoogt dit de zelfredzaamheid [44]		x			x	
Discussies kunnen gebruikt worden om leerlingen feedback te geven op de kwaliteit, bruikbaarheid en uitvoerbaarheid van hun ontwerpen d.m.v. het stellen van gerichte vragen over het denkproces van leerlingen [14]	x				x	
Feedback moet er op gericht zijn om toekomstige problemen zelfstandig op te lossen [9]	x	x			x	
Feedback vanuit een CRS moet vergezeld gaan met feedback van een docent [12]		x			x	
De kwaliteit van de vraagstelling van de docent is bepalend voor de kwaliteit van de feedback naar leerlingen [17]		x			x	x
Genereer evaluerende en beschrijvende feedback voor leerlingen [9]	x	x			x	
Time feedback zorgvuldig en integreer het in de les [26]		x			x	

6.1.6 Opvallende verschillen: Voorwaarden toets

De gevonden voorwaarden voor DT overlappen tot op zekere hoogte met die van OGW en AfL. Specifiek voor DT is de noodzaak om leerbehoefte in kaart te brengen [32] en om specifieke diagnostische conclusies te kunnen trekken over het leerproces [35]. Verder valt op dat de meeste kenmerken in Tabel 4 van toepassing zijn op OGW. Dit is niet verrassend gezien de centrale plaats die data systemen en data-analyse hebben bij deze vorm van formatief toetsen.

Verder gingen de meeste studies (met name in Tabel 4) over het VO, maar veel van de resultaten zijn waarschijnlijk ook van toepassing in het PO en MBO. Ook voor het PO en MBO kan bijvoorbeeld beargumenteerd worden dat het belangrijk is om verschillende soorten data te combineren, dat data van goede kwaliteit zijn en dat feedback niet alleen gegeven wordt op het niveau van het resultaat, maar ook op het niveau van het proces. Het belang van feedback wordt bij formatief toetsen binnen alle onderwijscontexten benoemd. Er waren meer studies (vijf) die wezen op het belang van eerlijke, gedetailleerde en beknopte feedback bij Afl, dan bij OGW (één studie) of DT (geen studies gevonden). Echter, feedback speelt een cruciale rol bij formatieve toetsing in alle onderwijssectoren, of het nu om OGW, Afl of DT gaat.

5.2 VOORWAARDEN DOCENT

In Tabellen 9, 10 en 11 worden alle gevonden factoren beschreven met betrekking tot de rol van de docent, die het gebruik van formatief toetsen in de klas beïnvloeden op basis van 52 studies (zie Tabel 8). De gevonden voorwaarden kunnen ingedeeld worden in drie categorieën; (1) kennis, vaardigheden en attitude, (2) samenwerken en het stellen van doelen en (3) het betrekken van leerlingen.

TABEL 8 ZOEKRESULTATEN ARTIKELEN VOORWAARDEN DOCENT

	Totaal	PO	VO	MBO	OGW	AfL	DT
Aantal studies	52	33	28	1	33	24	5

6.2.1 Voorwaarden docent: Kennis, vaardigheden en attitude

Veel studies hebben aangetoond dat kennis op het gebied van data-analyse en assessment cruciaal is voor het gebruik van formatief toetsen in de klas (zie Tabel 9). Docenten moeten deze kennis en vaardigheden bezitten om verschillende soorten gegevens te verzamelen, te analyseren en te interpreteren, en om maatregelen te kunnen treffen op basis van deze gegevens [2, 3, 4, 6, 7, 9, 11, 13, 16, 20, 25, 28, 33, 34, 35, 39, 58, 59, 61, 62, 64, 65, 70]. Het is belangrijk dat docenten formatief toetsen op een conceptueel niveau begrijpen en daarbij onderscheid kunnen maken tussen verschillende manieren van toetsen [11, 13]. Daarnaast moeten ze in staat zijn om toetsresultaten in instructie te integreren [39]. Drie studies [28, 34, 72] benadrukken dat docenten in staat moeten zijn zelf kwalitatieve toetsen te ontwikkelen.

Daarnaast is een positieve houding ten aanzien van formatief toetsen belangrijk [7, 9, 11, 30, 47, 51, 54, 56, 60, 61, 63, 64, 65, 74]. Overtuiging van de noodzaak om data te gebruiken voor het verbeteren van onderwijs en leerprestaties is hierbij van belang. Verder wordt intrinsieke motivatie en de bereidheid om kritisch te reflecteren op zowel leerresultaten als het eigen functioneren [9, 65] genoemd. Docenten zouden zich verantwoordelijk moeten voelen voor de resultaten [55], eigenaarschap moeten voelen ten opzichte van het toetsproces [74] en moeten geloven dat zij het verschil kunnen maken [60].

Echter, kennis van assessment en van data analyse en een positieve houding tegenover data gebruik zijn niet voldoende. Docenten hebben ook vakdidactische kennis nodig om onderwijs aan te kunnen passen op basis van toetsresultaten [7, 13, 14, 16, 20, 40, 69]. Docenten moeten beschikken over kennis en vaardigheden om leerlingen te begeleiden bij hun leren, docenten moeten bekend zijn met de leerstrategieën die leerlingen gebruiken en moeten leerlingen zinvolle feedback kunnen geven (zie ook

Tabel 7) die past bij verschillende leerstrategieën [16]. Verder moeten docenten in staat zijn om misvattingen van leerlingen te herkennen en op te lossen. Dit laatste is met name belangrijk voor DT [69].

Daarnaast speelt technologie op het gebied van informatiesystemen en digitale (adaptieve) toetsen een belangrijke rol bij formatief toetsen. Daarom zijn ook kennis, vaardigheden en een positieve houding ten opzichte van technologiegebruik in de klas van belang [7, 11, 32, 34, 47, 74].

TABEL 9 KENMERKEN VAN KENNIS, VAARDIGHEDEN EN ATTITUDE VAN DE DOCENT

	PO	VO	MBO	OGW	AfL	DT
Kennis en begrip van data en assessment [2, 3, 4, 6, 7, 9, 11, 13, 16, 20, 25, 28, 33, 34, 35, 39, 58, 59, 61, 64, 65, 70]	x	x		x	x	x
Positieve houding t.o.v. toetsdata gebruik [7, 9, 11, 30, 47, 51, 56, 60, 61, 63, 64, 65, 74]	x	x		x	x	
Vakdidactische kennis [1, 7, 13, 16, 40, 69]	x	x		x	x	x
Kennis, vaardigheden en een positieve attitude met betrekking tot technologie [7, 11, 32, 34, 47, 74]	x	x		x	x	x
Vaardigheden op het gebied van het schrijven van toetsvragen [28, 34]		x		x	x	
Het vermogen kritisch te reflecteren op data [65]	x	x		x		
Pas het onderwijs aan, maak gebruik van korte feedback cycli [7]		x		x		
Docenten moeten toegerust zijn om goede feedback te geven [20]		x			x	
Gebruik rubrics om aan te geven wat leerlingen moeten beheersen en om doelen te stellen [15]	x	x			x	
Benoem voorkennis [20]		x			x	
Het vermogen om de juiste vragen op het juiste moment te stellen [17]		x			x	x

6.2.2 Voorwaarden docent: Samenwerken en het stellen van doelen

Meerdere studies wijzen op het belang van samenwerking en onderling vertrouwen tussen docenten als het over het gebruik van formatief toetsen in de klas gaat (zie
Tabel 10). Het is belangrijk dat formatief toetsen niet geïsoleerd gebruikt wordt, maar dat docenten, administratief personeel en schoolleiding samenwerken bij het gebruik van toetsgegevens [2, 8, 10, 15, 16, 25, 28, 29, 31, 33, 36, 40, 45, 55, 56, 58, 59, 60, 61, 62, 66, 70, 74]. Een voorbeeld hiervan zijn professionele leergemeenschappen of datateams. Deze groepen zijn bedoeld om toetsresulaten te bespreken, informatie uit te wisselen, het werk van leerlingen te delen en gezamenlijk lesplannen te ontwikkelen. Uit verscheidene studies bleek dat docenten elkaar kunnen helpen en van elkaar kunnen leren, zowel binnen als tussen scholen [15, 45].

Het is belangrijk dat het gebruik van formatieve toetsen in de klas niet beperkt blijft tot een op zichzelf staande analyse van toetsgegevens, maar dat van te voren afgesproken wordt wat het doel van formatief toetsen is en welke vragen beantwoord moeten worden [1, 27, 31, 55, 58]. Docenten (maar ook schoolleiders en leerlingen) moeten gezamenlijk afspreken wat de doelen zijn op het niveau van de school, klas en individuele leerling. De toetsgegevens moeten ook voor die doelen gebruikt worden.

TABEL 10 – KENMERKEN VAN SAMENWERKEN EN DOELEN STELLEN

	PO	VO	MBO	OGW	AfL	DT
Samenwerken met andere docenten of met schoolleiders [2, 8, 10, 16, 25, 28, 29, 31, 33, 36, 40, 45, 55, 56, 58, 59, 60, 61, 62, 66, 70, 74]	x	x		x	x	
Definieer bij het begin van datagebruik een duidelijke set onderzoeks-vragen/doelen [27, 31, 55, 58]	x	x		x		

Netwerk met andere scholen [15, 45]	x	x			x	
Samenwerken met leerlingen [36]		x		x		
Vertrouwen onder docenten [2, 8]	x			x		
Beslissingen omtrent formatieve toetspraktijken moeten door alle betrokken docenten genomen kunnen worden (in plaats van een groep docenten die verantwoordelijk is voor toetsing en examinering) [72]			x		x	

6.2.3 Voorwaarden docent: Het betrekken van leerlingen

Een aantal studies gaat in op de vraag hoe je leerlingen kunt betrekken bij formatief toetsen. Tabel 11 geeft een overzicht van de resultaten. Omdat het betrekken van leerlingen vraagt om specifieke aandacht van de docent, is deze factor meegenomen in dit hoofdstuk. Docenten kunnen leerlingen op verschillende manieren bij formatief toetsen betrekken. Een aantal studies gaan in op het gebruik van *peer- of self-assessment* [5, 9, 19, 23]. Bij peer-assessment is het van belang dat docenten sturing geven aan het proces. De docent kan een faciliterende rol spelen [3, 5, 19, 23] en bepaalde feedbackmomenten organiseren zodat alle leerlingen ervan profiteren [19]. In andere studies worden leerlingen bij formatief toetsen betrokken door hen de doelen te laten stellen en te laten bepalen hoe ze die doelen kunnen bereiken [9,13].

TABEL 11 - KENMERKEN VAN HET BETREKKEN VAN LEERLINGEN

	PO	**VO**	**MBO**	**OGW**	**AfL**	**DT**
Ondersteun het geven van peer-feedback [3, 5, 19, 23] en organiseer de peer-feedback zodanig dat alle leerlingen ervan profiteren [19]	x	x		x	x	
Organiseer een klassengesprek over de toetsresultaten [15, 34] en houd in de gaten wanneer dit gesprek afdwaalt van het doel [38]	x	x		x	x	
Betrek leerlingen door hen bepaalde verantwoordelijkheden te geven [9, 13]	x	x		x	x	
Geef leerlingen de ruimte om hun planning en beslissingen te bespreken [14]	x				x	
Laat leerlingen fouten maken en ervan leren [9]	x	x			x	
Gebruik self-assessment [9]	x	x			x	
Bespreek hun leergedrag en hun leeromgeving met leerlingen zodat ze in staat zijn tot zelfstandig leren [22]	x				x	
Instrueer leerlingen over hoe ze feedback moeten geven aan hun medeleerlingen [5]	x				x	
Geef duidelijke criteria voor peer- of self-assessment [23]	x	x			x	
Leg uit wat de meerwaarde is van peer- of self-assessment [23]	x	x			x	

6.2.4 Opvallende verschillen: voorwaarden docent

De gevonden factoren uit dit onderzoek zijn van belang voor OGW en voor AfL (hoewel de factoren uitgebreider bestudeerd waren voor AfL). De factoren met betrekking tot samenwerken en het stellen van doelen zijn gevonden in het PO, VO en MBO. Samenwerking bleek vooral van belang voor het gebruik van OGW (16 studies OGW, 3 studies Afl). Het formuleren van duidelijke doelen en onderzoeksvragen werd ook vaker voor OGW gerapporteerd (vier keer) dan voor Afl (één keer) of DT (niet). Het betrekken van leerlingen bij formatief toetsen is van belang voor zowel het PO als VO. Het betrekken van leerlingen werd het meest gerapporteerd bij AfL, nauwelijks bij studies over OGW en niet in de DT studies.

5.3 VOORWAARDEN LEERLING

Achttien papers rapporteerden resultaten over de rol van leerlingen bij formatief toetsen. (zie Tabel 12). Deze papers kunnen ingedeeld worden in twee categorieën: (1) de kennis, vaardigheden en houding van de leerlingen en (2) de rol van leerlingen in het proces van formatieve toetsing.

TABEL 12 ZOEKRESULTATEN ARTIKELEN VOORWAARDEN LEERLINGEN

	Totaal	PO	VO	MBO	OGW	AfL	DT
Aantal studies	18	7	11	2	5	17	1

6.3.1 Voorwaarden leerling: Kennis, vaardigheden en houding

Ten aanzien van kennis, vaardigheden en houding van leerlingen (Tabel 13), benadrukken drie studies het belang van vaardigheden gerelateerd aan peer-assessment. Leerlingen moeten in staat zijn om het werk van hun klasgenoten te beoordelen, goede feedback te geven en verbeterpunten aan te geven [5, 41, 19]. Dit vereist oefening [41], maar het blijkt dat getrainde leerlingen in staat zijn om zelfstandig het werk van hun studiegenoten te beoordelen en veel meer kunnen dan het geven van oppervlakkige feedback. Verder benadrukken de resultaten het belang van self-assessment; leerlingen moeten in staat zijn om hun eigen werk te beoordelen als de docent daarvoor de richtlijnen aanreikt [46, 48].

TABEL 13 KENMERKEN VAN KENNIS, VAARDIGHEDEN EN HOUDING VAN DE LEERLING

	PO	VO	MBO	OGW	AfL	DT
Leerlingen moeten beschikken over vaardigheden ten aanzien van het geven van goede feedback aan hun medeleerlingen [5, 19, 41]	x	x			x	
Leerlingen moeten in staat zijn om hun eigen werk te beoordelen zodat ze beter met hun leertaken om kunnen gaan [46, 48]		x		x	x	
Leerlingen moeten in staat zijn om te reflecteren en hun leerstrategie aan te passen [22]	x				x	
Leerlingen hebben een houding nodig waarin ze serieus invulling geven aan formatief toetsen [32]		x				x
Leerlingen moeten de toetstaak in formatief toetsen als gepast en nuttig ervaren [71]			x		x	

6.3.2 Voorwaarden leerling: Rol bij het proces

Vijf studies benadrukken dat leerlingen een actieve rol moeten spelen in het proces van formatief toetsen [5, 9, 13, 11, 26, 28, 43, 26, 9]. Uit deze studies blijkt dat een actieve rol van leerlingen hun betrokkenheid en motivatie stimuleert [43], wat leidt tot betere leerresultaten [13].

TABEL 14 KENMERKEN VAN LEERLINGEN BIJ HET PROCES VAN FORMATIEF TOETSEN

	PO	VO	MBO	OGW	AfL	DT
Door leerlingen actief te betrekken bij het proces nemen betrokkenheid en motivatie toe [5, 9, 11, 13, 26, 28 43]	x	x		x	x	
Door leerlingen hun eigen werk te beoordelen nemen de motivatie en de leerresultaten toe [13, 43]	x			x	x	
De motivatie van leerlingen neemt toe als ze authentieke opdrachten krijgen [14]	x				x	
Docenten, ouders en leerlingen moeten met elkaar in gesprek gaan over de leerresultaten omdat leerlingen daar veel van leren [15]	x	x			x	
Leerlingen moeten niet afgeleid of onvoldoende toegerust worden omdat dit een negatieve invloed heeft op het gebruik van formatief toetsen in de klas [34]		x		x	x	
Naast de docent zouden leerlingen en actoren uit de beroepspraktijk ook verantwoordelijk moeten zijn voor de inhoud van formatieve toetsen [72]			x		x	

6.3.3 Opvallende verschillen: voorwaarde leerling

Bijna alle studies naar de rol van leerlingen bij formatief gaan over AfL. Dit past ook bij het niveau en proces van Afl: het vindt plaats als geïntegreerd onderdeel van de dagelijkse lespraktijk en benadrukt het belang

van de actieve deelname van de leerling in het leerproces. Tevens komt het merendeel van het bewijs uit het PO.

5.4 VOORWAARDEN CONTEXT

In 43 studies worden factoren beschreven die aangeven hoe de context van invloed is op het gebruik van formatief toetsen in de klas (zie Tabel 15). De gevonden factoren kunnen ingedeeld worden in vijf categorieën. De eerste vier categorieën vinden plaats binnen de school: (1) schoolkenmerken, (2) benodigde tijd en middelen, (3) ondersteuning van het professionaliseren van docenten en (4) schoolleiderschap. De vijfde categorie bevat factoren buiten de school.

TABEL 15 ZOEKRESULTATEN ARTIKELEN VOORWAARDEN CONTEXT

	Totaal	PO	VO	MBO	OGW	AfL	DT
Aantal studies	43	29	20	2	31	17	1

6.4.1 Context binnen de school: Kenmerken van de school

Een aantal kenmerken van de school blijken van belang voor het gebruik van formatieve toetsing in de klas (zie Tabel 16). In vijftien studies werd geconcludeerd dat een formatieve toetscultuur heel belangrijk is [2, 4, 7, 10, 15, 26, 30, 36, 38, 45, 54, 56, 59, 61, 62, 63]. Een formatieve toetscultuur wordt omschreven als gericht op samenwerking [26, 54] en continue verbetering [2, 7, 10, 30, 45, 56, 59, 61, 63] waarbij de docent uitgedaagd wordt om instructie strategieën aan te passen en waarbinnen ruimte is om te experimenteren met mogelijke verbeteringen [56].

Om tot een formatieve toetscultuur te komen is het nodig dat visie, normen, doelen en verwachtingen ten aanzien van het gebruik van toetsresultaten expliciet worden benoemd [7, 8, 10, 15, 31, 58, 60, 61, 73]. Bovendien wordt in verschillende studies [8, 31, 56, 66, 73] benadrukt dat het formatief gebruiken van toetsgegevens gezien moet worden als een gezamenlijke verantwoordelijkheid, terwijl andere studies juist wijzen op de autonomie van de individuele docent [29, 30, 54, 55, 59]. In deze studies [30, 54, 55, 59] wordt aangegeven dat de druk om bij te blijven ervoor zorgt dat docenten er meestal voor kiezen om het bestaande curriculum te volgen, in plaats van op basis van toetsresultaten de lesinhoud aan te passen aan de behoeften van leerlingen, omdat dit laatste er wel eens voor kan zorgen dat bepaalde onderdelen van de stof nog eens behandeld moeten worden. Docentautonomie kan bereikt worden door hen zeggenschap te geven over het toetsproces en hen de ruimte te geven om aanpassingen te maken aan het lesprogramma [30, 59].

TABEL 16 KENMERKEN VAN DE SCHOOL

	PO	VO	MBO	OGW	AfL	DT
Een cultuur van formatief toetsen [2, 4, 7, 10, 15, 26, 30, 36, 38, 45, 54, 56, 59, 61, 62, 63]	x	x		x	x	
Gedeelde visie, normen, doelen en verwachtingen ten aanzien van toetsdata gebruik [7, 8, 10, 15, 31, 58, 60, 61, 73]	x	x		x	x	
Het gebruik van toetsdata moet gezien worden als een gezamenlijke verantwoordelijkheid [8, 31, 56, 66, 73]	x	x		x	x	
Respecteer de autonomie en professionaliteit van docenten, geef hen zeggenschap over het toetsproces [29, 30, 54, 55, 59]	x	x		x	x	
Om het gebruik van toetsdata te stimuleren zouden beslissingen op verschillende niveaus in de school genomen moeten [45]	x	x		x		
Het curriculum zou niet alleen gericht moeten zijn op door de staat geformuleerde leerdoelen [7]		x		x		

6.4.2 Context binnen de school: Benodigde tijd en middelen

Tabel 17 laat de uitkomsten zien met betrekking tot de tijd en middelen die nodig zijn voor het succesvol gebruiken van formatief toetsen in de klas. In negen studies wordt de toegang tot vernieuwende methodes en hulpmiddelen genoemd [4, 8, 11, 14, 30, 31, 32, 34, 43, 47, 74]. Voorbeelden zijn de toegang tot data-discussie protocollen voor ondersteuning van de groepsdiscussies over toetsdata gebruik [8], hulpmiddelen die ondersteuning bieden op de gebieden die als zwak zijn geïdentificeerd [4] en moderne videoapparatuur [43]. De beschikbaarheid over technologie voor het opslaan, analyseren en rapporteren over toetsresultaten wordt ook gezien als een randvoorwaarde [30, 31, 61]. Verder is het belangrijk dat er tijd en geld geïnvesteerd wordt. Het belang van het investeren van tijd komt naar voren in negen studies [7, 8, 10, 29, 31, 55, 56, 58, 66, 70]. Deze tijd kan door docenten gebruikt worden om gezamenlijk te werken aan het verbeteren van het onderwijs op basis van de uitkomsten van de analyse van toetsgegevens [7, 8, 10, 31, 55, 56, 66, 70], om ervaringen uit te wisselen, om plannen te bediscussiëren [7] en voor het toewijzen van middelen aan groepen docenten [10]. Geld is nodig [3, 11] voor bijvoorbeeld het investeren in professionalisering van docenten [3]. Ook moet het gebruik van technologie binnen het onderwijs financieel mogelijk gemaakt worden [11, 34].

TABEL 17 KENMERKEN VAN BENODIGDE TIJD EN MIDDELEN

	PO	VO	MBO	OGW	AfL	DT
Toegang tot vernieuwende methodes, hulpmiddelen en systemen [4, 8, 11, 14, 30, 31, 32, 34, 43, 47, 61, 74]	x	x		x	x	x
Tijd voor (samenwerking) aan het toepassen van FA in de praktijk [7, 8, 10, 29, 31, 34, 55, 56, 58, 66, 70]	x	x		x		
Budget [3, 11]	x	x		x	x	

6.4.3 Context binnen de school: Het mogelijk maken van docentprofessionalisering

In zestien studies wordt gewezen op het belang van het investeren in professionalisering van docenten [7, 15, 30, 31, 42, 47, 55, 58, 59, 61, 62, 63, 65, 70, 71]. Het merendeel van deze studies geeft aan dat deze professionalisering zich zou moeten richten op het gebruik van data voor formatief toetsen (bijvoorbeeld, het formuleren van onderzoeksvragen, het selecteren van indicatoren, het interpreteren van de resultaten, en het ontwikkelen van oplossingen) [7, 15, 30, 31, 34, 42, 47, 58, 59, 61, 63, 65, 70]. Daarnaast wordt het ondersteunen van verdere professionalisering op het gebied van didactiek [7, 15] en vakdidactische kennis [15] genoemd. Naast deze onderwerpen voor professionalisering wordt in acht studies gewezen op de noodzaak van ondersteuning van de docent op de schoollocatie (bijvoorbeeld door coaches of specialisten) [4, 10, 16, 21, 34, 59, 63, 70]. Deze ondersteuning kan docenten helpen bij het implementeren van hun verbeterplannen [16], bij het gebruik van methodes en modellen [59] en bij het aanpassen van hun

lesprogramma op basis van de uitkomsten van formatief toetsen [4, 21].

Het ondersteunen van professionalisering moet een continu proces zijn en ondersteuning moet beschikbaar zijn gedurende lange tijd [15, 45, 62]. Als een heel schoolteam getraind is, kunnen collega's elkaar helpen bij het toepassen van nieuwe kennis en vaardigheden [62] en elkaar stimuleren bij het starten van professionele leergemeenschappen [15]. In vier studies kwam naar voren dat professionalisering van docenten afgestemd moet worden op de context waarbinnen ze werken [2, 38, 54, 55, 62]. Schoolpersoneel moet leren werken met de data die binnen hun school beschikbaar zijn [62]. Bovendien moeten de aanbieders van docentprofessionaliseringstrajecten geloofwaardig zijn en beschikken over grondige kennis van de context waarin de docenten werken [38, 54].

TABLE 18 KENMERKEN VAN HET FACILITEREN VAN PROFESSIONALISERING

	PO	VO	MBO	OGW	AfL	DT
Mogelijkheden tot professionalisering [7, 15, 30, 31, 34, 42, 47, 55, 58, 59, 61, 62, 63, 65, 70, 71]	x	x	x	x	x	
Ondersteuning van docenten op de school [4, 10, 16, 21, 34, 59, 63, 70]	x	x		x	x	
Professionalisering moet afgestemd zijn op de context waarin de docenten werken [2, 38, 54, 55, 62]	x	x		x	x	
Professionalisering is een traject dat gedurende lange tijd moet lopen [15, 45, 62]	x	x		x	X?	
De voorbereiding van het schoolpersoneel is van groot belang [62]	x			x		

6.4.4 Context binnen de school: Leiderschap

In eenentwintig studies wordt aangegeven dat leiderschap een belangrijke rol speelt bij het gebruik van formatieve toetsen in de klas. Volgens de meeste studies is het belangrijk dat leiders aandacht geven aan het gebruik van toetsgegevens binnen school [3, 8, 29, 30, 31, 54, 55, 58, 59, 60, 61, 63, 65, 66, 70, 74]. De directeur wordt vaak genoemd als de persoon die een cultuur van datagebruik kan scheppen [65, 63]. Daarom is het van belang dat schoolleiders kennis en vaardigheden hebben m.b.t. datagebruik [3, 63, 66, 70].

Meerdere studies benoemen de verantwoordelijkheid van de schoolleiding voor het formuleren van verwachtingen, het stellen van duidelijke doelen gericht op school-brede implementatie van formatief toetsen en het gebruik van data [10, 15, 36, 56, 70, 72]. Schoolleiders moeten bijvoorbeeld visie vertalen naar heldere doelstellingen voor hun eigen school [10] die recht doen aan de specifieke omstandigheden van de school en die van de gemeenschap [36]. Daarnaast blijkt dat schoolleiders direct invloed kunnen uitoefenen door tijd in te plannen voor samenwerking, door te beslissen wie aan de samenwerking mee doet, door specialisten toe te wijzen aan groepen docenten en door de samenwerking te monitoren en te faciliteren [10, 36].

Omdat formatief toetsen voor veel scholen een relatief nieuw fenomeen is, is het nodig dat schoolleiders beschikken over de juiste kennis en vaardigheden om formatief toetsen te implementeren [3, 56, 59, 70]. Ze moeten bijvoorbeeld weten welke middelen er nodig zijn, hoe het organisatorisch vormgegeven moet worden en welke routines nodig zijn [3]. Om dit te kunnen bereiken zouden schoolleiders toegerust moeten worden om formatief toetsen school-breed te implementeren [15].

TABEL 19 KENMERKEN VAN LEIDERSCHAP

	PO	VO	MBO	OGW	AfL	DT
Schoolleiders zouden zich moeten richten op datagebruik binnen de scholen en dit moeten ondersteunen [3, 8, 29, 30, 31, 54, 55,	x	x		x	x	

58, 59, 60, 61, 63, 65, 66, 70, 74]						
Schoolleiders zouden verwachtingen, heldere doelen en normen moeten formuleren voor schoolbrede implementatie van formatief toetsen en gebruik van data binnen scholen [10, 15, 36, 56, 70, 72]	x	x	x	x	x	
Schoolleiders hebben vaardigheden, een juiste houding en middelen nodig om veranderingen in toetsbeleid te implementeren [3, 15, 25, 56, 59, 70]	x	x		x	x	
Schoolleiders moeten kennis en vaardigheden hebben m.b.t. datagebruik [3, 63, 66, 70]	x	x		x		
Schoolleiders moeten enthousiast zijn over het formatieve gebruik van data en dit over kunnen brengen op hun personeel [59, 63]	x	x		x		
Schoolleiders moeten betrokken zijn [59]	x	x		x		
Schoolleiders moeten een structuur aanbieden waarbinnen data gebruikt kan worden [63]	x			x		
Schoolleiders moeten de professionele ontwikkeling van hun docententeam monitoren [10]	x			x		

6.4.5 Context buiten de school

Veertien studies refereerden aan context factoren van buiten de school. In zeven studies werd de druk om verantwoording af te leggen over de leerresultaten en expliciete doelen voor de leerresultaten genoemd als een factor die van invloed was op het gebruik van formatief toetsen in de klas [16, 25, 27, 29, 30, 54, 63, 65, 74]. Als er geen druk van buiten af is om de prestaties te verbeteren wordt er minder gebruik gemaakt van (formele) data [65]. Het moeten afleggen van verantwoording over leerresultaten wordt genoemd als een factor die van invloed is op het verbeteren van de leerling prestaties [65, 25]. Een belangrijke kanttekening is [25, 27, 54, 63] dat teveel externe druk op het afleggen van verantwoording het gebruik van formatief toetsen in de klas juist belemmert. Verder is er beleid nodig over het gebruik van formatief toetsen [7, 10, 54, 61, 70].

TABEL 20 KENMERKEN VAN DE CONTEXT BUITEN DE SCHOOL

	PO	**VO**	**MBO**	**OGW**	**AfL**	**DT**
Druk om verantwoording af te leggen [16, 25, 27, 29, 30, 54, 63, 65, 74]	x	x		x	x	
Beleid ten aanzien van formatief toetsen [7, 10, 54, 61, 70]	x	x		x	x	
Aandacht voor assessment bij de opleiding voor docenten [6, 64]	x			x		
Nationaal raamwerk voor formatief toetsen [72]			x		x	

6.4.6 Opvallende verschillen: Voorwaarden context

Voor bijna alle factoren die van invloed blijken op de context is evenveel bewijs gevonden in het PO als het VO. Alleen ten aanzien van de factoren ‘visie, normen, doelen en verwachtingen voor data gebruik’ is meer bewijs gevonden in het VO en voor ‘extra ondersteuning van de docenten’ bij het ondersteunen van docentprofessionalisering is meer bewijs gevonden in het PO. Slechts twee studies noemen de effecten van context op het gebruik van formatief toetsen binnen het MBO. Verder moet opgemerkt worden dat de meeste studies die onderzoek deden naar de effecten van de context van formatief toetsen gingen over OGW. Dit betekent niet dat deze factoren niet relevant zijn voor Afl en DT, maar dat wij weinig literatuur gevonden hebben die de invloed van deze factoren op het gebruik van Afl en DT in de klas aantoont.

6 RESULTATEN FOCUSGROEPEN

Om een link te leggen met de praktijk zijn de gevonden resultaten uit de literatuurreview besproken in focusgroepen. Dit had drie doelen: (1) het verifiëren van de resultaten uit het literatuuronderzoek, (2) het aanvullen van de informatie uit de literatuur en (3) het vertalen van de resultaten uit het literatuuronderzoek naar de praktijk. De resultaten uit een eerste literatuurzoektocht zijn voorgelegd aan focusgroepen met docenten uit het VO, docenten uit het PO en experts uit zowel het PO, het VO en het MBO. Na een aanvullende literatuurzoektocht zijn de resultaten nogmaals besproken met docenten uit het PO, het VO en het MBO. Met behulp van tabellen wordt een link gelegd tussen de resultaten van de focusgroep en bevindingen uit de literatuur, met betrekking tot PO, VO en/of MBO. Meer informatie over het proces rondom de focusgroepen is te vinden in bijlage 1. Tabel 21 geeft een overzicht van de focusgroepen en het aantal deelnemers.

TABEL 21 OVERZICHT FOCUSGROEPEN: AANTAL RESPONDENTEN PER FOCUSGROEP

Focusgroep-ronde	PO	VO	MBO	Experts
1	2	6	-	4
2	2	6	4	-

6.1 FOCUSGROEPEN PO

De docenten van beide PO focusgroepen gaven aan dat zij de term ‘formatieve toetsing’ niet gebruiken in het basisonderwijs. De docenten zijn wel bekend met formatieve toetsing. Een deelneemster benoemde bijvoorbeeld: *“Ik haal gegevens uit een toets en dan pas ik mijn lessen daar wel op aan, want je ziet wel dat iets onvoldoende aan bod is geweest bijvoorbeeld. Maar formatief toetsen als doel hebben we niet.”* Daarnaast werd genoemd dat docenten in gesprek gaan met collega’s of leerlingen over toetsresultaten. De docenten gaven aan dat ze begrijpen dat er verschillende typen toetsen zijn welke gebruikt kunnen worden voor formatieve toetsing, zoals Cito toetsen, methodegebonden toetsen, observaties en vraaggesprekken. Na het benoemen van de drie typen formatieve toetsen gaven de docenten aan dat, in wisselende mate, gebruik wordt gemaakt van zowel OGW, AfL als DT.

De docenten van beide focusgroepen benoemden meerdere ervaringen met formatieve toetsing. Een docent beschreef bijvoorbeeld: *“Ik ga nog wel eens met een kind samen zitten en dan gewoon vragen: hoe pak je dit nou aan? Om de ontwikkeling van het kind te stimuleren.”* Ook werd door deelnemers genoemd dat ze de leerlingen opdelen in verschillende niveaugroepen, om zo te bepalen welke leerlingen op welk onderdeel extra aandacht nodig hebben. De indeling in niveaugroepen wordt bepaald met behulp van diverse toetsen.

6.1.1 Voorwaarden toets: PO

PO Docenten gaven aan dat ze toetsdata steeds belangrijker vinden en het prettig vinden dat deze informatie uit het leerlingvolgsysteem kan worden afgelezen. De docenten gaven tevens aan dat ze het eens zijn met de voorwaarde dat rapportages vanuit een toetssysteem meer moet omvatten dan alleen een grafiek, omdat gesprekken en observaties moeten worden uitgeschreven en gerapporteerd. Wel worden problemen rondom het gebruik van ingevoerde gegevens genoemd. Een docent beschrijft bijvoorbeeld dat resultaten wel in het leerlingvolgsysteem worden ingevoerd, maar dat deze weinig worden gebruikt voor formatieve doeleinden. De aanwezigheid van een praktische handleiding vergemakkelijkt het gebruik van deze systemen. Verder moet het leerlingvolgsysteem goed afgestemd zijn op het type toetsresultaat dat wordt ingevoerd (bijvoorbeeld Cito resultaten).

De docenten omschreven dat ze door middel van toetsen de ontwikkeling van de vaardigheden van leerlingen proberen te monitoren. De docenten gaven echter aan dat ze als voorwaarde van toetsen graag zouden zien dat de verschillende vaardigheidniveaus van leerlingen in kaart kunnen worden gebracht. Daarbij vinden docenten het belangrijk dat er ook gericht wordt op sterke leerlingen in plaats van alleen op zwakke leerlingen. Ook zien de docenten als een voorwaarde van toetsen dat ze inhoudelijk aansluiten op elkaar, zoals de aansluiting van methodegebonden toetsen op Cito toetsen, of op de instructie van de docent.

Verder benoemden de docenten dat digitaal toetsen voordelen kan hebben ten opzichte van papieren toetsen, omdat bij digitaal toetsen een overzicht gegeven wordt van alle toetsonderdelen terwijl bij het analyseren van papieren toetsen er vaak alleen aandacht wordt besteed aan toetsonderdelen waarop laag is gescoord. Ook vastgestelde normeringen waaraan cijfers gerelateerd kunnen worden zien de docenten als voorwaarde van een toets. Over het afnemen van toetsen gaf een docent aan: *"Veel leerlingen hebben het helemaal niet door als ze een toets krijgen. Dan zeg ik oh, we gaan een toets maken en dan zeggen de leerlingen oh, oké. En dan zit iedereen klaar en dan gaat het beginnen. Ze doen er helemaal niet moeilijk over, omdat ze het gewend zijn om toetsen te krijgen."* De docenten omschreven dat zij zelfevaluaties (self-assessment) door leerlingen, bijvoorbeeld via portfolio's, een belangrijk onderdeel vinden van formatief toetsen en dat er meerdere toetsmomenten moeten plaatsvinden zodat het leerproces kan worden gemonitord. Ook vonden ze het observeren door middel van een checklist naast andere vormen van toetsing een voorwaarde voor formatieve toetsen.

6.1.2 Voorwaarden docent: PO

PO Focusgroepdeelnemers gaven aan dat docenten kennis en vaardigheden over toetsgegevens en datagebruik moeten hebben bij het gebruik van formatieve toetsing in de klas. Verder is het belangrijk dat de docenten open staan voor veranderingen van hun eigen onderwijspraktijk en dat zij kennis en vaardigheden ontwikkelen waarmee hun onderwijs aangepast kan worden. "*Mensen begrijpen wel wat je zegt, maar vervolgens zijn er ook vaardigheden nodig om het uit te kunnen voeren. Eerder vertelde jij als docent hoe het moet en nu moet je eigenlijk aan het kind vragen van hoe denk jij hoe het moet."* Kennis en vaardigheden over hoe onderwijs gedifferentieerd aangeboden kan worden is ook van belang. Een docent illustreert: *"Wanneer leerlingen uitblinken, geef je ze uitdagend materiaal. Als leerlingen niet meekomen met de gewone stof, dan geef je die een eigen leerweg."*

Verder gaven de docenten aan dat zij feedback heel belangrijk vinden, zodat er niet alleen gekeken wordt naar de resultaten, maar naar iedere leerling individueel. Een voorwaarde is dat deze feedback breed inzetbaar is en zich niet alleen richt op probleemoplossende vaardigheden, maar ook op het leerproces van leerlingen, zodat zij eigenaar worden van hun eigen ontwikkelproces. Een docent beschreef: "*Ja, het is bewust maken van de leerling. Je hebt een tien gehaald, wow, wat goed, maar hoe heb je het aangepakt dat je een tien hebt gehaald? Maar ook bij een kind wat een vijf heeft gehaald, hoe kan dat dan? Wat heb je gedaan met leren?"* Een andere voorwaarde is dat er aan het eind van een les tijd moet worden vrijgemaakt om feedback te kunnen geven aan leerlingen.

Verder bleek het betrekken van de leerling bij het leerproces door middel van gesprekken een belangrijke voorwaarde voor formatief toetsen. Ook werd aangegeven dat docenten het doel van een les moeten weten en dat de lesmethode hierin niet te sturend mag zijn. Gedurende een les moeten de juiste vragen gesteld worden en moet doorgevraagd worden op basis van de antwoorden van leerlingen. Een docent gaf als voorbeeld: "*Je vraagt wel regelmatig aan een kind, hoe ging het? Als dan het antwoord is 'goed', dan nemen we dat voor lief, van ja dat is fijn. Terwijl als we doorvragen, van wat ging er dan zo goed? Hoe heb je dit aangepakt? Wat is er zo speciaal aan jouw aanpak waardoor die som ook goed gelukt is?"*

Verder gaven de docenten aan dat een positieve attitude en gezamenlijke visie voor een verandering van de onderwijspraktijk een belangrijke voorwaarde is voor formatieve toetsing. Zo moeten docenten begrijpen waarom het bijhouden van resultaten in systemen belangrijk is, naast het bijhouden van de resultaten in hun eigen administratie. Daarnaast is het belangrijk dat docenten inzien dat wanneer een bepaalde vaardigheid bij een leerling minder goed ontwikkeld is, dit de gehele schoolloopbaan problemen op kan leveren en er dus aandacht aan besteed moet worden.

Daarnaast gaven de docenten aan dat ouders bij het onderwijsproces betrokken moeten worden om informatie uit te wisselen over de leerling. Ook moet er een veilig leerklimaat voor leerlingen worden gecreëerd, zodat leerlingen aan durven geven hoe ze aan bepaalde taken hebben gewerkt en fouten durven te maken. Daarnaast is overleg tussen docenten over de (kenmerken van de) leerlingen belangrijk. Verder noemden docenten dat het belangrijk is dat formatieve toetsen ook daadwerkelijk formatief worden gebruikt.

6.1.3 Voorwaarden leerling: PO

PO Docenten gaven aan dat leerlingen moeten leren hun eigen werk te beoordelen, op hun leerproces te reflecteren, en fouten te durven maken zodat de leerling ontdekt hoe hij/zij zichzelf kan ontwikkelen en het leerproces eigen maakt. Verder werd omschreven dat wanneer leerlingen zelf kunnen kiezen aan welk ontwikkelpunt ze willen werken dit motiverend werkt. Tot slot werd benoemd dat het belangrijk is dat leerlingen zelf doelen kunnen opstellen om opdrachten of gedrag te kunnen evalueren: "*Ik vind wel, de leerlingen moeten leren om ook zelf doelen te stellen. Die moeten wel een belangrijke rol spelen bij formatieve toetsing. Die zelfevaluatie*."

6.1.4 Voorwaarden context: PO

PO Docenten benoemden veel voorwaarden die nodig zijn voordat formatieve toetsing succesvol ingevoerd kan worden in een school. Zo is de manier waarop data gebruikt wordt binnen een school een belangrijke *voorwaarde voor de context*. Ook wordt opbrengstgericht werken soms gezien als een verplichting, welke in de ogen van de docenten wordt opgedragen door de inspectie.

Daarnaast gaven de docenten aan dat ontwikkelingen binnen een school niet te langzaam mogen gaan, maar dat het gehele team de visie moet delen zodat de ontwikkeling kan worden doorgevoerd. Zowel de schoolleiding als het team is verantwoordelijk voor het opstellen van rapportages. Daarnaast hebben de schoolleiders en de stichting een belangrijke rol bij het doorvoeren van nieuwe ontwikkelingen. Aandacht aan het zorgsysteem binnen de school, actieve deelname in en facilitering van bijvoorbeeld budget, en inzicht in de dagelijkse praktijk van docenten worden door de docenten benoemd als belangrijke voorwaarden. Verder gaven de docenten aan dat het bestuur van de school veel initiatieven stimuleert rondom onderwijsverbetering, zoals het opstellen van enquêtes en een borgingscyclus. Dit wordt als prettig ervaren.

Ook wordt professionele ontwikkeling belangrijk gevonden. Docenten noemden dat ze getraind worden door andere docenten en interne begeleiders van de school in het gebruik van bijvoorbeeld het leerlingvolgsysteem. Echter, docenten gaven ook aan dat ze behoefte hebben aan meer scholing, bijvoorbeeld over hoe onderwijs kan worden afgestemd op individuele leerlingen en over hoe leerlingen geleerd kan worden te reflecteren op hun eigen leerproces. Professionelere ondersteuning, vanuit bijvoorbeeld onderwijsadviesbureaus, is wenselijk, maar hiervoor is tijd en budget noodzakelijk. Een docent beschrijft: *"Het gaat heel vaak mis op geld. Dat je dan daadwerkelijk denkt van ik zou die cursus wel willen doen of ik zou die mensen wel willen binnenhalen, maar dan is er vaak geen budget om dat te realiseren."* Daarnaast gaven de docenten aan dat lerarenopleidingen meer aandacht moeten geven aan hoe je een onderwijsaanpak kunt aanpassen. Binnen professionele ontwikkeling speelt ook coaching een grote rol. Met behulp van coaching kan input gegeven worden aan verdere professionalisering. Een docent beschrijft dat in tegenstelling tot vijftien jaar geleden, stagiairs en docenten tegenwoordig worden gecoacht door

docenten, interne begeleiders of opleiders in de school.

De docenten beschreven dat zij het belangrijk vinden om samen te werken binnen het team door middel van teamvergaderingen en door te brainstormen met collega's over lesinstructies zodat men van elkaar kan leren. De docenten vinden het belangrijk dat het samenwerken tussen collega's gestructureerd verloopt om afhankelijkheid te creëren, bijvoorbeeld in de vorm van een coöperatieve werkvorm. Een docent beschreef: *"We hebben bijvoorbeeld ook mini-teams, twee nu, waarbij er een docent van groep 7/8, een van groep 5/6 en een van groep 3/4 een mini-team vormt om de doorlopende leerlijn in de gaten te houden."* Ook het belang van samenwerken in de klas wordt omschreven, bijvoorbeeld door gesprekken met leerlingen te voeren over het onderwijsproces.

6.1.5 Voorwaarden ouders: PO

Een factor die niet naar voren kwam in de literatuur review, maar wel van belang bleek in de focusgroepen is de rol van ouders. De docenten omschreven dat ouders vaak een attitude hebben waarin ze het moeilijk vinden als hun kind wat zwakker scoort dan een gemiddelde leerling in Nederland. De docenten zien het als een taak van ouders en de school om met leerlingen in gesprek te gaan over leerresultaten. Een docent beschreef: *"Ja, ik mis wel dat ouders ook in gesprek gaan van, hoe heb je dat gedaan? Dus het gesprek aangaan en bespreken hoe is dat nu gekomen? En wat is het effect ervan? En wat doet dat met een ander? En wat doet dat met jezelf? En wat kun je er in het vervolg mee doen?"* Daarnaast gaven de docenten aan dat ouders vaak alleen weten dat een toets van Cito in groep 8 wordt gebruikt als eindtoets, terwijl toetsen van Cito gedurende de hele schoolloopbaan tweemaal per jaar worden gebruikt om de resultaten van de leerlingen in kaart te brengen. De docenten omschreven dat het belangrijk is om ouders te betrekken in het onderwijsproces van hun kind, zodat de ontwikkeling van de leerling zo volledig mogelijk in kaart kan worden gebracht. Docenten verwachtten dat ouders deze betrokkenheid ook prettig vinden. Tabel 22 geeft een overzicht van de relatie tussen de resultaten uit de PO focusgroepen in relatie tot de literatuur review.

TABEL 212 SAMENVATTING FACTOREN FOCUSGROEPEN PO

Voorwaarde	**Gevonden in review**
Toets	
Goed functionerend informatie management systeem met heldere rapportages	x
Gebruik van verschillende soorten data (ook observaties, sociaal emotionele vaardigheden en groei van leerlingen i.p.v. vergelijking landelijk gemiddelde)	x
Mogelijkheden tot digitaal toetsen	x
Toegang tot verschillende soorten gegevens (toetsen, observaties) van goede kwaliteit	x
Frequent toetsen, maar niet te vaak	x
Groei en ontwikkeling kunnen meten van zowel sterke als zwakke leerlingen	
Toetsen moeten aansluiten op elkaar (bijvoorbeeld Cito en de methodetoetsen)	
Gebruik van benchmarks en vastgestelde normeringen en cijfers	x
Docent	
Positieve attitude: gebruik van toetsen kan leiden tot beter onderwijs, eigenaarschap over resultaten, vertrouwen dat docent invloed heeft op leren van leerlingen	x
Data en assessment literacy	x
Vakdidactische kennis (PCK)	x
Gebruik van discussies voor directe feedback en doorvragen	x
Heldere doelen per les, waarbij de methode niet te sturend is	
Ouders betrekken bij het proces, informatie uitwisselen	
Toepassen feedback strategieën m.b.t. resultaat en proces	x
Eind iedere les tijd vrij maken voor geven van feedback	

Voorwaarde	Gevonden in review
Veilig leerklimaat creëren waarin fouten gemaakt mogen worden	
Samenwerken rondom toetsen en resultaten	x
Leerling	
Fouten durven maken	
Aan de hand van duidelijke toetscriteria-feedback hun eigen werk analyseren en reflecteren op eigen resultaten	x
Actieve rol van leerlingen in toetsen, inclusief bepalen van doelen, kan betrokkenheid en motivatie beïnvloeden, wat de prestaties kan beïnvloeden.	x
Context	
Data die formatief bedoeld zijn (e.g. observaties van docenten in de klas) niet summatief gebruiken	
Professionele ontwikkeling nodig	x
Leiderschap voor toetsen en datagebruik (inclusief klasbezoeken)	x
Data experts, coaches, specialisten in de school	x
Samenwerking met andere scholen, van elkaar leren	
Tijd en geld beschikbaar voor formatief toetsen	x
Samenwerking (Professionele leergemeenschappen) gericht op verbeteren van onderwijs door middel van datagebruik	x
Meer aandacht binnen lerarenopleidingen voor formatief toetsen	
Ouders	
Gesprek met de school aangaan over resultaten leerling	
Betrokkenheid tonen bij onderwijsproces leerling	

6.2 FOCUSGROEPEN VO

OGW en DT zijn in beide VO focusgroepen bekende begrippen. AfL wordt, na uitleg, veelal geïdentificeerd met het dagelijkse "*onderwijsleergesprek*". Overigens betekent het bekend voorkomen van de begrippen geen volledige kennis van de begrippen: gedurende de focusgroepen komen herhaaldelijk misverstanden aan bod en wordt formatieve en summatieve toetsing door elkaar gehaald. Toetsing wordt daarnaast in het begin veelal gezien als een papieren toets, welke leidt tot een beoordeling. In de praktijk zijn volgens de deelnemers de drie vormen (OGW, AfL en DT) met elkaar verweven.

Alle deelnemende docenten proberen een vorm van formatief toetsen te gebruiken in de praktijk, ieder op een eigen manier en vaak zonder een vorm van scholing op het gebied van formatief toetsen. Een voorbeeld is een docente wiskunde waarbij de leerlingen elke week een kort toetsje maken, waardoor zij inzicht krijgt in wat de leerlingen al wel begrijpen en wat ze nog zouden moeten oefenen. De docente voorziet de leerlingen niet van een cijfer, maar van gedetailleerde feedback. De docenten hebben gemengde ervaringen met betrekking tot het doorvoeren van formatieve toetsing in hun scholen. Zo noemt een docent *"Ik voel me wel alleen staan binnen mijn sectie wat betreft formatief toetsen, andere docenten meekrijgen gaat heel langzaam".* Een andere docent licht toe: "*Het is de bedoeling dat wij dat langzaam maar zeker de school in brengen, een stroperig proces, het gaat uiteindelijk wel lukken.*"

6.2.1 Voorwaarden toets: VO

Wat betreft de voorwaarden voor de toets leggen de VO deelnemers vooral nadruk op de frequentie waarmee getoetst wordt. Te veel toetsen is ook niet goed. Dit kan niet alleen leiden tot demotivatie van leerlingen, maar deelnemers geven ook aan dat er verwachtingspatronen kunnen ontstaan waar de leerling of de docent niet altijd aan kan voldoen. Om deze reden is daarom ook twijfel over of data van de basisschool beschikbaar moeten zijn aan het begin van de middelbare school. Het zorgt ervoor dat leerlingen niet meer

opnieuw kunnen beginnen. Wel zou het handig zijn om in kaart te brengen of bepaalde leerlingen (onderdelen uit) de fundamentele basis missen zodat daar op voortgebouwd kan worden. Met betrekking tot tijdbesparing zien de deelnemers een meerwaarde in digitaal toetsen. Maar naast facilitering van middelen moet er ook sprake zijn van facilitering rondom de ontwikkeling en het leren werken met dit type materialen.

Over het wel of niet geven van een cijfer heerst onenigheid. Enerzijds hebben de deelnemers de ervaring dat wanneer er geen beloning gekoppeld wordt aan de toets (zoals een cijfer), de leerlingen de toets niet serieus nemen en leerlingen niet gemotiveerd worden om de toets te maken. Anderzijds hebben deelnemers de ervaring dat het inzicht krijgen in wat nog geoefend moet worden voldoende beloning is voor leerlingen en dat het koppelen van een cijfer aan een formatieve toets (onnodig) stress oplevert. Duidelijk wordt dat verwachtingen en gevolgen met betrekking tot het proces van formatief toetsen uitgesproken dienen te worden naar alle betrokken partijen, dus ook naar de leerlingen zelf en naar ouders.

6.2.2 Voorwaarden docent: VO

De attitude van VO docenten ten opzichte van de mogelijkheden van formatief toetsen is positief. Deelnemers vinden dat met behulp van formatief toetsen meer inzicht in het niveau van de leerling wordt gecreëerd, waardoor de leerling op een passend niveau uitgedaagd en ondersteund kan worden tijdens het leerproces. Daarnaast zijn een aantal deelnemers van mening dat het gebruik van formatieve toetsing noodzakelijk wordt, zeker met het oog op veranderingen in het Nederlandse onderwijs en ter ondersteuning van gedifferentieerd onderwijs.

Volgens de deelnemers heeft een schoolbrede aanpak een positieve invloed op de houding en visie van docenten met betrekking tot formatief toetsen. Deze positieve houding is volgens alle deelnemers belangrijk, maar lang niet altijd aanwezig. Wel merken de deelnemers op dat de literatuur een onrealistisch ideaalbeeld schetst over de docent: "*Als je alle competenties gaat doorlezen en als ik aan al deze competenties zou voldoen, dan mag ik haast wel een shirtje met super docent aan! En die bestaat dus niet!*" Deelnemers vinden sommige competenties belangrijker dan andere competenties en benadrukken dat het intermenselijke aspect van lesgeven niet vergeten moet worden.

Het belang van een basisniveau van toetsgeletterdheid komt ook aan bod, "*anders weet je niet waar je naar toe moet en welke vragen je moet stellen in de klas zelf*." Kennis met betrekking tot de echte ontwikkeling van een toets mag van een basisniveau zijn, aldus deelnemers, mits er voldoende ondersteuning plaats vindt. Toetsontwikkeling wordt gezien als een ambacht en men is van mening dat dit soms beter overgelaten kan worden aan mensen die er echt goed in zijn. Ondersteuning van toetsontwikkeling kan komen van externe partijen (zoals uitgeverijen of via digitale docenten platforms) of van binnen de school (bijvoorbeeld via collega's). Overigens zijn de docenten vooralsnog van mening dat toetsen die horen bij methodes over het algemeen van onvoldoende niveau zijn en niet voldoen aan wensen van de gebruikers, "*dan kan ik alsnog zelf aan de slag*". Daarnaast moet de docent kennis hebben van zowel de vakinhoud als van didactische strategieën om leerlingen te ondersteunen die bij een kennisgrens zijn en om een gevarieerde les aan te kunnen bieden.

Docenten erkennen dat er op de lerarenopleiding maar weinig aandacht wordt besteed aan alle benodigde vaardigheden, maar zijn ook van mening dat de kennis die een docent heeft, samenhangt met de professionele fase waarin een docent zich bevindt. Zo werd aangegeven dat een startende docent al blij zal zijn als hij een klas onder controle heeft. Daarna zal hij pas de structuur van het boek los kunnen laten om te zien wat er in de klas gebeurt. Verder vinden de deelnemers het belangrijk dat een docent zich verantwoordelijk voelt voor de resultaten van leerlingen, maar er moet wel rekening worden gehouden met het feit dat de docent dit niet altijd in de hand heeft. Docenten voelen zich tot op bepaalde hoogte verantwoordelijk voor een toets en de bijbehorende toetsresultaten: "*Tuurlijk snap ik wel dat ik invloed heb*

op een leerling, maar als een leerling het niet kan of het niet wil, he. Ja. Is het dan een voorwaarde van mij als docent dat ik moet zorgen dat hij doet wat ik wil? Dat vind ik wel een lastige, zeker als je kijkt naar resultaten."

6.2.3 Voorwaarden leerling: VO

De VO docent moet zijn best doen om leerlingen te stimuleren. Maar ook de leerling heeft een belangrijke rol in het leerproces en moet inzet tonen en willen leren. Volgens de docenten wordt er soms in de literatuur een ideaalbeeld geschetst wat niet helemaal reëel is. Ervaringen van docenten laten zien dat leerlingen het fijn en motiverend vinden om feedback te krijgen over hun eigen leerproces, mits ze dit gewend zijn. De mate waarin leerlingen het ontvangen van feedback gewend zijn en de manier waarop een leerling reageert op feedback verschilt. Docenten verwachten dat dit verschil voornamelijk voortkomt uit gewenning van de leerling aan de huidige cijfercultuur op scholen. De leerlingen hebben nu het beeld dat een toets vooral summatief is: *"Leerlingen zijn natuurlijk wel heel erg gewend dat ze een toets krijgen, een cijfer krijgen en dan weer verder gaan. Die denken 'ok ik ben niet zo goed in elektriciteit, volgende hoofdstuk, dan gaan we het nu hebben over iets anders'."* Dit beeld zou moeten veranderen, zodat het inzicht krijgen in het eigen leerproces voldoende stimulans biedt. Mogelijkheden hiervoor kunnen gevonden worden binnen de eigen school*, "als ze vanaf het begin dat ze hierop school 'opgevoed' worden met het idee dat deze toets ze inzicht kan geven in wat ik nog moet oefenen, dan kan dat ook voldoende stimulatie zijn."* Verder moeten leerlingen volgens de docenten kunnen reflecteren en vaardigheden aanleren om elkaar op een goede manier feedback te kunnen geven.

6.2.4 Voorwaarden context: VO

VO Docenten benoemden veel voorwaarden waaraan moet worden voldaan voordat formatieve toetsing succesvol ingevoerd kan worden in een school. *Voorwaarden voor de context* worden door deelnemers veelal als eerste aangehaald en veel herhaald. Om formatief toetsen op een succesvolle manier te implementeren in de praktijk moet aangestuurd worden op een passend beleid, met onder meer duidelijke verwachtingen, doelen en normen gericht op formatief toetsen. Ook moet de schoolleiding tijd en middelen beschikbaar stellen om dit beleid te bewerkstelligen. Er moet echter ruimte blijven voor zelfstandige invulling van het schoolbeleid en samenwerking tussen docenten moet gefaciliteerd worden. Deelnemers geven aan dat ze toewijzing van faciliteiten zien als 'erkenning van belang' door de schoolleiding. Daarnaast moet bewustzijn rondom de tijd die daadwerkelijk nodig is voor het doorvoeren van veranderingen aanwezig zijn.

Docenten zijn enthousiast over samenwerken en positieve ervaringen ondersteunen dit enthousiasme: "O*nze vakgroep Biologie heeft nu wel een soort netwerk opgezet met een aantal collega's van de andere zuilen binnen de scholengroep en die hebben hele leuke dingen ontwikkeld en die delen alles met elkaar*". Een bijkomend voordeel is dat het aldus de docenten veel oplevert, ook in verhouding tot de tijd die het kost. Daarnaast wordt het belang van samenwerking tussen verschillende vakgroepen belicht. Aan de ene kant om vakoverstijgende vaardigheden te ontwikkelen ("*Een kwart keer een kwart kunnen ze prima berekenen, maar als ze bij mij een kwart keer een kwart moeten doen bij een stamboom dan gaat het fout*."), maar ook bijvoorbeeld bij het gebruik van dezelfde middelen ("*Dan waren een heleboel docenten die dat hadden geïntroduceerd, en dan komt een havo 4 in de klas. Nou meneer, dat is de vierde keer vandaag*.").

Naast het beleid op school moet er ook bij de overheid en de inspectie passend beleid worden ontwikkeld. Docenten ervaren moeilijkheden bij het inpassen van formatief toetsen in de huidige eisen over het curriculum*: "Zo lang de overheid nog een monitorsysteem en gestelde targets bepaalt, kun je als school natuurlijk van alles willen op het gebied van formatief toetsen, op het gebied van individuele leerroutes, noem maar op: maar dan blijf je toch vast zitten aan dat examen. Dat we altijd daar naar toe moeten."* Vrijheid om het curriculum aan te passen en om daarbinnen formatieve toetsing toe te passen vereist daarom een heldere en doorlopende leerlijn. Tabel 23 geeft een overzicht van de relatie tussen de resultaten uit de VO

focusgroepen in relatie tot de literatuur review.

TABEL 223 SAMENVATTING FACTOREN FOCUSGROEPEN VO

Voorwaarde	Gevonden in review
Toets	
Mogelijkheden tot digitaal toetsen	x
Frequent toetsen, maar niet te vaak	x
Wel of geen cijfer geven?	x
Doelen van de toets en verwachtingen moeten helder zijn	x
Docent	
Positieve attitude: gebruik van toetsen kan leiden tot beter onderwijs, eigenaarschap over resultaten	x
Data en assessment literacy	x
Vakdidactische kennis (PCK)	x
Toetsen kunnen ontwerpen voor specifieke doeleinden, maar er moeten ook goede toetsen beschikbaar zijn	x
Leerling	
Positieve attitude ten opzichte van toetsen	x
Bewustzijn dat een toets niet alleen summatief, maar ook formatief kan zijn	
Mede leerlingen op een goede wijze beoordelen, aan de hand van heldere criteria en voorzien van feedback	x
Context	
Toetscultuur: visie, normen en doelen; samenwerking rondom toetsen, gezamenlijke verantwoordelijkheid; focus op verbetering	x
Helder beleid rondom formatief toetsen	x
Tijd en geld beschikbaar voor formatief toetsen	x
Leiderschap voor toetsen en datagebruik	x
Samenwerking (professionele leergemeenschappen gericht op verbeteren van onderwijs door middel van datagebruik)	x
Autonomie om beslissingen te nemen	x
Coherent beleid rondom formatief toetsen landelijk (onderwijsinspectie), bij schoolbesturen en scholen	x

6.3 FOCUSGROEP MBO

De term formatief toetsen wordt door de docenten geassocieerd met 'ontwikkelingsgericht toetsen' in het MBO. Daarbij benadrukten de docenten wel dat het gebruik van formatieve toetsen in het MBO nog in de kinderschoenen staat: "*…in de kern heeft het MBO leerling-gerichte formatieve toetsing gewoon nog niet doorgevoerd. Dat heeft alles te maken met de geschiedenis waarin we een afrekencultuur hadden en waarin we nu heel voorzichtig zijn met dat formatieve, omdat het vooral niet op examinering mag lijken. Daar lijkt nu weer wat ruimte in te komen maar dat betekent wel, we beginnen nog maar net met het formatief bouwen van onze toetsen. Eigenlijk is het pas een discussie die een jaar of twee geleden echt van de grond is gekomen om daar inhoud en vorm aan te geven. En dan heb ik het dus over discussie, nog niet eens over uitvoering.*"

Volgens de docenten wordt een leerlingvolgsysteem nog niet gebruikt, maar zijn er wel een aantal instrumenten waarmee de leervoortgang vastgelegd kan worden. Verder worden schoolbrede leerlinggegevens verzameld door middel van lesobservaties, panelgesprekken, intake gesprekken, de JOB enquête en een enquête die 'de leerling beoordeelt de docent' heet. Deze gegevens zijn beschikbaar maar worden op dit moment nauwelijks gebruikt. Komend najaar worden de gegevens beschikbaar gesteld voor

docententeams zodat zij het curriculum aan kunnen passen aan de behoefte van de leerling.

Naast de schoolbrede, OGW gerelateerde gegevens, kunnen de docenten hun praktijkervaringen vooral relateren aan AfL. Eén van de docenten vertelde over haar goede ervaring met het geven van opdrachten die direct gerelateerd waren aan een situatie in de beroepspraktijk. Een 'coach' begeleidde het proces waarin leerlingen ook van elkaar konden leren. In de opdrachten werd uitgelegd welke competenties beoordeeld werden, welke stappen leerlingen moesten ondernemen en waar informatiebronnen gevonden konden worden. Verder moeten alle MBO leerlingen een individueel leerplan maken, maar tijd ontbreekt vaak voor docenten om met elke individuele leerling aan de slag te gaan. In de beroepspraktijk zelf worden nu voornamelijk summatieve toetsen gebruikt. Wel voeren de docenten stagegesprekken met de leerlingen om nieuwe ontwikkelpunten te identificeren.

6.3.1 Voorwaarden toets: MBO

Het relateren van opdrachten en taken aan de relevante beroepspraktijk is inhoudelijks volgens de MBO docenten een belangrijke voorwaarde voor de toets. Dit is niet alleen belangrijk omdat het MBO specifiek is gericht op het functioneren in de beroepspraktijk, maar ook omdat het de motivatie van de leerling versterkt. Ook moet formatief toetsen vaker plaatsvinden in de beroepspraktijk in plaats van alleen op school. Een voorwaarde die volgens de docenten mist in de resultaten uit de literatuurreview is dat toetsen besproken moeten worden met leerlingen. Toetsen moeten niet alleen een cijfer teruggeven aan de leerling, maar moeten leiden tot inzicht in wat fout en goed is gedaan . Wel geven de docenten aan dat ze dit nog lastig toe te passen vinden in de praktijk. Daarnaast is het belangrijk meerdere gegevens te verzamelen over leerlingen en te denken aan het creëren van een veilige omgeving; "*…hoe creëer je veiligheid zodat leerlingen de waarheid in hun beleving willen vertellen, dat ze niet het wenselijke antwoord geven. Zodat jij objectieve informatie krijgt ter verbetering van je curriculum.*"

6.3.2 Voorwaarden docent: MBO

De MBO docent moet de beroepspraktijk kunnen integreren in toetsmomenten en instructie. Een docent van de opleiding tot klassenassistent zegt hierover: "*Ik laat ze bezig gaan met hoe ga ik kinderen breuken uitleggen. Alleen al die zin maakt dat ze allemaal rechtop gaan zitten. Daarom vind ik onze proeven van bekwaamheid zo leuk want die gebeuren eigenlijk alleen maar in de beroepspraktijk en daar zijn ze heel gemotiveerd voor want ze zien een directe link."*

Verder is eigenaarschap van docenten in het toetsproces een belangrijke voorwaarde volgens de docenten. Op dit moment worden veel instrumenten van bovenaf opgelegd. Docenten vinden dat het ontwikkelen van toetsen teruggegeven moet worden aan de docent. Dit betekent echter wel dat docenten over de kennis en vaardigheden moeten beschikken die hiervoor nodig zijn, terwijl het verschil in kennisniveau kenmerkend is voor het MBO: *"….in het MBO vind ik zo'n verschil zitten, want mensen die hier docent zijn die zijn daar niet allemaal voor opgeleid. Op mijn opleiding kreeg ik gewoon vier jaar training tot docent. En in het mbo zijn heel veel mensen die zijn ingestroomd met enkel een BVE cursus en zo hun bevoegdheid hebben gehaald tot docent."*

Naast het ontwikkelen van toetsen moeten docenten ook training krijgen in het analyseren en interpreteren van toetsgegevens. Een aantal deelnemers zijn het erover eens dat het teveel gevraagd is om van elke docent een bepaald niveau voor toetsontwikkeling te verwachten. Daarom zouden docenten in een team moeten werken waarin docenten met verschillende specialisaties zitten: *"…maar ik denk wel dat je binnen een team een aantal docenten moet hebben die dan affiniteit hebben met toetsen en sommige wat meer affiniteit met hele andere onderdelen van het vak en ik denk dat je daar een mooie mix van zou moeten hebben."* Hiervoor is het wel van belang dat docenten ook in een dergelijk team met elkaar kunnen samenwerken.

Daarnaast wordt benadrukt dat docenten het nut moeten inzien van formatief toetsen en het verzamelen van de gegevens die daarbij horen. Volgens de deelnemers komt het nog maar weinig voor dat docenten zelf gegevens opvragen om hun eigen doelen te monitoren. Vaak worden doelen door docenten gezien als iets dat is opgelegd door de inspectie, waardoor de intrinsieke motivatie ontbreekt.

6.3.3 Voorwaarden leerling: MBO

Om formatief toetsen succesvol te laten zijn is het volgens de MBO docenten belangrijk dat leerlingen vaardigheden met betrekking tot zelfreflectie ontwikkelen. Hierbij moeten leerlingen zelf hun groei vast kunnen leggen en kunnen bepalen wat ze geleerd hebben en wat ze nog moeten leren. Ook moeten leerlingen om kunnen gaan met het krijgen van feedback. "….*maar daar heb je zeker een bepaald niveau voor nodig van inzicht in jezelf en in je eigen ontwikkeling."*

Daarnaast moeten leerlingen ook een omschakeling maken van het cijfersysteem dat ze gewend waren naar formatief toetsen waarin eigen verantwoordelijkheid en zelfontwikkeling centraal staat. Verder moet rekening gehouden worden met leerlingen die niet altijd intrinsiek gemotiveerd zijn. *"…maar je hebt in de eerste instantie binnen het MBO te maken met een leeftijdsfase die gewoon met zich mee brengt dat het intrinsieke er nog niet in zit. Het is gewoon verplicht, ik moet naar school."* Echter merken docenten dat, als de leerling wel gemotiveerd is voor het beroep waarvoor ze gekozen hebben, toetstaken die direct gerelateerd zijn aan de beroepspraktijk effectief zijn. Het is daarom ook een belangrijke voorwaarde dat de leerling het nut van de toetsopdracht inziet voor de toekomstige beroepspraktijk.

6.3.4 Voorwaarden context: MBO

Hoewel de docenten de meerwaarde van maatwerk inzien en vinden dat docenten toetsen zelf moeten kunnen ontwikkelen, ervaren ze op dit moment hierin belemmering vanuit de context. Hierbij gaat het vooral om de beschikbaarheid van tijd en middelen, en over het beleid dat omtrent (formatief) toetsen wordt gevoerd. De docenten merken een spanning tussen de verdeling van het beschikbare budget en de verschillende belangen van betrokkenen in de school. Een docent wijst erop dat tijd samenhangt met het geld dat beschikbaar gesteld wordt. Ondanks dat de deelnemers het eens zijn dat de school docenten moet stimuleren en faciliteren in het verzamelen van gegevens om lessen te verbeteren, is dit volgens een deelnemer niet de realiteit: *"…scholen worden gestuurd door geld. En geld loopt onder andere naar aanleiding van allerlei overheidsregelingen waar we aan moeten voldoen."* De deelnemers concluderen dat teveel door de overheid vastgelegde regels en verantwoording averechts werkt voor een succesvolle implementatie. De nadruk komt dan te liggen op verantwoording en niet op verbetering.

Verder wordt training als een voorwaarde genoemd met betrekking tot de schoolbrede gegevens die beschikbaar worden gesteld voor docententeams. De schoolleiding moet het trainen van docenten in het analyseren en interpreteren van deze gegevens faciliteren. Ook moet de samenwerking tussen docenten gefaciliteerd worden door de school maar deze wordt volgens de docenten op dit moment steeds beperkter; alleen voor of na werktijd is ruimte voor samenwerking.

Daarnaast noemen de deelnemers dat een positieve houding ten opzichte van datagebruik van belang is, maar dat de huidige afrekencultuur daar niet aan mee werkt; *"Op de een of andere manier zit er in onze cultuur echt een afrekening. Door extern maar ook door onszelf. Als je daar in kunt winnen gaat er veel voor elkaar komen en kun je van data leren, maar als we daar niet in gaan winnen dan blijft data bedreigend."*

In die cultuurbepaling spelen leidinggevenden een belangrijke rol. Het is daarom volgens de deelnemers ook belangrijk dat een leidinggevende/directie dicht bij de inhoud van de lespraktijk staat. Tussen de deelnemers verschilt de eigen ervaring wat betreft de leidinggevenden. Een van de deelnemers noemt dat de leidinggevenden ook op inhoudelijk vlak met het curriculum bezig zijn, terwijl de leidinggevenden van een

andere deelnemer juist ver van de praktijk af staat en zich vooral bezig houden met het opstellen van de begroting. "*…dus elkaar ontmoeten op dat vlak, tussen directie, CvB en docenten, ja dat is veel moeilijker geworden. Want we praten vanuit twee werkelijkheden. De ene moet de begroting rond kijken en de andere de lessen."* Tabel 24 geeft een overzicht van de relatie tussen de resultaten uit de MBO focusgroepen in relatie tot de literatuur review.

TABLE 234 SAMENVATTING FACTOREN FOCUSGROEP MBO

Voorwaarde	**Gevonden in review**
Toets	
Toetstaken en opdrachten moeten gerelateerd zijn aan de beroepspraktijk om zo relevant te zijn voor het beroep waar de leerling voor opgeleid wordt	x
Moeten (meer) in de beroepscontext plaatsvinden	x
Moeten leiden tot inzicht in waarom antwoorden of opdrachten fout of goed zijn	
Docent	
Eigenaarschap voelen van het toetsproces	x
De beroepspraktijk kunnen integreren in toetsen	
Nut inzien van het verzamelen van gegevens en formatief toetsen	x
Eigen toetsen kunnen ontwikkelen	x
In staat zijn toetsgegevens te analyseren en interpreteren	x
Samenwerken in een multidisciplinair docententeam	x
Creëren van een veilige toetsomgeving waarin leerlingen eerlijk kunnen antwoorden.	
Positieve houding en intrinsieke motivatie t.o.v. datagebruik.	x
Leerlingen	
Vaardigheden m.b.t. zelfreflectie ontwikkelen.	x
Bewust zijn van een andere manier van toetsen	
Positieve houding t.o.v. het ontvangen van feedback	
Nut van de toetstaak inzien	x
Gemotiveerd zijn voor het beroep waarvoor ze gekozen hebben.	
Context	
Beschikbaarheid van tijd en middelen	x
Toetsen moeten niet van bovenaf opgelegd worden, eigenaarschap over het toetsproces moet teruggegeven worden aan de docent	
Directie/leidinggevenden moeten dicht bij de lespraktijk staan.	
Cultuur moet datagebruik stimuleren door datagebruik positief te benaderen (zonder afrekening)	
Directie/leidinggevenden moet samenwerken tussen docenten faciliteren.	x
Directie/leidinggevenden moet training voor docenten faciliteren m.b.t. data analyse en interpretatie.	x
Niet te veel regelgeving vanuit de overheid	x

6.4 FOCUSGROEP EXPERTS

Deelnemers gaven aan dat in hun beleving bij formatief toetsen meerdere zaken een rol spelen, namelijk: formatief toetsen is voor meerdere partijen (voor de docent en leerling, maar ook voor de schoolorganisatie), heeft als doel "*om van te leren*" en bevat daarom altijd feedback. Per onderwijstype en methode varieert de manier waarop formatief toetsen wordt gebruikt. De begripsomschrijvingen van OGW, AfL en DT worden herkend.

6.4.1 Voorwaarden toets: Experts

Toetsen moeten van een goede kwaliteit zijn en mogen niet misbruikt worden. Ook al zijn dit geen specifieke voorwaarden voor formatieve toetsen, toch worden deze voorwaarden door de experts extra benadrukt. Daarnaast moeten eventuele toetssystemen voorzien in heldere rapportages. Een voorwaarde

van de toets en het toetssysteem is verder dat het past bij het doel wat je als docent voor ogen hebt: "*Het is voor mij heel erg afhankelijk van wat de docent met de toets beoogt. Want volgens mij hoeft de toets helemaal niet creativiteit te stimuleren, als je gewoon wilt weten of de leerling zijn taallesje goed heeft geleerd.*" Experts gaven aan dat feedback in de vorm van een cijfer al haast per definitie summatief gebruik van de toets is, maar hier ontstaat wel enige discussie over. Het terugkrijgen van een cijfer geeft wel een indicatie van een niveau.

6.4.2 Voorwaarden docent: Experts

De deelnemers zijn van mening dat kennis over toetsen bij docenten van belang is. Niet alleen over zaken zoals validiteit en betrouwbaarheid, maar ook juist over waarom een toets op een bepaalde manier in elkaar zit. Kennis hierover moet docenten aangeleerd worden op de lerarenopleiding. *"Ja, en daar hoort dan ook een klein stukje toets theorie bij, en dan inderdaad, formatief en summatief, gewoon. Toetsen en testen. Ja, ze doen het iedere dag en iedere dag beoordelen ze kinderen op allerlei verschillende manieren en ze leren er veel te weinig over."* Overigens verwachten experts dat kennis bijbrengen op alleen de lerarenopleiding niet voldoende is. Toetskennis zou in kleine stapjes geleerd moeten worden, om uiteindelijk een onderwerp te vormen wat constant centraal staat. "*Als het bij de pabo aan de orde komt, dan moeten we er meer op gaan letten. Als het in een professionaliseringsprogramma nog een keer terug komt. Als het op het jaarcongres van de vo raad centraal staat. Als je allemaal van die hele kleine dingen doet, dan heb je uiteindelijk kans op resultaat.*" Met kennis kan ook eventueel aanwezige angst van docenten richting toetsen weggenomen worden. Er wordt aangegeven dat er eerst een goede definitie moet komen en dat vervolgens bepaald moet worden wat formatief toetsen inhoudt en wat een docent moet weten. "*Sowieso is toetsen of assessment heel erg veel en ik zou gewoon, voordat je de pabo en professionaliseringstrajecten in gaat, eerst zelf heel goed moeten weten wat nou de basis is die een docent of een docent moet kennen.*" Docenten in het mbo zijn vaak zij-instromers en niet iedere docent heeft een pedagogisch-didactische achtergrond. Ook deze zij-instromers horen kennis te ontwikkelen over toetsen, wil men gebruik maken van formatieve toetsen.

Maar volgens de experts is boven alles het betrekken van de leerling door de docent misschien wel het meest belangrijke punt: "*Hoe motiveer je nu die student en koppel je die bijvoorbeeld aan het stukje professionele ontwikkeling. En dan kijk ik bijvoorbeeld bij het VO. Van als je studenten, als je leerlingen, op een goede manier weet te motiveren, dan schep je automatisch al een aantal voorwaarden waardoor je zelf eigenlijk al professionaliseert waardoor je eigenlijk een soort cirkelbeweging creëert die alleen maar omhoog kan gaan.*" Verder moet volgens de experts de leerling door de leraar serieus worden genomen.

6.4.3 Voorwaarden leerling: Experts

Over opgestelde voorwaarden met betrekking tot de leerlingen zijn de deelnemers het unaniem eens: *"Hoezo, de leerling moet bereid zijn om te leren? Het is je uitdaging als leraar om er voor te zorgen. De voorwaarden van de leerling, sorry."*

6.4.4 Voorwaarden context: Experts

Tijdens de bijeenkomst werd er niet zozeer gefocust op toetstechnische zaken, maar op docentprofessionalisering en de schoolcultuur. Het belang van de juiste voorwaarden voor de context wordt daarmee aangestipt. De schoolleiding wordt geïdentificeerd als een belangrijke partij bij het initiëren van verandering, bijvoorbeeld op het gebied van formatief toetsen: "*Door het gesprek ook aan te gaan binnen de school, zal je ook zien dat je als schoolleiding je cultuur echt kan creëren en dat het niet gezien wordt als een soort dictatoriale omgeving maar dat het ook gezien wordt als kans om een gesprek aan te gaan.*" Verantwoordelijkheid bij een of meerdere personen wordt gezien als noodzakelijk, omdat zij zich op die manier "*wel moeten focussen op een bepaalde verandering.*" Hierbij is het wel van belang om duidelijke

doelen te stellen, zodat het traject richting het behalen van de doelen ingericht kan worden.

De aanwezigheid van instrumenten is een duidelijke voorwaarde: "*Het wel of niet aanwezig zijn van voldoende toetsen die ook voldoende gericht zijn op wat je er mee wilt doen*". Experts geven aan dat ondanks een aanwezige vraag om ondersteuning, het af en toe lastig is om in te schatten waar de vraag precies naar is en dat hierdoor een spanningsveld ontstaat tussen scholen en externe betrokkenen. Zo worden er bijvoorbeeld instrumenten ontwikkeld die niet overeenkomen met het gewenste gebruik van de docent: "*Docenten hebben behoefte aan zo'n instrument en gaan dan zelf maar wat ontwikkelen. En.. dat is natuurlijk wel een spanningsveld. Waar scholen of docenten behoefte aan hebben en wat je zelf biedt, en of dat dan ook wel verantwoord is*". Experts verwachten daarnaast dat wanneer docenten weten waar instrumenten te vinden zijn binnen de school, het gebruik van de instrumenten wordt verhoogd.

Eventuele kritiek op, en angst voor, toezicht vanuit de inspectie of de overheid zou anders bekeken moeten worden aldus de experts. De docenten hebben een defensieve houding, terwijl ze van data kunnen leren. Tabel 25 geeft een overzicht van de relatie tussen de resultaten uit de Experts focusgroepen in relatie tot de literatuur review.

TABEL 245 SAMENVATTING FACTOREN FOCUSGROEP EXPERTS

Voorwaarde	**Gevonden in review**
Toets	
Goed functionerend informatiemanagement-systeem met heldere rapportages	x
Doelen van de toets en verwachtingen moeten helder zijn	x
Docent	
Data en assessment literacy	x
Vakdidactische kennis (PCK)	x
Leerlingen motiveren om te leren	
Context	
Toetscultuur: Beleid, visie, normen en doelen; samenwerking rondom toetsen, gezamenlijke verantwoordelijkheid; focus op verbetering; mogelijkheden tot experimenteren	x
Leiderschap voor toetsen en datagebruik	x
Toegang tot en kennis van verschillende toetsmethodes en tools	x
Gesprek aangaan met onderwijsinspectie over toetsen	

7 CONCLUSIE

In deze review is onderzoek gedaan naar verschillende factoren die van invloed zijn op het succesvol implementeren van formatief toetsen in de klas. De resultaten zijn gebaseerd op wetenschappelijke literatuur en focusgroepdiscussies met docenten en experts. Dit hoofdstuk vat de belangrijkste resultaten samen, onderverdeeld in vier categorieën: voorwaarden met betrekking tot de toets, de docent, de leerling en de context.

7.1 LITERATUUR REVIEW

Voorwaarden Toets

Ten aanzien van toetsen wordt in de literatuur opgemerkt dat ze goed aan moeten sluiten bij het curriculum, de leerdoelen en het niveau van de leerlingen. Van tevoren moet duidelijk geformuleerd zijn welke informatie de toets moet opleveren en hoe deze informatie uit de toetsresultaten gehaald kan worden. Als daarbij gebruik gemaakt wordt van een toetssysteem, dan moet het systeem goed toegankelijk en gebruikersvriendelijk zijn, moeten de gegevens duidelijk en interpreteerbaar opgeleverd worden en moet de feedback die de systemen genereren aansluiten bij de informatiebehoefte. Ten aanzien van de toestvragen wordt opgemerkt dat docenten invloed moeten hebben op de keuze en formulering van de vragen en dat de vragen aangepast moeten worden aan de informatiebehoefte. Verder wordt aanbevolen dat de formatieve toetsen geregeld afgenomen worden (minimaal om de zes weken), dat leerlingen genoeg tijd krijgen om de vragen te beantwoorden en dat het bevorderlijk is om de toetsresultaten ook te bespreken. Tenslotte komt in de literatuur naar voren dat feedback op basis van de resultaten een aantoonbaar effect heeft op de motivatie om te leren. De genoemde factoren werden zowel in literatuur over PO, VO en MBO gevonden en waren voornamelijk afkomstig uit studies naar AfL en OGW.

Voorwaarden Docent

De resultaten wijzen uit dat voor het effectief inzetten van formatief toetsen in de klas docenten (en schoolleiders) moeten beschikken over voldoende kennis en vaardigheden om gegevens te verzamelen, te analyseren en te interpreteren en om concrete maatregelen te treffen op basis van de toetsresultaten. Bij dat laatste speelt ook didactische en vakinhoudelijke kennis een belangrijke rol. Tevens blijkt dat een positieve houding ten aanzien van formatief toetsen van belang is voor een succesvolle implementatie. Dit betreft de intrinsieke motivatie van docenten, bereidheid tot kritische reflectie, verantwoordelijkheids-gevoel en de overtuiging het verschil te kunnen maken. Omdat bij formatief toetsen vaak gebruik gemaakt wordt van informatie uit toetssystemen is een open houding m.b.t. en kennis van technologie bovendien van belang. Daarnaast blijkt samenwerking tussen docenten belangrijk bij formatief toetsen. Met name in de voorwaardelijke sfeer speelt de schoolleiding een cruciale rol. Docenten kunnen veel van elkaar leren en het uitwisselen van positieve ervaringen kan stimulerend en motiverend werken. Een essentiële competentie bij het implementeren van formatief toetsen, tenslotte, is het vermogen van docenten om goede feedback te geven. Deze factoren werden zowel bij PO, VO als MBO gevonden en hadden betrekking op AfL en OGW.

Voorwaarden Leerling

De resultaten benadrukken dat leerlingen in staat moeten zijn zelfstandig te reflecteren op hun leerproces. Ook hebben zij vaardigheden nodig om de leerprestaties van zichzelf en van hun medeleerlingen te kunnen beoordelen en vervolgens effectieve feedback te geven. Tevens moeten leerlingen actief betrokken zijn bij het proces van formatief toetsen, dit stimuleert motivatie en leidt tot betere leerresultaten. Om dit te kunnen bereiken is een open houding van de leerling van belang. Bijna alle literatuur over de leerlingen in formatief toetsen ging over AfL in het PO en VO.

Voorwaarden context

Ten aanzien van de context binnen de school wordt in de literatuur vooral gewezen op de cultuur. Een succesvolle implementatie gaat vaak samen met een cultuur van samenwerking, gericht op continue verbetering, waarin ruimte is om te experimenteren. Hierbij is het van belang dat de visie op formatief toetsen expliciet wordt geformuleerd. Daarnaast is het cruciaal dat er voldoende tijd en middelen voor formatief toetsen worden vrijgemaakt. De docenten en schoolleiders moeten ook in staat zijn om zichzelf te professionaliseren, met name op het terrein van het gebruiken, analyseren en interpreteren van toetsresultaten en om dit te vertalen naar interventies in de klas. Ondersteuning door specialisten wordt daarbij als wenselijk ervaren. Om dit te realiseren is het van belang dat de schoolleiding aandacht heeft voor formatief toetsen en samenwerkt met docenten op het gebied van formatieve toetsing. De schoolleiding wordt vaak genoemd als cultuurdrager en wordt in veel studies benoemd als degene die de doelen en verwachtingen moet formuleren (in samenspraak met docenten). Kennis en vaardigheden over formatief toetsen zijn daarom voor schoolleiders tevens van essentieel belang.

Ten aanzien van de context buiten de school wordt in veel studies gewezen op de invloed van het verantwoording afleggen op het implementeren van formatief toetsen. Enerzijds blijkt dat externe druk motiveert om met behulp van formatief toetsen de leerresultaten te verbeteren. Anderzijds blijkt teveel druk het effectief inzetten van formatief toetsen echter te belemmeren. De gevonden factoren komen zowel binnen het PO als het VO naar voren en komen met name uit studies gericht op OGW.

7.2 FOCUSGROEPEN

De resultaten van de literatuur review zijn in focusgroepen voorgelegd aan docenten uit verschillende onderwijscontexten (PO, VO, MBO) en experts op het gebied van formatief toetsen. Tijdens elke focusgroep is gediscussieerd over de voorwaarden die volgens de deelnemers van belang zijn voor een succesvolle implementatie van formatief toetsen in de onderwijspraktijk. De deelnemers uit de onderwijspraktijk reflecteerden op individuele ervaringen met formatief toetsen in hun eigen lespraktijk en de manier waarop formatief toetsen in de toekomst geïmplementeerd zou kunnen worden. In de onderwijspraktijk wordt de term formatief toetsen niet gebruikt maar wel inhoudelijk herkend. De expertgroep reflecteerde vanuit een wetenschappelijk perspectief en zijn zowel met formatief toetsen als de drie vormen van formatief toetsen bekend. We gaan eerst in op de aspecten die het meest naar voren kwamen bij de verschillende focusgroepen.

De meeste overeenkomsten tussen alle focusgroepen zijn gericht op voorwaarden voor de context en voorwaarden m.b.t. competenties van de docent. Alle deelnemers zijn het eens dat leiderschap een essentiële rol speelt bij de implementatie van formatief toetsen. Hierbij worden voorwaarden met betrekking tot de (toets)cultuur genoemd (bijv. helder beleid), maar ook voorwaarden die gericht zijn op de ondersteuning van formatief toetsen zoals het faciliteren in tijd, middelen, samenwerking en professionele ontwikkeling. Zowel de docenten uit het PO, VO en MBO ervaren de beschikbaarheid van tijd en middelen als een belangrijke voorwaarde. Vanuit een intern beleid moet ruimte vrij gemaakt worden voor docenten om effectief aan formatief toetsen te kunnen werken. Ook moet volgens de docenten tijd vrij gemaakt worden voor samenwerking op het gebied van formatief toetsen. Samenwerking moet gericht zijn op het verbeteren van onderwijs, gestructureerd verlopen en vakoverschrijdend zijn. De expertgroep merkte op dat het implementeren van formatief toetsen op dit moment op veel scholen een onderwijsvernieuwing betekent en dat schoolleiders een belangrijke partij zijn bij het initiëren van veranderingen. Ondanks dat een schoolleider verantwoordelijk is voor het doorvoeren van dergelijke vernieuwingen, werd zowel in het PO als het MBO benadrukt dat een nauwe samenwerking met docenten van belang is en dat een leidinggevende dichtbij de onderwijspraktijk moet staan.

Verder werd in elke focusgroep genoemd dat docenten en schoolleiders over kennis en vaardigheden moeten beschikken met betrekking tot het verzamelen, analyseren en interpreteren van toetsresultaten. Vervolgens moeten docenten in staat zijn op basis van deze resultaten hun lessen te verbeteren en daarmee de leeropbrengsten van leerlingen te verhogen. Zowel de docenten uit het PO en de experts noemen dat docenten ook didactische en vakinhoudelijke kennis nodig hebben om dit mogelijk te maken. In de focusgroep met MBO docenten werd in het verlengde hiervan genoemd dat docenten toetstaken moeten kunnen integreren in de beroepspraktijk. Daarnaast benadrukten de docenten ook dat een positieve houding ten opzichte van formatief toetsen van belang is. Docenten en schoolleiders moeten de mogelijkheid krijgen om benodigde kennis en vaardigheden via professionele ontwikkelingsactiviteiten te verwerven. In de focusgroepen met het PO en het MBO werd genoemd dat coaches of specialisten in de school voor verdere professionele ontwikkeling en ondersteuning zouden kunnen zorgen.

In de docentfocusgroepen werd ook nadruk gelegd op competenties die leerlingen nodig hebben om formatief toetsen succesvol te implementeren. Leerlingen moeten in staat zijn zelfstandig te reflecteren op hun leerproces en hebben vaardigheden nodig om medeleerlingen op een goede manier van feedback te kunnen voorzien. Bij de docentfocusgroepen voor het PO en het MBO werd ook specifiek het belang van een positieve houding van leerlingen genoemd. Daarnaast noemden PO en MBO docenten dat leerlingen zich bewust moeten worden van een andere (formatieve) manier van toetsen.

Met betrekking tot de voorwaarden voor de toets zijn er verschillen gevonden tussen de focusgroepen. Zowel de docenten uit het PO als het MBO vonden het belangrijk dat verschillende soorten gegevens beschikbaar zijn en gebruikt worden bij formatief toetsen. Zij benadrukten het gebruik van zowel kwantitatieve als kwalitatieve gegevens maar ook niet-cognitieve uitkomstmaten, zoals sociale competenties. Tijdens de focusgroepen met het PO en het VO werd aangestipt dat frequent (formeel) toetsen noodzakelijk is maar wel met mate moet gebeuren. Ook digitale toetssystemen vonden de docenten uit deze focusgroepen een belangrijk middel om formatief toetsen te kunnen realiseren. Aansluitend hierop merkten zowel docenten uit het PO en de expertgroep op dat het goed functioneren van dergelijke systemen een voorwaarde is. Ook noemden de docenten uit het VO en de experts dat doelen en verwachtingen met betrekking tot toetsen helder moeten zijn. Verder valt op dat specifiek voor het MBO als voorwaarde bij de toets genoemd werd dat deze opdrachten en taken moet omvatten die gerelateerd zijn aan de relevante beroepspraktijk. Ook zouden formatieve toetsen vaker in de beroepspraktijk plaats moeten vinden. Docenten uit het PO benoemden specifiek dat ouders betrokken moeten worden bij het onderwijsleerproces van de leerling.

7.3 RICHTLIJNEN VOOR DE PRAKTIJK

Uit de literatuur en de focusgroepen komen veel verschillende voorwaarden naar voren. De vraag is echter wat dit betekent voor scholen en docenten. Ondanks dat hier meer onderzoek naar nodig is, hebben we getracht om op basis van de resultaten een aantal tips te formuleren:

- **Maak het onderwerp toetsen bespreekbaar binnen de school**. Bespreek met elkaar dat het belangrijk is om niet alleen summatief te toetsen, maar ook formatief te toetsen. Bespreek met elkaar op welke wijze formatief toetsen (opbrengstgericht werken, assessment for learning en diagnostisch toetsen) op dit moment al dan niet wordt ingezet en hoe het ingezet zou kunnen worden binnen de school.
- **Kijk kritisch naar het toetsproces op school**. Hoe zit het met de kwaliteit van deze toetsen? Wordt er voldoende frequent getoetst? Sluiten de toetsen voldoende aan bij het curriculum, de leerdoelen en het niveau van de leerlingen? Zijn de beoordelingscriteria voor alle betrokkenen helder? Krijgen leerlingen voldoende feedback op basis van de toetsen?
- **Bekijk met elkaar of docenten de kennis en vaardigheden bezitten die noodzakelijk zijn voor**

formatief toetsen, zoals analyse vaardigheden, feedback vaardigheden en vaardigheden die nodig zijn om instructie aan te passen op basis van toetsresultaten. Realiseer hierbij dat formatief toetsen en bijbehorende vaardigheden niet vanzelfsprekend zijn. Bespreek of er behoefte is aan professionalisering en/of ondersteuning. En zo ja, welke professionalisering of ondersteuning past bij de behoefte van docenten.

- **Maak in het rooster tijd vrij om in teams aan de slag te gaan met toetsen in de school.** Deze teams kunnen zich dan bijvoorbeeld richten op het analyseren van toetsen en het bespreken van de resultaten in relatie tot het gegeven onderwijs. Belangrijke vragen waarop een dergelijk team zich kan richten zijn: Welke leerbehoeften hebben de leerlingen? En wat betekent dit voor onze instructie in de klas?
- **Betrek leerlingen bij het formatief toetsen, door ze bijvoorbeeld af en toe hun eigen werk of dat van medeleerlingen te laten beoordelen**. Dit vraagt wel om heldere toetscriteria aan de hand waarvan leerlingen op hun eigen werk kunnen reflecteren, om zo verbeterpunten en vervolgstappen in hun eigen leerproces te kunnen identificeren.
- **Een gedeelde visie, normen, doelen en verwachtingen zijn van belang voor een toetscultuur waarin het verbeteren van het onderwijs centraal staat en ruimte is om te experimenteren**. De schoolleider speelt een belangrijke rol in het creëren van deze cultuur. De schoolleider kan bijvoorbeeld zorgdragen voor de benodigde infrastructuur (systemen waarin toetsresultaten eenvoudig met elkaar gedeeld kunnen worden) en facilitering (in tijd om als docenten samen over toetsen en onderwijsinhoud te praten.

8 DISCUSSIE

Er is veel onderzoek gedaan naar formatief toetsen. Helaas wordt hier echter niet altijd op heldere wijze over gerapporteerd. In de literatuurreview is gezocht naar studies over formatief toetsen in de klas. Van de bijna 2000 publicaties gevonden in de literatuur zoektocht bleken slechts 72 studies relevant voor dit onderzoek en van voldoende kwaliteit. Veel mogelijk relevante studies zijn uitgesloten omdat essentiële informatie over de methode van onderzoek ontbrak, waardoor de kwaliteit van de resultaten niet beoordeeld kon worden. Daarnaast waren de gevonden studies voornamelijk gericht op OGW en Afl in het PO en VO en zijn er nauwelijks studies gevonden op het gebied van DT en nauwelijks studies in de MBO context, ondanks dat hiervoor aanvullende zoekacties zijn uitgevoerd. Er is meer onderzoek nodig naar DT en naar formatief toetsen binnen het MBO. Een voorzichtige conclusie uit de focusgroepen is bijvoorbeeld dat er tot op heden vooral summatief getoetst wordt in het MBO (bijvoorbeeld de proeven van bekwaamheid) en dat formatief toetsen nog in de kinderschoenen staat. Om dit beter in beeld te brengen is echter een verdere inventarisatie van formatieve toetsactiviteiten in het MBO nodig. Ook zal onderzocht moeten worden op welke manier MBO docenten hier verder in geprofessionaliseerd kunnen worden.

Uit de review en de focusgroepen komt naar voren dat er behoefte is aan professionalsering op het gebied van formatief toetsen, niet alleen in het MBO, maar ook in het PO en VO. Er is onderzoek nodig naar hoe deze professionalisering eruit moet zien. Welke kennis en vaardigheden (assessment en data literacy) hebben docenten bijvoorbeeld nodig? Moeten docenten zelf in staat zijn om toetsen te ontwikkelen? Uit de focusgroepen kwamen tegenstrijdige punten naar voren. Een aantal docenten was van mening dat docenten zelf goede toetsen moeten kunnen ontwikkelen. Andere docenten gaven aan dat ze vinden dat hier instanties voor zijn, die de benodigde expertise in huis hebben om betrouwbare en valide toetsen te ontwikkelen. Een belangrijke kanttekening hierbij is dat het begrip 'toets' vaak werd gezien als een formele activiteit die los plaatsvindt van instructie, bijvoorbeeld het afnemen van een pen-en-papier toets. Dit terwijl er heel veel verschillende vormen van toetsen zijn. Bij Afl gaat het er bijvoorbeeld om dat er op basis van verkregen bewijs over het leren van leerlingen continue aanpassingen worden gedaan om het leren te optimaliseren. De gebruikte toetsmomenten kunnen dan geïntegreerd plaatsvinden in de les. Het is mogelijk dat het begrip 'toetsen' een te grote lading heeft en veel docenten dit associëren met summatieve toetsen, terwijl zij wellicht onbewust gebruik maken van formatieve toetsmomenten. Een nuttige richting voor vervolgonderzoek is het vormgeven van professionele ontwikkeling rondom formatief toetsen. Waar moet deze zich op richten? Wat zijn de kenmerken van deze professionele ontwikkeling? Hoe kan deze geïmplementeerd en geëvalueerd worden? Onder welke condities is deze professionalisering effectief in de zin dat het tot het leren van docenten en leerlingen leidt?

Daarnaast is het van belang dat er onderzoek wordt uitgevoerd naar de manier waarop de verschillende formatieve toetsbenaderingen (OGW, AfL en DT) in de praktijk worden gecombineerd en geïntegreerd met instructie in de klas, en hoe dit mogelijk geoptimaliseerd kan worden om te komen tot meer valide beslissingen over leerlingen op alle niveaus in de school (Van der Kleij et al., 2013). Ook moeten vervolgens de effecten van deze toetspraktijken in kaart gebracht worden.

Tevens is er meer onderzoek nodig naar nieuwe toetsen in het PO, VO en MBO, zowel gericht op cognitieve als niet-cognitieve aspecten. Toetsen, zoals adaptieve of diagnostische toetsen, die adequaat o.a. leerprogressie en misconcepties van leerlingen in kaart kunnen brengen zijn noodzakelijk. De resultaten van de focusgroepen wijzen uit dat er behoefte is aan diagnostische toetsen, maar de beschikbaarheid en daarmee het gebruik hiervan in de praktijk staat nog in de kinderschoenen. Ook is er nog maar weinig literatuur beschikbaar over het gebruik van diagnostische toetsen in de onderwijspraktijk, dit is een belangrijke richting voor vervolgonderzoek. Tevens is er een gebrek aan instrumenten die niet-cognitieve

vaardigheden meten, zoals studievaardigheden, sociale vaardigheden en creativiteit. Vervolgonderzoek zou zich kunnen richten op het ontwikkelen, implementeren en evalueren van meetinstrumenten voor niet-cognitieve vaardigheden die formatief kunnen worden gebruikt.

Ook het belang van feedback komt naar voren in deze review. Echter, hoe de feedback er precies uit moet zien (wel of geen cijfers), wanneer deze gegeven moet worden (direct of vertraagd), in welke vorm (geschreven of mondeling), door wie deze gegeven moet worden (het toetssysteem, de docent, een medeleerling) blijft onduidelijk. Uit onderzoek komen ook tegenstrijdige resultaten naar voren. Zo bleek uit een onderzoek (Peterson en Irving, 2008) dat het geven van cijfers essentieel was voor bijvoorbeeld de motivatie van leerlingen om te leren. Een ander onderzoek (Lee, 2011) daarentegen concludeerde juist dat het geven van cijfers een negatieve impact had op de bereidheid van leerlingen om risico's te nemen en nieuwe dingen te leren. Ook in de focusgroepen waren docenten het onderling niet eens over het wel of niet geven van cijfers bij formatieve toetsen als vorm van feedback.

Uit de literatuurreview komt tevens naar voren dat leerlingen een belangrijke rol spelen bij formatief toetsen. De meerderheid van de docenten in de focusgroepen benadrukten dit ook. Zij gaven bijvoorbeeld aan dat docenten en leerlingen samen verantwoordelijk zijn voor het leerproces. Echter, hoe deze samenwerking vorm gegeven kan worden blijft onduidelijk. Verder onderzoek kan zich richten op de rol van docenten en leerlingen die samen gebruik maken van formatief toetsen in de klas.

Tot slot valt op dat de context, zowel binnen als buiten de school, een grote invloed heeft op de praktijk van formatief toetsen in de klas. Echter rapporteerden veel van de gevonden studies over de invloed van bijvoorbeeld de Amerikaanse context op formatief toetsen. Deze resultaten zijn niet zonder meer te generaliseren naar de Nederlandse context door het verschillende onderwijssysteem, de verschillende schaal en verschillende beleidsniveaus. De invloed van context op de praktijk van formatief toetsen in de Nederlandse klas moet daarom verder onderzocht worden.

Concluderend, deze review heeft een globaal beeld opgeleverd van wat het implementeren van formatief toetsen vergt van scholen. Hierbij vonden we een grote mate van overeenstemming tussen de wetenschap en de praktijk. Het belang van formatief toetsen is duidelijk, maar de toepassing hiervan in de klas vraagt nog veel van scholen, schoolleiders, docenten en leerlingen. Tot slot willen we benadrukken dat de benoemde factoren in isolatie niet zullen leiden tot een succesvolle implementatie van formatief toetsen in de klas. Om te komen tot een succesvolle implementatie zal een goede balans nodig zijn tussen de verschillende factoren op het niveau van de toets, docent, leerling en context in en buiten de school. Vervolgonderzoek moet duidelijk maken op welke wijze OGW, AfL en DT geïmplementeerd kunnen worden zodat het leidt tot de beoogde resultaten: het leren van leerlingen.

9 REFERENTIES

9.1 REFERENTIES GEBRUIKT VOOR THEORIE IN HET RAPPORT

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*, 5–25. doi:10.1080/0969594X.2010.513678

Crisp, G. T. (2012). Integrative assessment: Reframing assessment practice for current and future learning. *Assessment & Evaluation in Higher Education*, *37*, 33–43. doi:10.1080/02602938.2010.494234

Fullan, M., & Watson, N. (2000). School-based management: Reconceptualizing to improve learning outcomes. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, *11*, 453–473. doi:10.1076/sesi.11.4.453.3561

Keeley, P., & Tobey, C. R. (2011). *Mathematics formative assessment*. Thousand Oaks, CA: Corwin.

Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, *16*, 263–268. doi:10.1080/09695940903319646

Mandinach, E. B., & Honey, M. (2008). *Data-driven school improvement: Linking data and learning*. Teachers College Press, NY: New York.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide.* Oxford, UK: Blackwell.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, *48*, 4–11. doi:10.1080/00405840802577536

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144. doi:10.1007/BF00117714

Sanders, P. (2011). Het doel van toetsen. In P. Sanders (Ed.), *Toetsen op school* (pp. 9–20). Arnhem: Cito.

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, *26*, 482–496. doi:10.1016/j.tate.2009.06.007

Schildkamp, K., Lai, M.K., & Earl (Eds.) (2013). Data-based decision making in education: challenges and opportunities. Dordrecht: Springer.

Sluijsmans, D., Joosten-ten Brinke, D., & Van der Vleuten, C. (2013). *Toetsen met leerwaarde*. Een reviewstudie naar de effectieve kenmerken van formatief toetsen. Review Studie uitgevoerd in opdracht van en gesubsidieerd door NWO-PROO.

Stobart, G. (2008). *Testing times: The uses and abuses of assessment.* Abingdon, England: Routledge.

Van der Kleij, F. (2013). *Computer-based feedback in formative assessment* (Doctoral dissertation). University of Twente, Enschede, the Netherlands. doi:10.3990/1.9789036535809

Van der Kleij, F., Vermeulen, J. A., Schildkamp, K., & Eggen, T. (2013). *Towards an integrative formative approach of data-based decision making, assessment for learning, and diagnostic testing.* In 26th International Congress for School Effectiveness and Improvement, ICSEI 2013, 03-01-2013 - 06-01-2013.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*, 3–14. doi:10.1016/j.stueduc.2011.03.001

9.2 REFERENTIES LITERATUUR REVIEW

Referentie	PO	VO	MBO	OGW	AfL	DT
1. Aschbacher, Pamela, & Alonzo, Alicia. (2006). Examining the utility of elementary science notebooks for formative assessment purposes. *Educational Assessment, 11*(3-4), 179-203.	x				x	
2. Birenbaum, M., Kimron, H, & Shilton, H. (2011). Nested contexts that shape assessment "for" learning: School-based professional learning community and classroom culture. *Studies in Educational Evaluation, 37*(1), 35-48.	x	x			x	
3. Blanc, Suzanne, Christman, Jolley Bruce, Liu, Roseann, Mitchell, Cecily, Travers, Eva, & Bulkley, Katrina E. (2010). Learning to learn from data: Benchmarks and instructional communities. *Peabody Journal of Education, 85*(2), 205-225.	x			x		
4. Brown, Launcelot I., De Four-Babb, Joyanne, Bristol, Laurette, & Conrad, Dennis A. (2014). National tests and diagnostic feedback: What say teachers in Trinidad and Tobago? *Journal of Educational Research, 107*(3), 241-251.	x			x		
5. Bryant, D. A., & Carless, D. R. (2010). Peer assessment in a test-dominated setting: Empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice, 9*(1), 3-15.	x				x	
6. Christoforidou, M., Kyriakides, L., Antoniou, P., & Creemers, B. P. M. (2014). Searching for stages of teacher's skills in assessment. *Studies in Educational Evaluation, 40*(1), 1-11.	x			x		
7. Datnow, A, Park, V., & Kennedy-Lewis, B. (2012). High school teachers' use of data to inform instruction. *Journal of Education for Students Placed at Risk, 17*(4), 247-265.		x		x		
8. Datnow, A., Park, V., & Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *Journal of Educational Administration, 51*(3), 341-362.	x			x		
9. Davis, D. S., & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment. *Journal of Educational Research, 104*(3), 202-215.	x	x			x	
10. Farley-Ripple, E. N., & Buttram, J. L. (2014). Developing collaborative data use through professional learning communities: Early lessons from Delaware. *Studies in Educational Evaluation, 42*(1)*,* 41-53.	x			x		
11. Feldman, A. & Capobianco, B.M. (2008). Teacher learning of technology enhanced formative assessment. *Journal of Science Education and Technology, 17*(1), 82-99.		x		x	x	
12. Ferster, B., Hammond, T.C., Alexander, R. C., & Lyman,H. (2012). Automated formative assessment as a tool to scaffold student documentary writing. *Journal of Interactive Learning Research, 23*(1), 81-99.		x			x	
13. Fletcher, A., & Shaw, G. (2012). How does student-directed assessment affect learning? Using assessment as a learning process. *International Journal of Multiple Research Approaches, 6*(3), 245-263.	x			x	x	
14. Fox-Turnbull, W. (2006). The influences of teacher knowledge and authentic formative assessment on student learning in technology education. *International Journal of Technology and Design Education, 16*(1), 53-77.	x				x	
15. Frey, C.A. (2009). *Teachers' understanding and use of formative assessment strategies: A multiple embedded case study in K-12 urban ring schools of a mid-size city in Rhode Island.* ProQuest LLC.	x	x			x	
16. Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. *American Educational Research Journal, 36*(3), 609-646.	x			x	x	
17. Furtak, Erin Marie, & Ruiz-Primo, Maria Araceli. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education, 92*(5), 799-824.		x			x	x
18. Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching, 49*(9), 1181-1210.		x			x	x

Referentie	PO	VO	MBO	OGW	AfL	DT
19. Gamlem, S. M., & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy and Practice, 20*(2), 150-169.		x			x	
20. Gottheiner, D. M., & Siegel, M.A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education, 23*(5), 531-557.		x			x	
21. Graney, S. B., & Shinn, M. R. (2005). Effects of reading curriculum-based measurement (R-CBM) teacher feedback in general education classrooms. *School Psychology Review, 34*(2), 184-201.	x			x	x	
22. Hargreaves, E.. (2013). Inquiring into children's experiences of teacher feedback: reconceptualising Assessment for Learning. *Oxford Review of Education, 39*(2), 229-246.	x				x	
23. Harris, Lois R., & Brown, Gavin T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education, 36*(1), 101-111.	x	x			x	
24. Harris, L. R., Brown, G. T. L., & Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, *26*(2),1-27.	x	x			x	
25. Hubbard, L., Datnow, A., & Pruyn, L. (2014). Multiple initiatives, multiple challenges: The promise and pitfalls of implementing data. *Studies in Educational Evaluation, 42*(1)*,* 54-62.	x			x		
26. Havnes, A., Smith, K., Dysthe, O.,, & Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation, 38*(1), 21-27.		x			x	
27. Jimerson, J. B. (2014). Thinking about data: Exploring the development of mental models for "data use" among teachers and school leaders. *Studies in Educational Evaluation, 42*(1), 5-14.	x			x		
28. Kay, Robin, & Knaack, Liesel. (2009). Exploring the use of audience response systems in secondary school science classrooms. *Journal of Science Education and Technology, 18*(5), 382-392.		x		x	x	
29. Kennedy, Brianna L., & Datnow, Amanda. (2011). Student involvement and data-driven decision making: Developing a new typology. *Youth & Society, 43*(4), 1246-1271.		x		x		
30. Kerr, Kerri A., Marsh, Julie A., Ikemoto, Gina Schuyler, Darilek, Hilary, & Barney, Heather. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and ussons from three urban districts. *American Journal of Education, 112*(4), 496-520.		x		x		
31. Lachat, Mary Ann, & Smith, Stephen. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk, 10*(3), 333-349.		x		x		
32. Lai, Ah-Fur, & Chen, Deng-Jyi. (2010). Web-based two-tier diagnostic test and remedial learning experiment. *International Journal of Distance Education Technologies, 8*(1), 31-53.		x				x
33. Lee, I. (2011). Bringing innovation to EFL writing through a focus on assessment for learning. *Innovation in Language Learning and Teaching, 5*(1), 19-33.		x			x	
34. Lee, H., Feldman, A., & Beatty, I.D. (2012). Factors that affect science and mathematics teachers' initial implementation of technology-enhanced formative assessment using a classroom response system. *Journal of Science Education and Technology, 21*(5), 523-539.		x		x	x	
35. Leighton, Jacqueline P., Gokiert, Rebecca J., Cor, M. Ken, & Heffernan, Colleen. (2010). Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: Implications for assessment literacy. *Assessment in Education: Principles, Policy & Practice, 17*(1), 7-21.		x				x
36. Levin, James A., & Datnow, Amanda. (2012). The principal role in data-driven decision making: Using case-study data to develop multi-mediator models of educational reform. *School Effectiveness and School Improvement, 23*(2), 179-201.		x		x		

Referentie	PO	VO	MBO	OGW	AfL	DT
37. Lysaght, Zita, & O'Leary, Michael. (2013). An instrument to audit teachers' use of assessment for learning. *Irish Educational Studies, 32*(2), 217-232.	x				x	
38. Maeng, J.L.C. (2011). *Differentiating science instruction: Success stories of high school science teachers.* ProQuest LLC. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED534590&site=ehost-live		x			x	
39. McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research, 95*(4), 203-213.	x	x			x	x
40. McNaughton, Stuart, Lai, Mei Kuin, & Hsiao, Selena. (2012). Testing the effectiveness of an intervention model based on data use: a replication series across clusters of schools. *School Effectiveness and School Improvement, 23*(2), 203-228.	x			x		
41. Newby, L., & Winterbottom, M. (2011). Can research homework provide a vehicle for assessment for learning in science lessons? *Educational Review, 63*(3), 275-290.		x			x	
42. Ní Chrónín, D., & Cosgrave, C. (2013). Implementing formative assessment in primary physical education: Teacher perspectives and experiences. *Physical Education and Sport Pedagogy, 18*(2), 219-233.	x				x	
43. O'Loughlin, Joe, Ni Chroinin, Deirdre, & O'Grady, David. (2013). Digital video: The impact on children's learning experiences in primary physical education. *European Physical Education Review, 19*(2), 165-182.	x				x	
44. Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806-813.		x			x	
45. Park, Vicki, & Datnow, Amanda. (2009). Co-constructing distributed leadership: District and school connections in data-driven decision-making. *School Leadership & Management, 29*(5), 477-494.	x	x		x		
46. Pat El, R., Tillema, H., & Van Koppen, S. W. M. (2012). Effects of formative feedback on intrinsic motivation: Examining ethnic differences. *Learning and Individual Differences, 22*(4), 449-454.		x			x	
47. Penuel, W. R., Boscardin, C.K., Masyn, K., & Crawford, V.M. (2007). Teaching with student response systems in elementary and secondary education settings: A survey study. *Educational Technology Research and Development, 55*(4), 315-346.	x	x		x	x	
48. Peterson, E.R., & Irving, S.E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238-250.		x		x	x	
49. Peterson, Euguenia, & Siadat, M. Vali. (2009). Combination of formative and summative assessment instruments in elementary algebra classes: A prescription for success. *Journal of Applied Research in the Community College, 16*(2), 92-102.			x	x	x	
50. Phelan, J. C., Choi, K., Niemi, D,, Vendlinski, T.P., Baker, E.L., & Herman, J. (2012). The effects of POWERSOURCE[C] Assessments on middle-school students' math performance. *Assessment in Education: Principles, Policy & Practice, 19*(2), 211-230.	x	x		x	x	
51. Rakoczy, K., Klieme, E., Bürgermeister, A., & Harks, B. (2008). The interplay between student evaluation and instruction: Grading and feedback in mathematics classrooms. *Zeitschrift für Psychologie/Journal of Psychology, 216*(2), 111-124.		x			x	
52. Riggan, Matthew, & Olah, Leslie Nabors. (2011). Locating interim assessments within teachers' assessment practice. *Educational Assessment, 16*(1), 1-14.	x				x	
53. Ruiz-Primo, Maria Araceli, & Furtak, Erin Marie. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3-4), 205-235.	x	x			x	
54. Sach, E. (2013). An exploration of teachers' narratives: What are the facilitators and constraints which promote or inhibit 'good' formative assessment practices in schools? *Education 3-13*.	x	x			x	
55. Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in The Netherlands: A comparison. *Educational Research and Evaluation, 14*(3), 255-282.	x			x		

Referentie	PO	VO	MBO	OGW	AfL	DT
56. Schildkamp, K., & Visscher, A. (2009). Factors influencing the utilisation of a school self-evaluation instrument. *Studies in Educational Evaluation, 35*(4), 150-159.	x			x		
57. Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement, 20*(1), 69-88.	x			x		
58. Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education, 26*(3)*,*482–496.		x		x		
59. Schildkamp, K., & Visscher, A.J. (2010). The utilisation of a school self-evaluation instrument. *Educational Studies, 36*(4), 371-389.	x			x		
60. Schildkamp, Kim, Rekers-Mombarg, Lyset, & Harms, Truus. (2012). Student group differences in examination results and utilization for policy and school development. *School Effectiveness and School Improvement, 23*(2), 229-255.		x		x		
61. Schildkamp, K., Karbautzki, L., & Vanhoof, J. (2014). Exploring data use practices around Europe: Identifying enablers and barriers. *Studies in Educational Evaluation, 42*(1)*,* 15-24.	x	x		x		
62. Staman, L., Visscher, A. J., & Luyten, H. (2014). The effects of professional development on the attitudes, knowledge and skills for data-driven decision making. *Studies in Educational Evaluation, 42*(1)*,* 79-90.	x			x		
63. Sutherland, Stephanie. (2004). Creating a Culture of Data Use for Continuous Improvement: A Case Study of an Edison Project School. *American Journal of Evaluation, 25*(3), 277-293.	x			x		
64. Van der Kleij, F. M., & Eggen, T. J. H. M. (2013). Interpretation of the score reports from the computer program LOVS by teachers, internal support teachers and principals. *Studies in Educational Evaluation, 39*(3), 144-152.	x			x		
65. Vanhoof, J., Verhaeghe, G., Van Petegem, P., & Valcke, M. (2012). Flemish primary teachers' use of school performance feedback and the relationship with school characteristics. *Educational Research, 54*(4), 431-449.	x			x		
66. Wayman, Jeffrey C., Cho, Vincent, Jimerson, Jo Beth, & Spikes, Daniel D. (2012). District-wide effects on data use in the classroom. *Education Policy Analysis Archives, 20*(25).	x	x		x		
69. Yin, Y., Tomita, M. K., & Shavelson, R. J. (2013). Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *International Journal of Science Education, 36* (4), 531-552.	x				x	x
70. Young, Viki M. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American Journal of Education, 112*(4), 521-548.	x			x		
71. Torrence, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice, 14*(3), 281-294.			x		x	
72. Gulikers, J. T. M., Biemans, H. J. A., Wesselink, R., & van der Wel, M. (2013). Aligning formative and summative assessments: A collaborative action research challenging teacher conceptions. *Studies in Educational Evaluation, 39*(2)*,* 116–124.			x		x	
73. Hawe, E. M., & Dixon, H. R. (2014). Building students' evaluative and productive expertise in the writing classroom. *Assessing Writing, 19*, 66-79.	x				x	
74. Schildkamp, K., & Visscher, A. (2009). The use of performance feedback in school improvement in Louisiana. *Teaching and Teacher Education, 26*(7), 1389-1403.	x	x		x		

BIJLAGE 1: METHODE/PROCES

De literatuur review is uitgevoerd aan de hand van de methode van Petticrew en Roberts (2006). De eerste stap betrof het formuleren van onderzoeksvragen en het definiëren van zoektermen. Vijf verschillende wetenschappelijke databases zijn geraadpleegd: Education Resources Information Center (ERIC), Web of Science, Scopus, PsychINFO en Picarta. In al deze databases is met dezelfde zoekstrategie gezocht. Er is gezocht met termen zoals "*formative assessment*", "*data-based decision making*", "*assessment for learning*", "*diagnostic testing*" en aanverwante termen zoals gevonden in de thesaurus en publicaties. Om te komen tot een selectie van relevante publicaties zijn de volgende inclusiecriteria geformuleerd:

1. De studie is gepubliceerd in een wetenschappelijk tijdschrift of is een dissertatie. Boeken, boekhoofdstukken en conferentiepapers worden niet meegenomen omdat de kwaliteit van deze publicaties niet gewaarborgd kan worden.
2. Het betreft een wetenschappelijk onderzoek. Publicatietypen als theoretische beschouwingen en opiniestukken worden niet meegenomen omdat deze review studie zich richt op factoren die een bewezen invloed hebben op het gebruik van formatieve assessment in de onderwijspraktijk.
3. Het onderzoek is uitgevoerd in de context van het primair onderwijs (PO), voortgezet onderwijs (VO) of middelbaar beroepsonderwijs (MBO).
4. Het onderzoek betreft het gebruik van formatieve toetsing (zoals gedefinieerd in deze studie) in de klas. Onderzoek gericht op andere doeleinden, bijvoorbeeld het gebruik van resultaten van formatieve toetsing door schoolleiders voor school zelfevaluatie, wordt buiten beschouwing gelaten.

De eerste zoektocht leidde tot 2533 publicaties. Na het verwijderen van niet relevante publicaties en het scannen van abstracts op relevantie bleven er 117 publicaties (wetenschappelijke artikelen en doctorale dissertaties) over. Deze 117 gevonden publicaties zijn geanalyseerd met behulp van een data extractie formulier zodat van iedere publicatie soortgelijke data verzameld werd. Het data extractie formulier bevatte de volgende onderdelen:

- Algemene informatie: auteurs, titel, journal, context (po, vo, mbo) etc.
- Onderzoeksdesign: onderzoekvragen, korte omschrijving van het onderzoek, design (kwalitatief, kwantitatief of mixed methods), vorm van formatieve assessment (OGW, AfL of DT), niveau van onderzoek (school, klas, leerling), vak, duur van het onderzoek, onderzoeksinstrumenten, analyse methoden
- Onderzoekspopulatie: aantal scholen, docenten en leerlingen en wijze van steekproef
- Resultaten: antwoord op de onderzoeksvraag, voorwaarden toets, voorwaarden proces, voorwaarden docent, voorwaarden context binnen de school, voorwaarden context buiten de school, overige beïnvloedende factoren
- Kwaliteit van het onderzoek: duidelijk onderzoekdoel, geschikte methode gezien de onderzoeksvraag, kwaliteit steekproef, kwaliteit methode, kwaliteit data analyse, kwaliteit conclusie. Elk artikel kreeg een kwaliteitsscore aan de hand van elf vragen. Per vraag werd 0, 0.5 of 1 punt toebedeeld. Alle publicaties die bij de kwaliteitscontrole een score tussen de 5 en 7 haalden (schaal 0-11) zijn gelezen door twee onderzoekers en vervolgens besproken. Alle artikelen met een kwaliteitsscore van 7 of hoger zijn vervolgens geselecteerd voor analyse.

Om de betrouwbaarheid van het data extractie formulier en de kwaliteitsselectie te waarborgen is 50% van de studies met een 7 of hoger door twee beoordelaars gecodeerd. De dubbelgecodeerde data heeft geleid tot een overeenstemming van 80% en een interbeoordelaarsbetrouwbaarheid van .620, welke door Eggen en Sanders (1993) wordt beoordeeld als acceptabel.

Alleen de studies die van voldoende kwaliteit waren zijn geselecteerd (47 stuks). Vervolgens zijn de voorwaarden van deze publicaties samengebracht en samengevat. Vanwege een tegenvallend resultaat op het gebied van studies over OGW en DT en studies die zijn uitgevoerd binnen het MBO hebben de onderzoekers verder gezocht naar literatuur. Hiermee zijn nog 27 artikelen gevonden, wat het totaal aantal artikelen geschikt bevonden voor de review op 72 heeft gebracht.

De conclusies van de eerste zoekronde zijn verwerkt in meerdere tabellen, opgedeeld aan de hand het PO en het VO, en aan de hand van de gedefinieerde voorwaarden: van de toets en het proces, de docent, de leerling en de context. Om een link te leggen met de praktijk zijn de tabellen besproken in focusgroepen. Dit had drie doelen: (1) het verifiëren van de resultaten uit het literatuuronderzoek, (2) het aanvullen van de informatie uit de literatuur en (3) het vertalen van de resultaten uit de literatuur naar de praktijk. In de eerste ronde focusgroepen zijn op basis van de eerste literatuurzoektocht tabellen over het PO besproken met PO docenten en VO tabellen met VO docenten. De PO als de VO tabellen zijn besproken met experts uit allerlei onderwijsgebieden.

Tijdens de tweede zoektocht is specifiek gezocht naar DT, OGW en MBO studies. Aanvullende studies zijn, samen met de studies die gevonden waren in de eerste zoektocht, wederom verwerkt in nieuwe tabellen. Deze tabellen zijn voorgelegd aan PO, VO en MBO docenten. De focusgroepen zijn gezamenlijk geanalyseerd en de resultaten zijn ter controle voorgelegd aan de focusgroep deelnemers.

Een gedetailleerdere versie van de procesbeschrijving en de instrumenten zijn op te vragen bij de hoofdonderzoeker (Schildkamp, K.).