5 maart 2026

Waardering praktijkgericht onderwijs III

Eindrapportage

Bernard Veldkamp
Erik Roelofs
Lientje Maas
Suzanne Gerritsen
Surya Carrilho
Maren van de Vrie
Marijke Veugen
Judith Conijn

Inhoudsopgave

1. Inleiding 3
2. Theoretisch kader 4
Realistische zelfevaluatie als noodzakelijke vaardigheid voor zelfsturing en zelfregulatie 4
Brede toepasbaarheid van kalibratievaardigheid 5
Sport: Kalibratie van sportprestaties 5
Tandheelkunde: Zelfbeoordeling en kalibratie 5
Autorijden: Kalibratie van rijvaardigheid 5
Het meten van kalibreervaardigheid ter ondersteuning van verder leren 5
3. Prototyping 7
Uitgangspunten 7
User flow 7
Startscherm 7
Doelen stellen 8
Beoordeling doelen door leerling en leraar 10
Vervolgstappen bepalen 11
Dashboards 11
Leerling dashboard 11
Leraar dashboard 12
Technische specificaties 13
4. Overdraagbaarheid 14
5. Opzet evaluatie 15
Methode 15
Doel en functie van de vignetten 15
Ontwikkelen vignetten 15
Focusgroep met schoolleiders 15
Analyse 15
Focusgroepen 15
Werving 15
Uitvoering 16
6. Uitkomsten evaluatie 17
Praktijkgericht onderwijs in de huidige situatie 17
Hoe krijgt praktijkgericht onderwijs momenteel vorm? 17
Welke plek krijgen doelen stellen en evalueren op de scholen momenteel? 17
Huidig AI-beleid op school 18
Algemeen perspectief op de tool 18

Doelen stellen 19
Meerwaarde 19
Aandachtspunten 20
Evalueren of doelen behaald zijn 21
Meerwaarde 21
Aandachtspunten 21
Vervolgstappen bepalen 22
Meerwaarde 23
Aandachtspunten 23
7. Conclusies en aanbevelingen 24
Hoe ziet dit eruit als proof of concept voor één praktijkgerichte vaardigheid? 24
Overdraagbaarheid 25
Praktijkgerichte vaardigheden in de dagelijkse praktijk van het onderwijs. 25
Concrete aanbevelingen 26
Tot slot 26
8. Referenties 27
Bijlage 1. Systeemprompt feedback opstellen doelen 28
Bijlage 2. Vignetten leraren 30
Introductie 30
Voorstelronde 30
Situatie 30
Voorbeeldopdracht 30
Evalueren of doelen behaald zijn 32
Overkoepelende vragen 33
Bijlage 3. Vignetten Schoolleiders 35
Introductie 35
Voorstelronde 35
Situatie 35
Voorbeeldopdracht 35
Evalueren of doelen behaald zijn 37
Overkoepelende vragen 38

1. Inleiding

In een Kamerbrief van september 2023 over de waardering van praktijk in het funderend onderwijs wordt gesproken over het meer stimuleren van praktijkonderwijs en het verbeteren van de waardering hiervan in po, vmbo, havo en vwo. Gezien er veel vraag is naar praktisch opgeleide mensen nu en in de toekomst, is het belangrijk dat er meer waardering komt voor praktijkgerichte vaardigheden. Met praktijkgerichte vaardigheden worden met name *metacognitieve vaardigheden* en *breed toepasbare vaardigheden* bedoeld, zoals leren samenwerken, leren plannen en werk structureren.

In dit project wordt onderzocht of een digitaal toetsmodel ondersteunend kan zijn voor het meten van praktijkgerichte vaardigheden. In overleg met de opdrachtgever is één praktijkgerichte vaardigheid gekozen, waarvoor we uitgewerkt hebben hoe deze getoetst kan worden in een digitale omgeving en hoe gerapporteerd kan worden over de leerlingvaardigheid. De probleemstelling waaraan gewerkt is luidt:

Hoe kan een toetsmodel worden ontworpen waarmee praktijkgerichte vaardigheden op valide, betrouwbare en haalbare wijze meetbaar en toetsbaar worden gemaakt, zodat deze vaardigheden kunnen worden geïntegreerd in het bredere systeem van leerlingvolgsystemen en toetsgestuurde overgangsbeslissingen in po, vmbo, have en vwo?

Om deze vraag te beantwoorden hebben we de probleemstelling geoperationaliseerd in een aantal onderzoeksvragen die we met dit onderzoek beogen te beantwoorden.

Onderzoeksvragen:

1. Welke praktische vaardigheden zijn waardevol voor het leerproces van leerlingen in het funderend onderwijs en hebben betekenis voor het vervolgonderwijs?

2. Op welke manier kunnen praktijkgerichte vaardigheden zo geoperationaliseerd worden dat ze meetbaar zijn en handvatten geven voor formatief handelen?

3. Hoe ziet dit eruit als proof of concept voor één praktijkgerichte vaardigheid?

4. In hoeverre is een dergelijke toetssystematiek overdraagbaar naar andere vaardigheden of zijn vaardigheden zo specifiek dat voor elke vaardigheid een andere toetssystematiek moet worden ontworpen?

5. Is het voor scholen in hun dagelijkse praktijk mogelijk om praktijkgerichte vaardigheden op te nemen in de leervorderingen die systematisch worden gemonitord en op basis waarvan beslissingen rond overgang naar andere leerroutes worden genomen?
    a. Hoe valide is de beoordeling?
    b. In hoeverre is de bijkomende werkdruk realistisch?

2. Theoretisch kader

Onderzoek laat zien dat er weinig aandacht is voor het gebruik van metacognitieve strategieën in praktijkgericht onderwijs (Khaled et al., 2026; Mejeh & Grieder, 2025). Tegelijkertijd wordt erkend dat het van groot belang is dat leerlingen naast beroepsgerichte kennis ook deze zogenoemde hogere orde vaardigheden leren, mede omdat deze belangrijk zijn voor *lifelong learning* (Kirschner & Stoyanov, 2020). In de praktijk zullen leerlingen in realistische werksituaties voortdurend hun leerproces moeten sturen. Hierbij moet een situatie worden ingeschat, besloten worden wat voor actie nodig is en de resultaten moeten worden gemonitord (Eraut, 2004). Hiervoor zijn metacognitieve vaardigheden essentieel.

Jossberger, Brand-Gruwel, Van de Wiel en Boshuizen (2020) hebben onderzoek gedaan naar zelfsturing en zelf-regulerend leren bij studenten in het middelbaar beroepsonderwijs (mbo). De resultaten lieten zien dat de studenten zich vaak richten op de uitkomst van een praktische opdracht in plaats van op het leerproces. Studenten stellen niet vaak expliciete (leer)doelen en zelfbeoordeling vindt meestal plaats op basis van het product, zonder het proces te beschouwen. Het evalueren van het proces is een metacognitieve vaardigheid die studenten in staat stelt om fouten te herkennen en ermee om te gaan, zodat prestaties kunnen verbeteren.

Realistische zelfevaluatie als noodzakelijke vaardigheid voor zelfsturing en zelfregulatie

De oorsprong van inzichten over zelfevaluatie, zelfsturing en zelfregulatie ligt in het begrip metacognitie: het bewustzijn van eigen denkprocessen en het vermogen deze te monitoren en bij te sturen. Dit metacognitieve vermogen blijkt een sterke voorspeller van schools leren en presteren.
In schoolse vaardigheden zoals schrijfvaardigheid (Deane, 2011) komt cognitieve zelfregulatie duidelijk naar voren. Leerlingen houden hierbij standaarden in de gaten voor helderheid en samenhang, en hechten belang aan structuur, aansluiting bij de voorkennis van de lezer, en consistentie van de tekst. Algemene regulatieactiviteiten zijn onder meer het reflecteren op het behalen van schrijfdoelen, het herordenen van ideeën, en het plannen, bijstellen, redigeren, herformuleren en corrigeren van teksten.

De leerpsychologische literatuur biedt verschillende modellen voor cognitieve zelfregulatie. Een invloedrijk model is dat van Winne en Hadwin (1998), dat vier fasen beschrijft in zelfgereguleerd leren:

- Taakbegrip – De leerling vormt een persoonlijke interpretatie van de leertaak, inclusief het doel en de sociale context.
- Doelstelling en planning – Op basis van de taakinterpretatie formuleert de leerling leerdoelen en plant strategisch hoe deze te bereiken.
- Taakuitvoering – De leerling voert de taak strategisch uit en past daarbij verschillende strategieën toe, terwijl hij/zij het eigen handelen kritisch monitort.
- Aanpassing – Op basis van zelfevaluaties wordt het eigen gedrag, de taakaanpak of het doel bijgesteld.

Volgens Winne (1995) vormt onnauwkeurigheid in deze processen een risico: als leerlingen hun prestaties verkeerd inschatten, kunnen zij ook verkeerde keuzes maken, bijvoorbeeld over de benodigde tijd, het gebruik van strategieën, of het vragen om hulp. Het missen van hiaten kan zo een negatieve leerspiraal in gang zetten. Winne en Jamieson-Noel (2002) benadrukken daarom het belang van kalibratie: de afstemming tussen het vertrouwen in eigen prestaties en de feitelijke nauwkeurigheid daarvan. Onjuiste zelfinschatting kan leiden tot zelfonderschatting of zelfoverschatting, beide schadelijk voor het leerproces. Goed kalibrerende leerlingen hebben accuratere informatie over hun prestaties en kunnen hun leerproces beter bijsturen, wat leidt tot hogere leerresultaten.

Zelfonderschatting kan leiden tot onzekerheid over doelstellingen en tot kritiekloze acceptatie van andermans oplossingen (Efklides, 2006). Zelfoverschatting leidt juist tot minder aandacht voor kritische aanwijzingen, afleiding, en aandacht voor irrelevante details of doelen. Wanneer deze verkeerde inschattingen niet worden gecorrigeerd, kunnen ze uitgroeien tot stabiele maar ongewenste persoonlijkheidskenmerken. Kalibratievaardigheid wordt daarom in de literatuur gezien als een essentiële basis voor het stellen van realistische leerdoelen en effectief leren (Boekaerts, 2009; Hadwin & Webster, 2012; Stone, 2000).

Brede toepasbaarheid van kalibratievaardigheid

Er zijn tal van studies verricht naar de rol van kalibratie, kort door de bocht gedefinieerd als: het vermogen om de eigen prestaties accuraat in te schatten, bij het aanleren van praktische vaardigheden. Hieronder noemen we er enkele.

Sport: Kalibratie van sportprestaties
Binnen de sportpsychologie wordt kalibratie onderzocht in het kader van de vraag hoe atleten hun prestaties inschatten ten opzichte van hun werkelijke prestaties. Het bleek dat atleten vaak een "better-than-average" effect vertonen (Măirean & Turliuc, 2013), waarbij ze hun eigen prestaties als beter dan gemiddeld beschouwen, ongeacht de werkelijke uitkomst. Deze overschatting bleek invloed te hebben op trainingsintensiteit en motivatie. Het bevorderen van nauwkeurige zelfbeoordeling kan atleten helpen realistischere doelen te stellen en effectievere trainingsstrategieën te ontwikkelen.

Tandheelkunde: Zelfbeoordeling en kalibratie
In een studie binnen het tandheelkundig onderwijs werd onderzocht hoe zelfbeoordeling en kalibratie van leraren invloed hebben op de prestaties van studenten tijdens praktische examens (Oh et al., 2017). Studenten die gebruikmaakten van een instructieve beoordelingsrubriek voor zelfbeoordeling, en waarbij leraren voorafgaand aan de beoordeling gekalibreerd waren, vertoonden betere prestaties en lagere faalpercentages. Dit suggereert dat duidelijke beoordelingscriteria en kalibratie van beoordelaars bijdragen aan een nauwkeurigere zelfinschatting en betere leerresultaten bij praktische vaardigheden.

Autorijden: Kalibratie van rijvaardigheid
In de context van autorijden verwijst kalibratie naar de mate waarin bestuurders hun eigen rijvaardigheid correct inschatten (Horrey et al., 2015). Een goed gekalibreerde bestuurder heeft een nauwkeurige perceptie van zijn of haar rijvaardigheid, wat essentieel is voor verkeersveiligheid. Onderzoek heeft aangetoond dat bestuurders vaak hun rijvaardigheid overschatten, wat kan leiden tot risicovol rijgedrag. Het verbeteren van kalibratie bij bestuurders kan bijdragen aan veiliger rijgedrag en betere besluitvorming op de weg.

Het meten van kalibreervaardigheid ter ondersteuning van verder leren

De voorbeelden uit verschillende toepassingsdomeinen benadrukken dat kalibratie niet alleen een theoretisch concept is, maar een praktisch relevante vaardigheid die het leerproces wezenlijk beïnvloedt. Binnen modellen voor zelfregulatie (Winne & Hadwin, 1998) vormt kalibratie een terugkerend en bepalend element in alle fasen van zelfgereguleerd leren: van het juist interpreteren van de taak en het stellen van haalbare doelen, tot het monitoren van de uitvoering en het bijstellen van strategieën en doelen op basis van zelfevaluatie. Onnauwkeurige zelfinschatting kan deze cyclus verstoren en daarmee zowel motivatie als leerresultaten negatief beïnvloeden.

Om die reden is in dit project gekozen om kalibratievaardigheid als centrale metacognitieve vaardigheid te verkennen voor instrumentontwikkeling. Een meetinstrument dat niet alleen de kwaliteit van een uitgevoerde

taak beoordeelt, maar ook de juistheid van de zelfbeoordeling ten opzichte van eigen leerdoelen, biedt inzicht in de mate waarin leerlingen hun leren effectief kunnen reguleren. Hiermee sluiten we direct aan bij de theoretische kaders over zelfgereguleerd leren, waarin het cyclische proces van doelstelling, monitoring en evaluatie als basis wordt gezien voor duurzame zelfregulatie.

Naast de waarde voor leerlingen biedt het systematisch verzamelen van zelfevaluaties belangrijke aanknopingspunten voor leraren om het leerproces te optimaliseren. Wanneer leerlingen herhaaldelijk hun prestaties en doelen zelf beoordelen, ontstaat er een zichtbaar patroon van kalibratieontwikkeling: leraren kunnen zien of leerlingen leren om hun prestaties realistischer in te schatten en welke denk- en doe stappen zij daarbij wel of niet maken. Deze informatie kan worden benut om feedback meer op maat te geven, gerichte reflectiegesprekken te voeren en het aanbod van leerstrategieën of begeleiding te differentiëren (Andrade & Valtcheva, 2009; Panadero et al., 2016).

Tot slot fungeren leerlingzelfevaluaties als een vorm van *assessment as learning*: de gegevens die daaruit voortkomen helpen leraren niet alleen bij het beoordelen van leerresultaten, maar vooral bij het begrijpen van hoe leerlingen leren (Earl, 2013). Wanneer leraren signaleren dat leerlingen structureel hun eigen prestaties over- of onderschatten, kan dit aanleiding zijn om expliciet aandacht te besteden aan metacognitieve strategieën zoals monitoring, feedbackgebruik en doelgericht plannen (Hattie & Timperley, 2007).

Het systematisch inzetten van zelfevaluaties biedt zo niet alleen inzicht in individuele leerprocessen, maar ook in de kwaliteit van de leeromgeving en de effectiviteit van instructie. Het stelt leraren in staat om feedbackcycli te verkorten, eigenaarschap van leren te vergroten en het onderwijs adaptiever in te richten — precies de functies die beoogd worden binnen het model van zelfgereguleerd leren van Winne en Hadwin (1998).
Deze inzichten vormden het uitgangspunt voor de volgende fase van het project, waarin CitoLab een eerste *proof of concept* ontwierp voor een digitale tool die leerlingen in praktijkgericht onderwijs ondersteunt bij het reguleren van hun leerproces.

3. Prototyping

Om onderzoeksvraag 3 te beantwoorden, heeft CitoLab een proof of concept ontworpen voor een tool die leerlingen in praktijkgericht onderwijs ondersteunt bij het reguleren van hun leerproces. Het proof of concept is geen werkend prototype, maar bestaat wel uit klikbare schermen om de mogelijke functionaliteit de demonstreren. In dit hoofdstuk beschrijven we de werkwijze tijdens deze fase en beschrijven we de keuzes die tijdens het ontwikkelproces gemaakt zijn.

Het prototype ondersteunt drie stappen in het leerproces van leerlingen: 1) het stellen van doelen, 2) de zelfbeoordeling van een leerling t.o.v. deze doelen en 3) het maken van keuzes voor vervolgstappen. Hierbij beogen we aan te sluiten bij de doelpopulatie (po, vmbo, havo en vwo) door een tool te bouwen zonder hoog abstractieniveau. Bij het stellen van doelen geeft een AI feedback aan de leerling om via een iteratief proces tot een goed doel te komen. We hebben ter illustratie gekozen voor toepassing van deze vaardigheden in een specifieke praktijkgerichte context, namelijk een taak waarbij leerlingen een product van hout moeten maken. De voorbeelden zijn gebaseerd op opdrachten uit de Centraal Schriftelijk en Praktisch Examens (CSPEs) van het vak *Dienstverlening en producten*.[1] Hoewel het prototype voor dit domein uitgewerkt is, zijn de onderliggende principes generaliseerbaar en ook inzetbaar bij andere domeinen.

Het prototype is te vinden op https://praktijkvaardigheden.web.app/. In dit hoofdstuk gebruiken we schermafbeeldingen om de tool toe te lichten.

Uitgangspunten

Bij het ontwerpen van het prototype is rekening gehouden met de volgende uitgangspunten:

1. De focus ligt op formatief gebruik.
2. Het proces moet een zekere mate van autonomie creëren of stimuleren bij de leerling.
3. Het gaat om een praktische vaardigheid en niet om een cognitief/theoretische vaardigheid. Metacognitieve vaardigheden bevatten ook cognitieve aspecten; daar richten we ons in mindere mate op.
4. Er moet een evaluatiemogelijkheid bestaan voor scholen en leraren om te beoordelen of leerlingen zich ontwikkelen in het stellen van doelen, het beoordelen van doelen en het bepalen van vervolgstappen.

User flow

Binnen deze paragraaf beschrijven we de user flow van het proof of concept. Een user flow is een visuele weergave van de stappen die een gebruiker (in dit geval leerling en leraar) doorloopt binnen de tool.

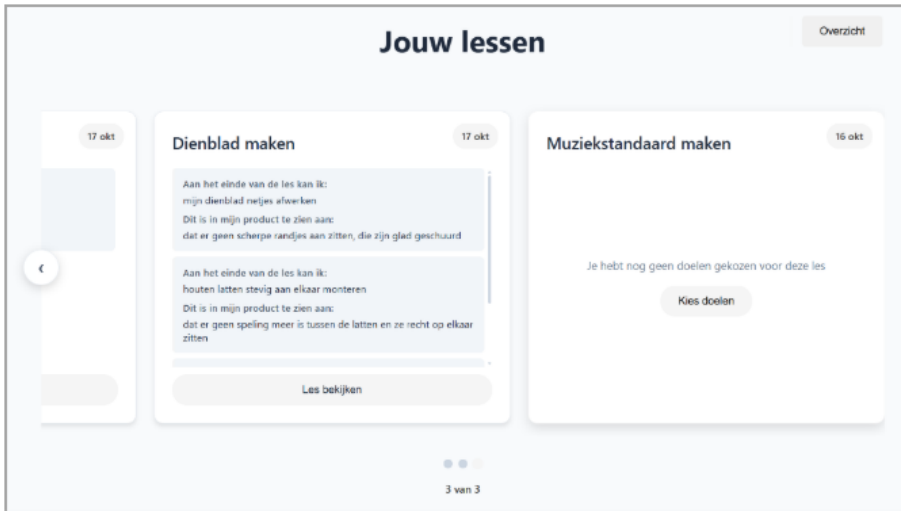
Startscherm
Het startscherm voor de leerling geeft een overzicht van de lessen waaraan een leerling heeft gewerkt en/of gaat werken (zie Figuur 1). In het huidige prototype staan drie opdrachten uit oude CSPEs van *Dienstverlening en producten*[2]; in een uitgewerkt prototype zal een leraar zelf lessen kunnen aanmaken. Voor afgeronde lessen

[1] In hoofdstuk 3.2 van de syllabus worden de exameneisen van dit toepassingsdomein beschreven.

[2] In deze opdrachten moesten verschillende producten gemaakt worden: een waxinelichthouder (CSPE 2024 vmbo-bb, versie rood, onderdeel C), een dienblad (CSPE 2023 vmbo-bb, versie blauw, onderdeel C) en een muziekstandaard (CSPE 2024 vmbo-bb, versie blauw, onderdeel C). De opdrachten en correctievoorschriften zijn te vinden op www.examenblad.nl.

is te zien aan welke doelen de leerling heeft gewerkt en voor nieuwe lessen kan de leerling doelen gaan opstellen.

Figuur 1: Startscherm voor de leerling



Doelen stellen

Een leerling start met een les door op “Kies doelen” te klikken bij de desbetreffende les. Als er bij een eerdere les al doelen zijn opgesteld, ziet de leerling dan een overzicht met de eerder gestelde doelen en kan deze al dan niet selecteren om aan te werken in de huidige les. Daarnaast kan de leerling één of meerdere nieuwe doelen opstellen en selecteren voor de huidige les. De user flow bij het opstellen van doelen is weergegeven in Figuur 2. Deze flow wordt onder de figuur in detail toegelicht.

Figuur 2: Leerling user flow voor het opstellen van doelen bij een les



a. *Kies een onderwerp*



b. *Stel een doel op a.d.h.v. de gegeven bouwstenen*



c. *Lees en verwerk feedback (meer iteraties mogelijk)*



d. *Kom tot een doel voor deze les*

Bij het opstellen van een nieuw doel start de leerling met het kiezen van een onderwerp (Figuur 2a). Dit geeft de leerling richting bij het nadenken over een goed doel. In het huidige prototype zitten onderwerpen die gebaseerd zijn op de correctievoorschriften van de voorbeeldopdrachten uit de CSPEs. Deze correctievoorschriften bevatten onderwerpen die gebaseerd zijn op het *eindproduct* en onderwerpen die gebaseerd zijn op het *proces*. In het huidige prototype richten we ons op eindproduct-onderwerpen. Dit maakt het eenvoudiger voor de leraar om aan het eind van de opdracht te beoordelen of de doelen behaald zijn, zonder dat de leraar de leerling continu moet observeren. Bovendien zal het opstellen van doelen rondom het eindproduct voor de leerling minder gericht zijn op cognitieve processen (zie uitgangspunt 4 in sectie 2.1). We nemen dit als startpunt voor het proof of concept met mogelijkheden voor uitbreiding in de toekomst.

Na het kiezen van een onderwerp gaat de leerling een doel opstellen (Figuur 2b). Om zicht te krijgen op hoe dit proces in de praktijk gaat, hebben we gesproken met verschillende leraren en onderwijskundigen. Daarnaast hebben we aan leerlingen uit drie klassen (brugklas vwo, brugklas havo en derde klas mavo) gevraagd om doelen op te stellen voor praktijkopdrachten waar ze die les aan zouden gaan werken. Hieruit kwam naar voren dat de doelen die leerlingen opstellen vaak niet specifiek zijn en niet altijd haalbaar binnen een opdracht. Vaak werden woorden als “beter” en “sneller” gebruikt[3], zonder te specificeren wat dit betekent of hoe dit zichtbaar wordt. Deze informatie hebben we meegenomen in zowel de opzet van het stellen van doelen als de feedback die door een AI gegeven wordt. Een doel in het prototype bestaat uit de volgende bouwstenen:

- Aan het eind van de les kan ik: […]
- Dit is in mijn product te zien aan: […]

Op die manier wordt de leerling gestuurd een doel te stellen dat haalbaar is binnen de les en dat specifiek genoeg is om te kunnen beoordelen of het doel behaald is. Als een leerling een doel heeft ingevuld en op “controleer” klikt, wordt het doel geëvalueerd met behulp van generatieve AI. Dit houdt in dat er wordt gecontroleerd of een doel specifiek genoeg is en of het haalbaar is binnen de opdracht. Als dat zo is, wordt het doel goedgekeurd en kan de leerling door met het opstellen van een nieuw doel óf met de opdracht starten. Als het doel nog niet goed genoeg is, zal de AI feedback geven (Figuur 2c). Op basis van deze feedback kan de leerling het gestelde doel aanpassen en opnieuw op “controleer” klikken. Het nieuwe doel wordt dan geëvalueerd en deze kan weer goedgekeurd worden of er wordt feedback gegeven. De leerling krijg maximaal vier keer feedback op een doel; daarna zal de AI hele gerichte feedback geven die de leerling kan (maar niet moet) overnemen. Het doel wordt na deze iteratie afgerond en opgenomen in het overzicht van doelen voor deze les.

Voor de evaluatie met generatieve AI is een systeemprompt gebruikt. De volledige systeemprompt is weergegeven in Tabel 1 in de Bijlage. De prompt bevat instructies ver hoe de evaluatie moet plaatsvinden, wat aandachtspunten zijn, op welk taalniveau feedback gegeven moet worden, voorbeelden van goed en minder goed geformuleerde doelen en voorbeelden van feedback teksten. Deze systeemprompt is iteratief ontwikkeld aan de hand van tests door het projectteam zelf en op basis van de verzamelde doelen in de drie vo-klassen. Met name het toevoegen van voorbeelden van doelen zorgt ervoor dat de AI de gestelde doelen adequater evalueert en betere feedback genereert. De voorbeelden in de huidige systeemprompt zijn gericht op opdrachten rondom houtbewerking. Voor toepassing in een ander domein zal het relevant zijn om domein specifieke voorbeelden toe te voegen; hiervoor zou een feature kunnen worden toegevoegd die de leraar in staat stelt om een set voorbeelddoelen te uploaden zodat de AI betere feedback geeft. Bovendien is onderzoek in de praktijk nodig, aanvullend aan dit project, om het evalueren van doelen en het geven van feedback met AI te verbeteren. Wat maakt dat een doel goed is? Hoe formuleren leerlingen doelen? Wat voor feedback zouden leraren hierop geven? (Hoe) verschilt dit voor verschillende domeinen? Met inzicht daarin kan de systeemprompt sterk verbeterd worden.

[3] Enkele voorbeelden van gestelde doelen die inzicht geven in wat voor type doelen leerlingen stelden zijn: “sneller werken”, “beter kunnen ontwerpen” en “leren documenten te ordenen”.

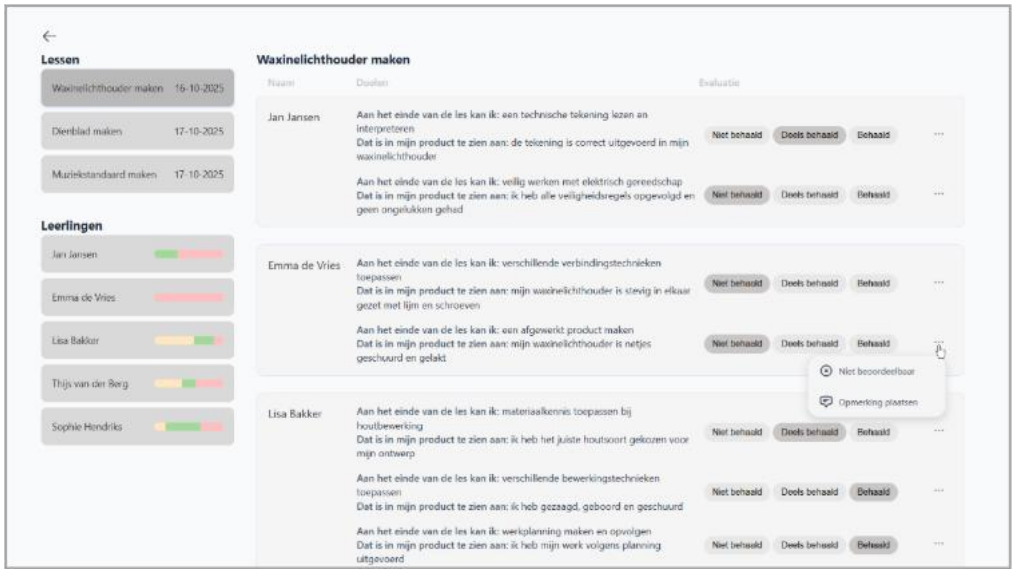
Beoordeling doelen door leerling en leraar

Wanneer de gewenste doelen opgesteld en geselecteerd zijn, gaat de leerling aan de slag met de opdracht. Dit vindt buiten de tool plaats. Na het uitvoeren van de opdracht komt de leerling terug naar de tool om te reflecteren op de gestelde doelen. Voor elk doel beoordeelt de leerling of het niet, gedeeltelijk of helemaal behaald is (zie Figuur 3). Als de leerling de evaluatie voltooid heeft, kan de leraar ook beoordelen of de leerling het doel behaald heeft. De leraar kan hierbij niet zien hoe de leerling het doel zelf beoordeeld heeft, zodat voorkomen wordt dat de beoordeling van de leerling de beoordeling van de leraar beïnvloedt. De leraar heeft dezelfde drie keuzes als de leerling: niet, gedeeltelijk of helemaal behaald (zie Figuur 4). Het kan echter voorkomen dat een doel niet goed geformuleerd is en daardoor niet te beoordelen is, doordat de leerling de feedback van AI niet overgenomen heeft, of omdat de feedback van AI niet duidelijk was; de leerling heeft bij het opstellen van doelen immers eigenaarschap over het gestelde doel. Als een doel niet te beoordelen is, kan de leraar dit aangeven in het uitklapmenu achter de drie puntjes. Hier is ook ruimte voor de leraar om een opmerking te plaatsen. Hoewel de knop hiervoor zichtbaar is, heeft het huidige prototype deze functionaliteit nog niet.

Figuur 3: Leerling overzicht om het behalen van de gestelde doelen te beoordelen.



Figuur 4: Leraar overzicht om het behalen van de gestelde doelen te beoordelen

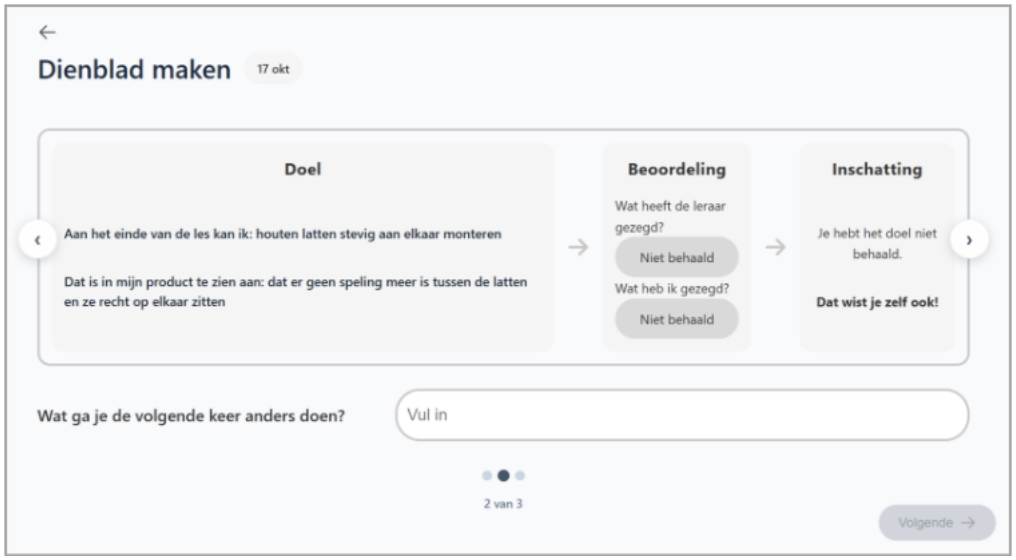


Vervolgstappen bepalen

Als zowel de leerling als de leraar de doelen beoordeeld heeft, kan de leerling aan de slag met het bepalen vervolgstappen. Voor elk doel van die les krijgt de leerling een overzicht van de overeenstemming tussen de beoordeling van de leerling en de leraar een bepaald doel. Als het doel niet of gedeeltelijk behaald is, krijg de leerling hierbij de vraag: “Wat ga je de volgende keer anders doen?”. Dit is weergegeven in Figuur 5. Als het doel behaald is, krijg de leerling de vraag: “Wat heeft ervoor gezorgd dat het lukte?”.

Wat de leerling hier invult wordt in het huidige prototype niet geëvalueerd of teruggekoppeld naar de leraar. Bij doorontwikkeling zouden we dit graag verder uitwerken. Hiervoor willen we eerst onderzoeken hoe dit nu in de praktijk gaat. Gaan leraren hierover in gesprek met leerlingen, en zo ja, gebeurt dit individueel, in groepjes of klassikaal? Het uitgangspunt is dat het bepalen van vervolgstappen een zekere mate van autonomie stimuleert bij de leerling. Idealiter kan met behulp van het prototype geëvalueerd worden of de leerling weet wat een goede vervolgstap is om zichzelf te verbeteren.

Figuur 5: Leerling scherm voor het bepalen van een vervolgstap per gesteld doel bij een opdracht



Dashboards

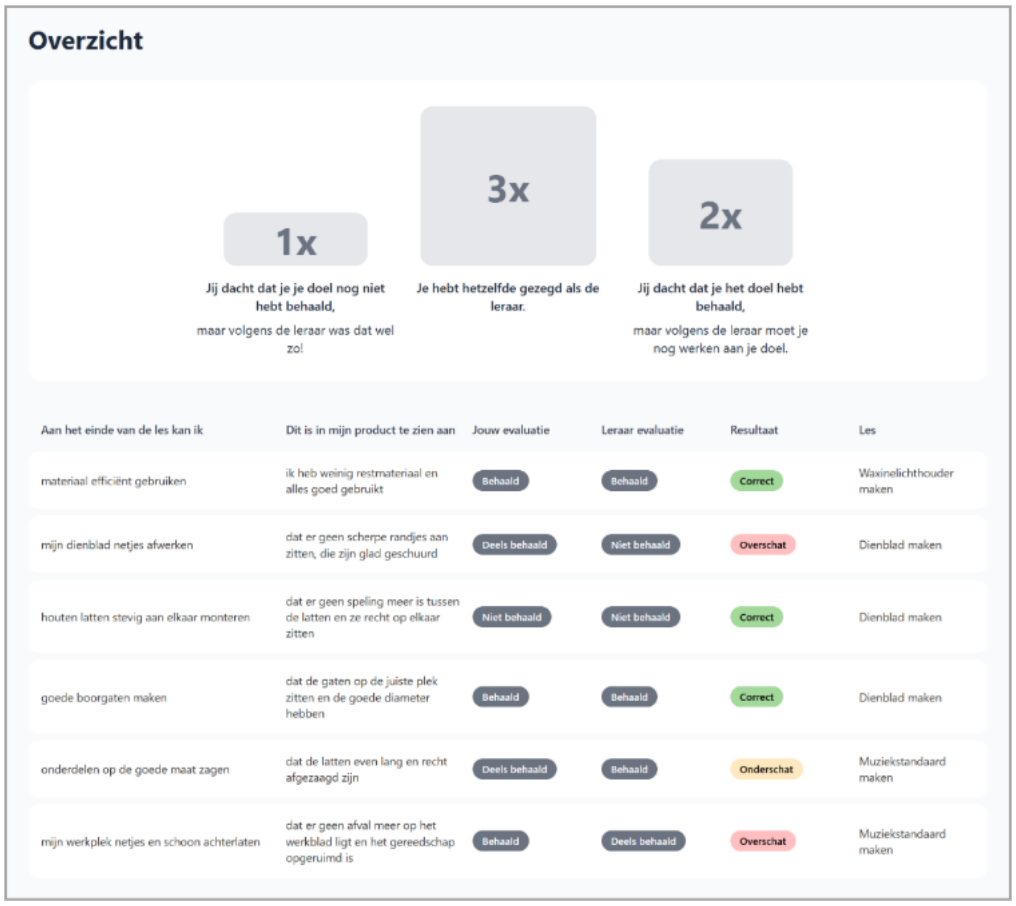
In de tool zitten twee verschillende dashboards. Een voor leerlingen en een voor de leraar. Hieronder lichten we toe hoe en waarvoor de dashboards gebruikt worden.

Leerling dashboard

De leerling kan resultaten op twee niveaus bekijken: per les of een totaaloverzicht met resultaten van alle afgeronde lessen. Een voorbeeld van een totaaloverzicht is weergegeven in Figuur 6. In de leerling dashboards staat een overzicht van de gestelde doelen per les, de beoordeling van de leerling en de leraar en de overeenstemming hiertussen. Doordat de beoordelingen op een 3-puntsschaal zijn, is informatie beschikbaar over onder- en overschatting. Als de leerling het behalen van een doel *lager* inschat dan de leraar, wordt dit gezien als een *onderschatting* (bijv. de leerling beoordeelt het doel als niet of gedeeltelijk behaald en de leraar als wel behaald). Als een leerling het behalen van een doel *hoger* inschat dan de leraar, wordt dit gezien als een *overschatting* (bijv. de leerling beoordeelt het doel als behaald en de leraar als niet of gedeeltelijk behaald). Op het dashboard is weergegeven hoe vaak een leerling zich onderschat, correct inschat en overschat. Omdat de woorden ‘onderschatten’ en ‘overschatten’ niet tot het basale woordenschatniveau horen, is hier in het

dashboard uitleg bij gezet (zie het staafdiagram in Figuur 6). Dit maakt het overzicht redelijk talig. Aanvullende gebruikerstesten zouden moeten uitwijzen of dit overzichtelijk en begrijpelijk is voor de doelgroep.

Figuur 6: Totaaloverzicht van het leerling dashboard



Leraar dashboard

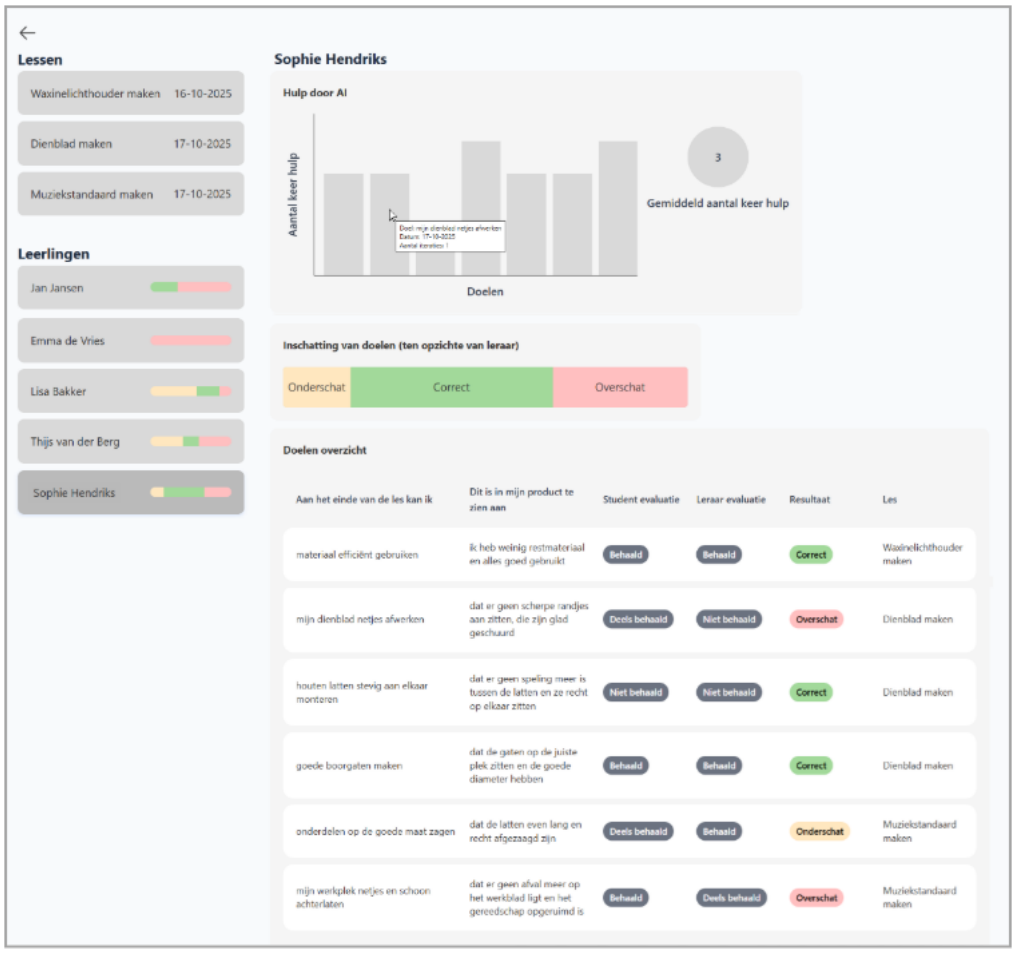
Voor de leraar is een dashboard ontworpen met een overzicht van de lessen en een overzicht van de leerlingen. Het overzicht van de lessen is het scherm waarin leraren het behalen van de gestelde doelen beoordelen. Deze zagen we eerder al in Figuur 4. Het overzicht van de leerlingen geef informatie over hoe goed de leerlingen doelen opstellen en hoe goed zij zelf kunnen inschatten of ze hun doelen behaald hebben. Een voorbeeld van dit dashboard is gepresenteerd in Figuur 7.

Om informatie te geven over hoe vaardig leerlingen zijn in het stellen van doelen, is per gesteld doel weergegeven hoe vaak de leerling hulp van AI heeft gehad. Hoewel dit geen directe meting van deze vaardigheid is, geeft het wel een indicatie: weinig iteraties betekent dat een leerling sneller zelf tot een specifiek en haalbaar doel is gekomen. In het diagram is te zien of het aantal iteraties afneemt over tijd. Wanneer met de muis over een staaf bewogen wordt, is te zien over welk doel het gaat, op welke datum deze opgesteld en beoordeeld is, en hoeveel iteraties feedback van AI zijn gebruikt bij het opstellen van het doel. Zoals eerder beschreven kunnen er per doel maximaal vijf iteraties zijn. Naast het aantal iteraties met AI, hebben we ook enige informatie over de kwaliteit van doelen doordat de leraar kan aangeven wanneer een doel niet te beoordelen is. Hierbij is de aanname dat als een doel te beoordelen is, het doel niet van voldoende kwaliteit is. In het dashboard zou een weergave kunnen worden toegevoegd die presenteert hoe vaak een leerling niet-beoordeelbare doelen heeft opgesteld.

Om informatie te geven over hoe vaardig leerlingen zijn in het inschatten of de gestelde doelen behaald zijn, geeft een gesegmenteerd staafdiagram weer bij hoeveel procent van de doelen een leerling zichzelf onderschat

(geel), correct inschat (groen) of overschat (rood). Dit overzicht geeft niet direct een beeld van de ontwikkeling over tijd. Wel kan de leraar onderaan in het overzicht van de doelen van een leerling scannen hoe de inschatting over tijd verandert. Een mogelijke aanvullende oplossing is het toevoegen van een datumfilter bij het gesegmenteerde staafdiagram (bijv. kies voor gegevens van de laatste week, laatste maand, laatste 3 maanden of het gehele schooljaar). In het huidige prototype is gekozen voor simpele weergaven, maar er zijn ook uitgebreidere, complexere visualisaties mogelijk met een hogere informatiedichtheid. Als die geïmplementeerd zouden worden, is het van groot belang om bij gebruikers te onderzoeken of de weergaven duidelijk en begrijpelijk zijn.

Figuur 7: Leraar dashboard met leerling resultaten



Technische specificaties

Het prototype is een front-end applicatie gebouwd in React en wordt gehost op Firebase. Gegevens worden uitsluitend in de local storage van de browser opgeslagen. Aangezien er geen inlogfunctionaliteit is, gaan gegevens verloren bij het gebruik van een ander apparaat, een andere browser of incognito modus. De local storage van de browser kan ook handmatig geleegd worden, direct in de browser of via de knop “Reset”. Via de back-end wordt een AI-model (GPT-4o) aangeroepen, dat binnen de beveiligde Azure-omgeving van Cito gehost wordt. Dit betekent dat gegevens die omgeving niet verlaten en ook niet worden gebruikt om het AI-model verder te trainen.

4. Overdraagbaarheid

In het prototype is gekozen om zelfsturing te beoordelen bij het maken van een concreet eindproduct door te beoordelen in hoeverre leerlingen doelen kunnen stellen, door de kalibratievaardigheid van leerlingen te beoordelen en door het bepalen van vervolgstappen te evalueren. De verschillende elementen van zelfsturing komen daarmee aanbod.

Het huidige prototype laat drie verschillende eindproducten zien waarvoor de leerling doelen kan stellen en kan beoordelen in hoeverre die doelen zijn behaald. Door de manier waarop deze eindproducten in het prototype een rol spelen, kan het prototype eenvoudig uitgebreid worden naar andere eindproducten die, bijvoorbeeld, in de CSPE's aan bod komen. Als het prototype omgezet wordt in een concrete applicatie, krijgt de leraar hiermee een instrument in handen voor formatief gebruik in de klas.

Bij de vaardigheid om doelen te stellen, wordt gebruik gemaakt van generatieve AI om feedback te geven. Daarvoor is een systeemprompt opgesteld, die ook terug te vinden is in de Bijlage. Een groot voordeel van deze manier van werken is dat de leerling directe feedback kan krijgen zonder dat dit de leraar extra belast. De prompt is gebaseerd op specifieke voorbeelden voor houtbewerking. Door deze voorbeelden te vervangen door voorbeelden uit andere domeinen, kan AI feedback geven bij andere eindproducten. Binnen dit project hebben we de kwaliteit van de feedback niet met leerlingen kunnen evalueren. Daarom is deze vorm van beoordeling in principe overdraagbaar, maar zal hij in de praktijk uitgebreid getest moeten worden, voordat hij toegepast kan worden. Ook om in lijn te blijven met de AI Act.

In het algemeen kan opgemerkt worden dat dit prototype laat zien dat ook meta-cognitieve vaardigheden in een digitale omgeving getoetst kunnen worden. Dit brengt een bijvoorbeeld grote administratieve voordelen met zich mee. Het prototype blijft alleen wel vaardigheid specifiek. De systematiek is weliswaar overdraagbaar, maar de concrete operationalisering is afhankelijk van de concrete vaardigheid.

5. Opzet evaluatie

Methode

Om te onderzoeken of het beoordelen van praktijkgerichte vaardigheden door middel van het prototype haalbaar is, zijn vignetten voorgelegd aan leraren in drie focusgroepen en zijn twee focusgroepen met schoolleiders gehouden. Door deze twee methoden te combineren, konden we inzicht krijgen in zowel gedeelde perspectieven tussen leraren van verschillende scholen als diepgaandere contextspecifieke ervaringen binnen scholen. In de analyse van de vijf focusgroepen is rekening gehouden met verschillen tussen het po en het vo en tussen leraren en schoolleiders.

Doel en functie van de vignetten
Een vignet is een beschrijvende schets van een praktijksituatie die bedoeld is om reacties, overwegingen, twijfels, voorkeuren en redeneringen van deelnemers op te roepen. In dit onderzoek namen we leraren mee in een fictieve, maar realistische situatie waarin een school werkt aan het meten van een praktijkgerichte vaardigheid.

Ontwikkelen vignetten
Het vignet werd ontworpen nadat het prototype grotendeels was afgerond, zodat de inhoud en formulering goed op elkaar aansloten. De gespreksleidraad is in Bijlage 2 te vinden, en wordt hieronder verder toegelicht.

Binnen het besproken prototype stond de voorbeeldopdracht houtbewerken centraal, waarbij leerlingen een muziekstandaard maken. Deze opdracht vraagt van leraren niet alleen dat zij beoordelen of leerlingen een kwalitatief goed product kunnen opleveren, maar ook dat zij nagaan in hoeverre leerlingen in staat zijn om doelen te stellen, te evalueren of deze doelen zijn behaald, en op basis daarvan passende vervolgstappen kunnen formuleren.

Focusgroep met schoolleiders
In de focusgroep met schoolleiders stond met name de praktische haalbaarheid van het meten van praktijkgerichte vaardigheden centraal. Ook doorliepen we kort het prototype. Deze focusgroep biedt inzicht in de bereidheid en (praktische) bezwaren van schoolleiders om praktische vaardigheden te toetsen. De gespreksleidraad is opgenomen in Bijlage 3.

Analyse
Na de afname ontwikkelden de onderzoekers een analysekader waarmee de gespreksverslagen systematisch konden worden geanalyseerd. Deze analyse maakte het mogelijk om patronen en verschillen in de perspectieven van leraren inzichtelijk te maken.

De bevindingen uit deze interviews zijn meegenomen in het analysekader en verwerkt in de rapportage. Daarbij is onderscheid gemaakt tussen de perspectieven van schoolleiders en leraren en tussen verschillende onderwijstypen, daar waar verschillen in opvattingen en aandachtspunten zichtbaar waren.

Focusgroepen

Voor de werving en uitvoering van de focusgroepen is de volgende aanpak gehanteerd.

Werving
Er is geworven via de directe kanalen en netwerken van de onderzoeksbureaus. Zo zijn berichten geplaatst op de LinkedIn kanalen van Kohnstamm Instituut, Oberon en Cito waarbij een oproep werd gedaan aan scholen

om deel te nemen. Daarnaast zijn warme contacten ingezet met scholen waar praktijkonderwijs aandacht krijgt en zijn zij benaderd met de vraag voor deelname. Verschillende verenigingen waarbij praktijkonderwijs de nadruk krijgt, zoals vereniging voor Agora, Monteressori en Jenaplan onderwijs als ook Stichting Platforms Vmbo zijn gevraagd om een oproep voor deelname aan het onderzoek te verspreiden. Daarnaast zijn scholen in verschillende regio's direct aangeschreven per mail met verzoek om deelname.
Waar eerst het plan was per school een schoolteam te werven, bleek dit in de praktijk voor scholen vaak lastig. Daarom is gekozen voor het organiseren van schooloverkoepelende focusgroepen.

Uitvoering
De vijf focusgroepen vonden plaats eind november en begin december 2025 via Teams. De focusgroepen duurden 1 uur. De gesprekken zijn opgenomen en getranscribeerd. Deze opnames worden vernietigd na 2 maanden. De focusgroep werd geleid door een onderzoeker uit het onderzoeksteam en de verslaglegging werd gedaan door een andere onderzoeker uit het team. In Tabel 1 staat een overzicht van de deelnemers aan de focusgroepen. In totaal hebben we acht schoolleiders en twaalf leraren gesproken, verdeeld over de sectoren.

Aan het eind van iedere focusgroep werden deelnemers geïnformeerd over wat er met de data gebeurt en dat zij op de hoogte worden gebracht zodra het rapport klaar is. Iedere deelnemer ontving een kleine attentie in de vorm van een cadeaubon van 25 euro.

Tabel 1. Overzicht deelnemers focusgroepen.

	Schoolleiders	Leraren	Totaal
Primair onderwijs	-	3	**3**
Praktijkonderwijs	-	1	**1**
vmbo-b/k/gl/tl	1	3	**4**
vmbo-tl	2	1	**3**
havo/vwo	1	1	**2**
vmbo/havo/vwo	3	3	**6**
MBO	1	-	**1**
Totaal	**8**	**12**	**20**

6. Uitkomsten evaluatie

In dit hoofdstuk presenteren we de resultaten met betrekking tot de focusgroepen. In de focusgroepen hebben we gevraagd naar de perspectieven van leraren en schoolleiders op de haalbaarheid van het gebruik van het prototype bij het meten van de vaardigheden van leerlingen doelen stellen, evalueren of doelen behaald zijn en vervolgstappen bepalen. Waar relevant maken we onderscheid tussen de opvattingen van onderwijsprofessionals in het po en vo, evenals tussen leraren en schoolleiders. Wanneer geen onderscheid wordt gemaakt, gaat het om een breed gedragen beeld.

Praktijkgericht onderwijs in de huidige situatie

Om genoemde aandachtspunten en meerwaarde met betrekking tot het prototype goed te kunnen duiden, hebben we allereerst gevraagd naar hoe praktijkgericht onderwijs in de huidige situatie op de scholen vorm krijgt, hoe op dit moment gewerkt wordt aan het opstellen en evalueren van doelen en wat het huidige AI-beleid is op scholen. Hieronder werken we de bevindingen op deze verschillende thema's verder uit.

Hoe krijgt praktijkgericht onderwijs momenteel vorm?

Uit de focusgroepen bleek dat de definitie van praktijkgerichte vaardigheden kan verschillen per onderwijstype en opvattingen van individuele onderwijsprofessionals of scholen. Zo waren deelnemers uit het po werkzaam op dalton- of montessorischolen. Binnen deze onderwijsconcepten is expliciet aandacht voor het ontwikkelen van vaardigheden als plannen, bijvoorbeeld via een weektaak. Daarnaast kom praktijkgericht onderwijs volgens deelnemers terug in bijvoorbeeld talentklassen en in lessen waarin leren expliciet wordt verbonden aan de praktijk, bijvoorbeeld door rekenactiviteiten buiten het klaslokaal te laten plaatsvinden.

In het vo gaven deelnemers aan dat praktijkgerichte vaardigheden regelmatig door de war gehaald worden met executieve functies en soft skills (zoals plannen en organiseren), en metacognitieve vaardigheden (zoals evalueren). Sommige scholen zien praktijkgericht onderwijs als onderwijs waarbij concrete materialen en situaties in het onderwijs betrokken worden of waarin leerlingen 'veldwerk' uitvoeren, net als in het po, zoals hierboven beschreven. Op een vwo-school van een van de deelnemers in de focusgroepen worden praktijkgerichte vaardigheden geassocieerd met 'hbo-vaardigheden' zoals presenteren, onderzoeken en samenwerken. Leraren in het vo geven aan dat het meest voor de hand liggend is om praktijkgerichte vaardigheden in te bedden in het praktijkonderwijs en op vmbo b/k, aangezien daar praktijkvakken gegeven worden. Tegelijkertijd signaleren schoolleiders dat er steeds meer aandacht komt voor onderzoeksvaardigheden en vaardigheden als samenwerken, waarbij de focus verschuift van uitsluitend kennisoverdracht naar de bredere ontwikkeling van de leerling.
In ieder geval was de tendens in de focusgroepen dat er behoefte is aan een duidelijke afbakening wat onder praktijkgerichte vaardigheden wordt verstaan.

Welke plek krijgen doelen stellen en evalueren op de scholen momenteel?

Hoe het doelen stellen en evalueren vorm krijgt binnen het onderwijs, verschilt eveneens per sector en per school. In het po komt in de focusgroepen terug dat dit een combinatie is van leraargestuurd en leerlinggestuurd. Zo worden lesdoelen bijvoorbeeld door de leraar vastgesteld, maar formuleren leerlingen wel zelf weekdoelen. Leerlingen reflecteren op hun leerproces via gesprekken of portfolio's. Deze reflectie is afgestemd op de leeftijd, wat volgens deelnemers belangrijk is. Jonge leerlingen kiezen bijvoorbeeld meer eenvoudige doelen, die gaan om bijvoorbeeld het afmaken van een taak. Oudere leerlingen formuleren meer persoonlijke doelen die gaan om bijvoorbeeld concentreren.

Binnen het praktijkonderwijs zijn de leerdoelen voor de leerlingen vastgelegd in onder andere werkkaarten. Leerlingen beoordelen zichzelf op deze werkkaarten, bijvoorbeeld op aspecten van veiligheid, leeropbrengst en

werkhouding. Vervolgens bespreken de leraar en leerling deze beoordeling samen. Deelnemers zijn positief over deze aanpak, omdat dit volgens hen bij de leerling tot meer inzicht in hun leerproces leidt.

Op het vmbo wordt door deelnemers genoemd dat er geïnvesteerd is in het werken met rubrics, gericht op vaardigheden en competenties. Het idee van de rubrics is dat dit voor leerlingen transparant maakt wat er van hen verwacht wordt. Zowel de leerling als de leraar vult de rubrics in. Tijdens een gesprek vergelijken de leraar en leerling hun beoordelingen met elkaar. Op mentorniveau wordt er gewerkt met doelenkaarten, waarin leerlingen per periode zelf hun doelen vastleggen. Deelnemers zien wel dat de kwaliteit van deze doelen verschilt per leerling.

Ook op havo/vwo zien deelnemers grote verschillen tussen leerlingen. Vaak hebben leerlingen moeite met het opstellen van doelen, waardoor dit over het algemeen niet gebeurt. Tegelijkertijd is het ingewikkeld om als leraar voor iedere leerling gepersonaliseerde doelen op te stellen en formatief te toetsen, zeker in grote klassen. Deelnemers zien op scholen dat de nadruk sterk ligt op toetsen en cijfers, in plaats van op formatief evalueren. Zij vinden dit wel belangrijk.

Schoolleiders in het vo noemen dat er binnen technasia veel waardevolle instrumenten en deskundigheid zijn ontwikkeld rondom het meten van praktijkgerichte vaardigheden, zoals reflecteren. Kennisoverdracht naar andere scholen lijkt echt nog beperkt. Voorbeelden van tools en programma's dit momenteel op vo-scholen gebruikt worden zijn OMZA (staat voor Organisatie, Meedoen, Zelfvertrouwen en Autonomie), Skills-meter, Remind Learning (om eigenaarschap bij leerlingen te ontwikkelen), RTTI (staat voor Reproductie, Training, Transfer, Inzicht & Innovatie) en Catwise (waarbij reflecteren en leerlinggerichte actieplannen centraal staan).

Huidig AI-beleid op school

Binnen het po wordt AI als minder urgent thema gezien. Deelnemers van po-scholen geven dan ook aan dat AI momenteel niet gebruikt wordt op school. Er wordt wel verkend wat de mogelijkheden zijn, maar hiervoor is behoefte aan expertise van buitenaf. Een van de deelnemers geeft aan een verschuiving te zien waarbij minder gebruik wordt gemaakt van technologische apparaten, zoals Chromebooks, en meer wordt gewerkt op papier.

In het vo wordt AI daarentegen wel breed gezien als urgent thema. Dit komt mede door het intensieve gebruik door leerlingen. Overwegend zit het AI-beleid op scholen in de ontwikkelfase. Veel scholen hebben (nog) geen vastgesteld of overkoepelend beleid, mede door de snelle ontwikkelingen rondom AI. Op sommige scholen bestaan al beperkte richtlijnen, bijvoorbeeld rond het gebruik van AI bij profielwerkstukken. Op meerdere scholen is een werkgroep actief met het ophalen van input over het gebruik van AI bij leraren en leerlingen om deze te vertalen naar beleidskaders. Het gebruik van AI wordt wel op verschillende manieren geagendeerd als gespreksonderwerp. Zo worden er informatie- en scholingsbijeenkomsten georganiseerd voor leraren, met aandacht voor verantwoord gebruik van AI en het verlagen van werkdruk met behulp van AI.

Algemeen perspectief op de tool

Zoals hierboven beschreven, bleek uit de focusgroepen dat er op verschillende scholen wel aandacht is voor het opstellen en/of evalueren van leerdoelen, maar dat scholen ook nog worstelen met de manier waarop dit vormgegeven moet worden. In de focusgroepen zijn we door het prototype heen gegaan en hebben vragen gesteld over de mogelijke meerwaarde en aandachtspunten om deze tool in te zetten op scholen. In deze paragraaf beschrijven we eerst het perspectief van deelnemers op het prototype in het algemeen. Daarna zoomen we verder in op de drie onderdelen: doelen stellen, doelen evalueren, vervolgstappen bepalen.

Tijdens de focusgroepen gaven deelnemers aan dat zij potentiële meerwaarde zien van het gebruik van het prototype op school. Een belangrijke voorwaarde die zij noemen is dat het 'doelen opstellen en evalueren' ingebed moet worden in het curriculum van de school, voordat de meerwaarde zichtbaar kan worden.

Deelnemers denken dat het prototype leraren aan het denken kan zetten over hoe leerdoelen en leerprocessen meer bij de leerlingen kunnen worden belegd, waardoor leerlingen meer eigenaarschap over hun leerproces krijgen.

Hoewel het prototype momenteel sterk taakgericht is, zien deelnemers vooral meerwaarde van het prototype bij de inzet voor vakoverstijgende doelen. Deze vakoverstijgende doelen zijn dan niet gekoppeld aan een specifieke les, maar kunnen bijvoorbeeld voor een langere periode opgesteld worden. Zo ontstaat er volgens deelnemers makkelijker een doorlopende leerlijn omdat reflectie over een langere periode meer samenhang kan krijgen en het leerproces cyclisch wordt in plaats van losstaand per les. Het zou daarbij wenselijk zijn dat de leraar of leerling binnen de tool de een tijdsperiode geselecteerd kan worden, waarbij de leerling of leraar bepaalt gedurende welke periode de leerling aan een doel gaat werken. Voorbeelden van vakoverstijgende doelen die genoemd werden, zijn: '*beter worden in het schrijven van teksten*', '*actiever vragen stellen tijdens de les*' of '*beter concentreren tijdens de les*'. Aan dergelijke doelen kan vakoverstijgend gewerkt worden en bijvoorbeeld binnen het mentoraat worden opgesteld en besproken. Bij vakoverstijgende doelen sluiten termen als 'les' en 'product' zoals deze nu in het prototype staan, niet aan.

Een risico van het werken met doelen per les, is volgens deelnemers dat leerlingen het lesdoel uit de methode letterlijk overnemen, bijvoorbeeld: '*Aan het einde van de les kan ik verschillen benoemen tussen klimaat en weer.*' Deelnemers geven juist aan dat zij het belangrijk vinden dat leerdoelen voortkomen uit de leerling zelf en dat er ruimte is voor wat de leerling persoonlijk wil ontwikkelen. Door te werken met vakoverstijgende doelen wordt het spanningsveld verminderd tussen ruimte voor eigen leerdoelen van leerlingen en de lesdoelen die door de school worden voorgeschreven.

Tot slot noemen deelnemers dat er binnen de tool gedifferentieerd moet kunnen worden in taalniveau van de leeftijd en het niveau van de leerling. Zo is het voor leerlingen in de onderbouw van het po belangrijk om rekening te houden met het feit dat zij nog niet allemaal kunnen lezen en schrijven. Er zou gewerkt kunnen worden met een voorleesfunctie of de leraar moet een ondersteunende rol innemen door vooraf geformuleerde doelen aan te bieden. Ook binnen het vo zien deelnemers verschillen. Zo wordt een voorbeeld genoemd: *"Een leerling uit de eerste klas van het vmbo vraagt om een ander taalniveau dan een leerling uit de vijfde klas van vwo."*

Doelen stellen

Tijdens de focusgroepen hebben we specifiek gevraagd naar welke meerwaarde deelnemers zien van het gebruik van het prototype bij het opstellen van doelen. Naast de meerwaarde, is ook een aantal aandachtspunten naar voren gekomen.

Meerwaarde

Over het algemeen zien deelnemers meerwaarde van het prototype bij het opstellen van doelen voor zowel leerlingen als leraren. Zo noemen zij dat de tool leerlingen stimuleert om actief te reflecteren en doelgericht leerdoelen te formuleren, in samenhang met het gewenste eindresultaat van de opdracht. Deelnemers geloven dat leerlingen daadwerkelijk beter kunnen worden in het opstellen van doelen door herhaald gebruik van het prototype. Dit komt volgens hen door verschillende factoren. Allereerst de inzet van AI in het geven van feedback. Waar leerlingen AI momenteel vaak nog ongericht gebruiken, biedt dit prototype juist een doelgerichte inzet van AI. Daarbij is het positief dat AI na vier pogingen zelf een leerdoel formuleert. Volgens deelnemers voorkomt dit dat leerlingen blijven hangen in de feedback en het formuleren van leerdoelen, wat ten koste kan gaan van hun motivatie. Ten tweede zien deelnemers meerwaarde van de mogelijkheid voor leerlingen om te kiezen over welk onderwerp zij een doel willen opstellen (bijvoorbeeld 'montage'). Het aanbieden van keuzemogelijkheden binnen lessen kan er namelijk aan bijdragen dat leerlingen bewuster keuzes maken, wat hun motivatie kan verhogen.

De meerwaarde voor leraren op dit punt is dat het genereren van feedback door AI tijd bespaart voor leraren. De leraar hoeft zo niet iedere leerling afzonderlijk te spreken, maar krijgen leerlingen toch feedback op maat. Een ander voordeel van de inzet van AI voor leraren is dat zij een minder corrigerende rol hoeven in te nemen. Deze corrigerende rol kan de relatie tussen leerling en leraar in de weg staan omdat de correctie dan persoonsgebonden wordt.

Aandachtspunten

Naast de meerwaarde die deelnemers zagen, werden er tijdens de focusgroepen ook verschillende aandachtspunten/risico's genoemd wanneer het prototype gebruikt wordt bij het opstellen van doelen. Deze lichten we hieronder verder toe.

1. De opgestelde doelen sluiten onvoldoende aan bij de lesinhoud.

Deelnemers zien een risico dat het leerdoel dat een leerling opstelt, inhoudelijk niet voldoende past bij de formele kerndoelen en/of de lesinhoud die de leraar op dat moment centraal wil stellen: "*Anders kan een doel een hele andere/verkeerde kant op gaan dan de docent voor ogen heeft."* Deelnemers vinden het wenselijk dat de leraar daar voldoende regie over houdt en vragen zich af in hoeverre AI in de tool intelligent genoeg is om daar selectief op te zijn. Een suggestie die in de focusgroepen werd genoemd om dit risico te ondervangen is dat de leraar aan het begin van de les de inhoud en bijbehorende kaders duidelijk maakt, eventueel in de vorm van een klassikale brainstormsessie. De leerlingen formuleren vervolgens binnen de gestelde kaders zelf hun leerdoelen.

2. Leerlingen focussen te veel op (een aspect van) het einddoel.

Bij het formuleren van leerdoelen maakt het prototype volgens deelnemers onvoldoende onderscheid tussen proces- en productcriteria. Hiermee bestaat het risico dat leerlingen minder nadenken over *hoe* zij het doel gaan bereiken en te veel focussen op het einddoel. Het prototype richt zich volgens deelnemers nu vooral op feedback en deels op feed forward vanuit de leerling, maar zij zouden het waardevol vinden om ook een expliciete feed-up stap toe te voegen. Zo kan duidelijk gemaakt worden welke criteria er zijn om het doel te behalen. Leraren vinden het belangrijk om over meerdere lessen heen een langetermijndoel helder te houden en tegelijkertijd concrete tussenstappen te behouden. Zo begrijpen leerlingen volgens hen beter hoe groei richting het doel eruitziet. In het prototype zou dit volgens deelnemers bijvoorbeeld vorm kunnen krijgen via een logboek of actieplan, waarbij na het formuleren van het doel de vraag volgt: 'hoe ga je dit aanpakken en welke stappen ga je zetten?' Dit zou het leerproces inzichtelijker maken.

Daarnaast vinden leraren het belangrijk dat leerlingen oog hebben voor meerdere aspecten van een eindproduct. Leerlingen moeten dus werken aan meerdere leerdoelen, in plaats van richten op één leerdoel/onderdeel. Een deelnemer illustreerde dit met: *"Bij het maken van een muziekstandaard vind ik het belangrijk dat deze niet wiebelt, maar ik vind het ook belangrijk dat het standaard er mooi uitziet."* Praktisch gezien vinden deelnemers het daarbij belangrijk dat de (tussen)doelen van verschillende lessen, vakken of thema's per leerling niet door elkaar staan.

3. Leerlingen worden gemakzuchtig.

Omdat AI het leerdoel uiteindelijk zal formuleren, zien deelnemers een risico dat leerlingen snel afhaken in het *zelf* formuleren van leerdoelen of gemakzuchtig gedrag vertonen bij het opstellen van leerdoelen. Zo denken deelnemers dat een deel van de po-leerlingen geneigd zijn om minimale input te geven, herhaaldelijk feedback van AI te ontvangen totdat AI uiteindelijk een leerdoel geeft. In het vo denken deelnemers dat een deel van de leerlingen een andere AI tool zal gebruiken om hun eerste leerdoel op te stellen. In de focusgroepen wordt de suggestie gedaan om de mogelijkheid in te bouwen dat leraren AI handmatig in of uit kunnen schakelen, zodat leerlingen in de ene les wel feedback van AI ontvangen en in de andere les niet. Zij denken op deze manier de aandacht van de leerlingen vast te houden.

Daarnaast kunnen leerlingen gemakzuchtig worden, omdat ze het opstellen van doelen op den duur teveel als een invuloefening beschouwen. Naar mate ze de tool vaker gebruiken, weten leerlingen steeds beter wat er van hen verwacht wordt. Deelnemers zien een risico dat leerlingen hun leerdoelen daarop gaan aanpassen in plaats van oprecht te reflecteren.

4. Onjuiste feedback van AI door verschillende taal en terminologie.

Onder deelnemers leeft de vraag in hoeverre AI gevoelig is voor het verkeerd interpreteren van leerdoelen wanneer leerlingen spelfouten maken of afwijkende terminologie gebruiken en in hoeverre dit invloed heeft op de kwaliteit van de feedback. Het is volgens deelnemers positief dat AI geen expliciete feedback geeft op spelling en grammatica, omdat dit afleidend kan zijn bij het opstellen van een leerdoel, mits dit niet ten koste gaat van de kwaliteit van de feedback. Daarnaast moet de terminologie die AI gebruikt en die leerlingen in hun leerdoelen gebruiken, aansluiten bij de terminologie die op school gehanteerd wordt. Een deelnemer legt uit: "*Op onze school gebruiken we bij rekenen de term 'verliefde harten', terwijl andere scholen kunnen spreken over 'de vriendjes van 10'. Het gaat hierbij om hetzelfde rekenconcept, maar het is belangrijk dat op school de eigen terminologie wordt gebruikt en dus ook in de doelen wordt verwerkt."*

Evalueren of doelen behaald zijn

Tijdens de focusgroepen vroegen we de deelnemers naar welke mogelijkheden voor het evalueren van doelen het prototype biedt. Daarnaast bespraken we mogelijke knelpunten en risico's wanneer leerlingen via het prototype hun leren op deze manier evalueren.

Meerwaarde

Deelnemers zien op twee punten vooral de meerwaarde van het prototype wat betreft het evalueren van de doelen. Zo vinden leraren het cruciaal om het gesprek te voeren over verschillen en overeenkomsten in de evaluatie van het leerdoel. Volgens deelnemers zijn ruwe gegevens op zichzelf niet veel waard, maar gaat het juist om het gesprek. Zij zien de meerwaarde van het prototype als ondersteuning bij zulke evaluatiegesprekken over voortgang.

Daarnaast kan het terugkijken op behaalde doelen voor leerlingen succeservaringen opleveren en motiverend werken. Zeker in grotere klassen kan het persoonlijk contact tussen leraar en leerling beperkt zijn. Daar biedt het prototype een manier om waardering en complimenten van leraren beter zichtbaar te maken.

Aandachtspunten

Ook werden verschillende aandachtspunten/risico's genoemd rondom dit onderdeel, die we hieronder verder toelichten.

1. Te weinig diepgang in de evaluatie.

Doordat een leerling nu enkel beoordeeld of een doel niet, deels of wel behaald is, missen deelnemers een bepaalde diepgang in het evalueren van de doelen. Leraren zouden voor zowel de leerling als de leraar ruimte willen hebben voor een toelichting op de evaluatie. Een van de suggesties waarop diepgang gecreëerd kan worden is door een keuzemenu toe te voegen met redenen waar leerlingen uit kunnen kiezen waarom een doel niet, deels of wel behaald is. Deelnemers zien het risico dat leerlingen bij een open vraag oppervlakkig blijven antwoorden, zoals *"lekker gewerkt."*

Een andere suggestie om diepgang te creëren was om evaluatievragen toe te voegen. Daarbij is een positieve insteek van belang volgens deelnemers, door niet alleen te focussen op wat beter kan, maar ook te focussen op wat al wél goed is gegaan. Leerlingen vinden het zelf ingewikkeld om te evalueren, dus gerichte vragen zijn hierbij helpend. Voorbeelden die genoemd werden, zijn:

- Niet behaald: Je hebt het doel nu niet behaald, wat kan jou de volgende keer helpen? Wat ging er al wel goed? Wat is nog niet gelukt?
- Deels behaald: Wat is er wel gelukt? Wat ging er goed? Wat is nog niet gelukt? Hoe ga je hier de volgende keer aan werken?
- Wel behaald: Wat ging er goed? Waar ben je trots op? Wat neem je mee naar de volgende keer? Wat zou je de volgende keer anders doen? Welke lessen heb je geleerd die je in de toekomst ook gaat toepassen?

Deelnemers geven aan het belangrijk te vinden dat deze reflectiestap direct plaatsvindt na het evalueren van de doelen, in plaats van (zoals het nu in het prototype is) na de beoordeling van de leraar.
Tot slot wordt genoemd dat het evalueren van doelen op dit moment nog sterk cognitief ingestoken is en dat er weinig aandacht is voor de beleving en emoties van leerlingen tijdens het leerproces. Dit vinden de deelnemers wel belangrijk. De hierboven genoemde suggesties om diepgang te creëren kunnen dus ook focussen op emoties en beleving. Daarnaast werd de suggestie genoemd om leerlingen met behulp van smileys hun gevoel aan te laten geven. Zo krijgt de leraar ook beter zicht in hoe de leerling het leerproces ervaart.

2. Te grote belasting voor leraren: toename van werkdruk.

Deelnemers zien een risico in een grote toename van werkdruk bij leraren. Volgens hen is het niet haalbaar om alle doelen van elke leerling bij elke les te beoordelen. Zeker gezien het aantal lessen per dag in het vo in combinatie met de beperkte lestijd (“*dat zijn dan 125 invoermomenten per dag”*). Dit zorgt volgens deelnemers voor een te grote tijdsbelasting naast de reguliere lesactiviteiten. Het evalueren van leerdoelen op langere termijn, bijvoorbeeld per periode, sluit beter aan bij de praktijk. Een andere voorgestelde oplossing hiervoor is om leerlingen tijdens groepsopdrachten gezamenlijk doelen te laten formuleren, en deze ook samen te laten evalueren.

3. De evaluatie zegt weinig bij te eenvoudige doelen.

Deelnemers zien een risico wanneer leerlingen onzeker zijn over hun eigen vaardigheden, zij te eenvoudige leerdoelen opstellen. Op die manier doen zij sneller succeservaringen op, want de leraar zal het leerdoel sneller beoordelen als ‘behaald’. Daarnaast noemen deelnemers dat dit ook kan spelen bij leerlingen met een verschillend niveau. Zij geven aan differentiatie tussen leerlingen van belang vinden. De ene leerling heeft meer uitdaging nodig dan de andere leerling en zij zullen dus andere leerdoelen op moeten stellen. AI moet hier volgens hen rekening mee houden in de feedback. Deelnemers vragen zich af wat in die gevallen de toegevoegde waarde is van deze informatie en hoe dit bijdraagt aan het leerproces van leerlingen.

4. Leerdoelen ingewikkeld te beoordelen.

In sommige gevallen kan het voor leraren ingewikkeld zijn om te beoordelen of een leerdoel behaald is. Zij hebben hier niet voor elke leerling evenveel zicht op. Zo legt een van de deelnemers uit: “*Jan Jansen had bijvoorbeeld zes doelen en werkte heel stil in een hoek van de klas, hoe kun je dan al die doelen beoordelen als docent?*”.

Ook zijn doelen gericht op soft skills, zoals samenwerken, ingewikkeld te beoordelen. Er is onduidelijkheid wanneer zo’n doel als volledig behaald geldt. De beoordeling is leraarafhankelijk, waarbij de ene leraar strenger beoordeelt dan de andere. Om te bepalen of een doel bereikt is, kan een onderliggende rubric helpend zijn, waarin duidelijk omschreven staat wat als niet, deels, of wel behaald wordt beschouwd Een rubric draagt bij aan een gedeelde basis voor consistente beoordeling.

Vervolgstappen bepalen

Tijdens de focusgroepen gaven deelnemers aan welke resultaten zij als school graag zien van verschillende leerlingen en in hoeverre dit verschilt van het huidige overzicht in het toetsmodel.

Meerwaarde

Deelnemers zien in het huidige prototype (nog) geen duidelijke meerwaarde voor het bepalen van vervolgstappen. Zij erkennen wel het belang van het bepalen van vervolgstappen, maar niet op de manier waarop dit in het huidige prototype gebeurt. De grootste meerwaarde van de tool in het algemeen vinden deelnemers de inzet van AI. Een suggestie om de tool ook bij het bepalen van vervolgstappen meerwaarde te laten hebben, is om AI ook in te zetten bij deze stap. Door leerlingen gerichte adviezen te geven op basis van hun evaluatie, wordt het leerproces van de leerling ondersteund en wordt de leraar tegelijkertijd ontlast in het begeleiden van de leerlingen.

Aandachtspunten

Er worden weer verschillende aandachtspunten/risico's genoemd, die we hieronder toelichten.

1. De leergroei wordt onvoldoende zichtbaar en gestimuleerd.

Wanneer een leraar het doel als behaald heeft beoordeeld en de leerling als (deels) behaald, krijgt de leerling een vraag als '*wat heeft ervoor gezorgd dat het lukte*'. Deelnemers zijn van mening dat deze vraag geen vervolgstap is, terwijl het ook in deze situatie belangrijk is dat de leerling vooruitkijkt. Uit een behaald leerdoel ontstaat vaak een nieuw doel. Deelnemers geven aan dat er behoefte is aan een manier van evalueren die meer aansluit bij een *growth mindset,* in plaats van een summatieve benadering "*behaald, klaar".* Deelnemers vinden het belangrijk dat in de tool duidelijk wordt dat leren een doorlopend proces is. Zij doen een suggestie voor een volgens hen logische vervolgvraag: '*Waar ga je je nu op richten?*' eventueel met het verzoek om direct een nieuw leerdoel te formuleren. Dit is helpend voor de leerling, om het nut en de samenhang van het leerproces beter te begrijpen en ook om de eigen ontwikkellijn inzichtelijk te maken. Maar ook voor de leraar kan dit helpend zijn om de ondersteuning op af te stemmen.

2. Kleurindicaties leiden tot verwarring

Er worden verschillende aandachtspunten genoemd met betrekking tot de kleurindicaties. Allereerst interpreteren de deelnemers de kleuren op het eerste oog als *leerdoel niet, deels of wel behaald*, terwijl deze in werkelijkheid aangeven of de beoordeling van de leerling overeenkomt met die van de leraar. Dit onderscheid is niet duidelijk en leidt tot verwarring. Daardoor biedt het overzicht voor de leraar en leerling geen inzicht in de inhoudelijke groei op leerdoelen, terwijl leraren dit wel wenselijk vinden.

Ten tweede wordt *overschat* momenteel in het rood weergegeven en *onderschat* in het oranje. Deze kleurkeuzes kunnen een oordelende of summatieve indruk wekken, terwijl deelnemers dit niet passend vinden wanneer het gaat om het reflecteren op doelen. Minder oordelende kleuren in het prototype zouden wenselijker zijn.

Tot slot zien deelnemers het risico dat de tool te veel gaat lijken op bestaande digitale methodes, waarin de resultaten vooral zichtbaar zijn via de kleuren. Hierdoor verschuift volgens hen de aandacht naar het behalen van vinkjes in plaats van naar het leerproces.

3. Opvolging bij een afwijkende boordeling is leraarafhankelijk.

Binnen het huidige prototype is er geen moment waarop leerlingen expliciet reflecteren wanneer hun beoordeling afwijkt van de beoordeling van de leraar. Deelnemers geven aan dat dit essentieel is voor het kwaliteitsbesef bij leerlingen: bij overschatting kan bijvoorbeeld onduidelijk zijn wat 'goed genoeg' is, ligt het aan gemakzucht bij de leerling of komt dit door onduidelijkheid door de leraar? Dat is het punt waar het gesprek tussen leerling en leraar zou moeten starten, maar dat ontbreekt momenteel. Doordat hier nu geen vaste werkwijze voor is, wordt de opvolging van ongelijke beoordelingen leraarafhankelijk. In het prototype zou bijvoorbeeld het volgende opgenomen kunnen worden: '*je beoordeling komt niet overeen met die van de leraar, ga naar je leraar toe en bespreek dit*.' Hierbij bestaat wel het gevaar dat er onrust in de klas ontstaat.

7. Conclusies en aanbevelingen

In dit hoofdstuk worden de onderzoeksvragen beantwoord en wordt een integrale analyse gedaan van het onderzoek. Deze analyse resulteert in een serie aanbevelingen en richtlijnen voor beleid en praktijk.

Hoewel “praktische vaardigheden” in onderwijscontexten vaak worden geassocieerd met concrete psychomotorische handelingen zoals een fietsband repareren of sociaal-interactieve handelingen zoals een productklacht indienen, is in dit project gekozen voor een andere invalshoek. Wij zijn uitgekomen bij een vaardigheid die binnen praktische taken zichtbaar wordt, maar niet gebonden is aan één domein, en die dus breed toepasbaar is in het funderend onderwijs. Vanuit deze overweging is kalibratievaardigheid (zelfbeoordeling) geselecteerd: het vermogen van leerlingen om hun eigen prestaties accuraat in te schatten en hun leerproces daarop bij te sturen. Deze vaardigheid vormt een kerncomponent van zelfgereguleerd leren (Winne & Hadwin, 1998) en is daarmee essentieel voor het stellen van realistische doelen, het benutten van feedback en het ontwikkelen van duurzame leervaardigheden.

Door kalibratie te meten binnen praktische taakcontexten – zoals opdrachten in praktijkgerichte vakken – ontstaat een concreet en toetsbaar kader waarin leerlingen leren reflecteren op hun handelen en de kwaliteit van hun werk. Dit maakt kalibratie zowel een praktische vaardigheid in uitvoering als een metacognitieve vaardigheid in denken. De keuze om kalibratievaardigheid centraal te stellen biedt zo mogelijkheden om het leren in uiteenlopende domeinen te versterken én om leerlingen beter voor te bereiden op vervolgonderwijs en beroepspraktijk, waarin zelfbeoordeling en bijsturing cruciale voorwaarden zijn voor levenslang leren.

Hoe ziet dit eruit als proof of concept voor één praktijkgerichte vaardigheid?

Het prototype levert een helder en realiseerbaar proof of concept voor het meten van de vaardigheden doelen stellen, zelfbeoordeling en vervolgstappen bepalen in een houtbewerkingscontext. Naar verwachting maakt de combinatie van vaste bouwstenen voor doelen en iteratieve AI-feedback (o.b.v. domeinspecifieke voorbeelden in de systeemprompt) het voor leerlingen concreet om specifieke, haalbare en meetbare eindproduct-doelen te formuleren. Bij doorontwikkeling is het essentieel om ook leerlingen mee te nemen in het ontwerpproces. Als blijkt dat het voor hen te moeilijk is om op deze manier doelen te stellen, kunnen alternatieven worden ingezet. Zo zou in een beginfase een set van doelen te gegeven kunnen worden waar een leerling uit kan kiezen, of een visuele aanpak waarbij een leerling foto’s van eindproducten ziet en kan aangeven welke producten goed zijn en waarom.

In het prototype wordt het aantal keer hulp van AI gebruikt als indicatie voor de vaardigheid *doelen stellen* (en evt. een signaal van de leraar als een doel niet beoordeelbaar is). Dit geeft echter niet direct informatie over de kwaliteit van de gestelde doelen. In vervolgonderzoek kunnen we de automatische evaluatie van de kwaliteit van gestelde doelen onderzoeken, bijvoorbeeld in termen van specificiteit en haalbaarheid, zodat daar informatie over kan worden verzameld en teruggekoppeld aan de gebruikers.

Evaluatie van de vaardigheid *zelfbeoordeling* is o.a. afhankelijk van de beoordeling van de leraar. Het is hierbij van belang dat de leraar een betrouwbare inschatting kan maken over het al dan niet behalen van de doelen. We verwachten dat wanneer een doel goed geformuleerd is, een leraar in staat is om dit te evalueren. Omdat het om een formatieve context gaat, is een extra beoordelaar niet nodig.

Om de vaardigheid *vervolgstappen bepalen* te evalueren, is aanvullend onderzoek nodig. We willen onderzoeken hoe leraren hier in de praktijk mee omgaan en hoe ze samen met leerlingen reflecteren op behaalde en nog niet behaalde doelen. Op basis van die informatie kunnen we de evaluatie van vervolgstappen verder uitwerken.

Overdraagbaarheid

In het onderzoek kwam naar voren dat de systematiek van het prototype goed overdraagbaar is naar andere vaardigheden. Het prototype kan dan ook eenvoudig uitgebreid worden naar andere vaardigheden, omdat doelen stellen, evalueren en het formuleren van vervolgstappen ook op andere vaardigheden kan worden toegepast. De inhoudelijke uitwerking is alleen vaardigheid specifiek.

Praktijkgerichte vaardigheden in de dagelijkse praktijk van het onderwijs.

Om te onderzoeken in hoeverre het prototype aansluit bij de wensen van de onderwijspraktijk, zijn vijf focusgroepen afgenomen met leerkrachten/leraren en met schoolleiders uit po en vo. De focusgroepen laten zien dat praktijkgericht onderwijs in het primair en voortgezet onderwijs op uiteenlopende en vaak niet eenduidige manieren wordt vormgegeven. Er is daarom behoefte aan een duidelijke afbakening en een gemeenschappelijke taal. Ook de manier waarop doelen worden gesteld en geëvalueerd varieert sterk. In het po en praktijkonderwijs is meer aandacht voor leerlingreflectie en formatief werken, terwijl dit in havo/vwo lastig blijkt door tijdgebrek, grote klassen en de nadruk op toetsen en cijfers. Het AI-beleid verschilt eveneens: in het po speelt AI nauwelijks een rol, terwijl het vo dit als urgent ziet, maar het beleid zich daar nog in een ontwikkelfase bevindt.

Deelnemers zien duidelijk de meerwaarde van het prototype als het structureel wordt ingebed in het curriculum. Het kan leraren stimuleren om leerdoelen en leerprocessen meer bij leerlingen te beleggen, wat het eigenaarschap van leerlingen vergroot. De meeste meerwaarde wordt gezien bij vakoverstijgende doelen die over een langere periode worden opgesteld, waardoor reflectie samenhangender wordt. Dit verkleint ook het risico dat leerlingen standaard lesdoelen kopiëren. Daarnaast is differentiatie in feedback noodzakelijk, passend bij leeftijd en onderwijsniveau, om de tool toegankelijk en effectief te maken voor alle leerlingen.

Het prototype kan de vaardigheid doelen stellen effectief ondersteunen, doordat het leerlingen stimuleert tot een gerichte formulering van leerdoelen en reflectie daarop. Als de tool vaker wordt ingezet, kan dit leerlingen helpen deze vaardigheid te ontwikkelen. Wel is het belangrijk dat de leraar de aansluiting bij de leerdoelen waarborgt. Zonder aanvullende ondersteuning bestaat het risico op het formuleren van oppervlakkige doelen, op een eenzijdige focus op het eindproduct, op gemakzuchtig gedrag. Ook zijn respondenten beducht voor onjuiste AI-feedback.

Bij het evalueren van leerdoelen is de tool vooral nuttig als ondersteuning van het gesprek tussen leraar en leerling en als middel om succeservaringen zichtbaar te maken en motivatie te vergroten. De huidige manier van evalueren geeft nog niet genoeg informatie: meer reflectie, gerichte vragen, aandacht voor emoties en directe terugkoppeling zijn nodig om diepgang te creëren. Daarnaast zijn er zorgen over werkdruk voor leraren, haalbaarheid in grote klassen, het beperkte nut van te eenvoudige doelen en de moeilijkheid van consistente beoordeling, vooral bij soft skills. Rubrics en evaluatie over langere perioden worden als essentieel gezien.

Deelnemers aan de focusgroepen onderschrijven het belang van het formuleren van vervolgstappen, maar zij zien weinig meerwaarde in de huidige uitwerking. Daarvoor is meer een focus op het doorlopende leerproces nodig: na het behalen van een individueel doel wordt te weinig vooruitgekeken. Deelnemers pleiten voor een growth-mindsetbenadering, waarbij behaalde doelen leiden tot nieuwe doelen. Daarnaast is het kleurgebruik voor sommige deelnemers verwarrend. Ook is onduidelijk wat er moet gebeuren als er verschillen zijn tussen leerling- en leraarbeoordelingen, waardoor reflectie en begeleiding te afhankelijk zijn van de interpretatie van de leraar.

Samengevat kan uit de focusgroepen geconcludeerd worden dat het prototype positief is ontvangen en dat het vooral ingezet moet worden op het niveau van leerlijnen. De tool kan eigenaarschap en reflectie bevorderen,

maar vraagt om goede afstemming door leraren, differentiatie en verdiepende reflectie. Voor evalueren en vervolgstappen is verdere doorontwikkeling nodig, met aandacht voor haalbaarheid, onderscheid po en vo, werkdruk, growth mindset en voor consistente begeleiding.

Concrete aanbevelingen

Dit onderzoek begon met de vraag hoe een toetsmodel kan worden ontworpen waarmee praktijkgerichte vaardigheden op valide, betrouwbare en haalbare wijze meetbaar en toetsbaar worden gemaakt, zodat deze vaardigheden kunnen worden geïntegreerd in het bredere systeem van leerlingvolgsystemen en toetsgestuurde overgangsbeslissingen in po, vmbo, have en vwo. In het onderzoek hebben we ons met name gericht op het ontwikkelen van een prototype en het evalueren daarvan met focusgroepen van leraren en schoolleiders. Op basis van onze ervaringen tijdens het ontwikkelen van het prototype en op grond van de feedback komen we tot een aantal aanbevelingen:

1. Zorg voor een duidelijke en gemeenschappelijke taal met betrekking tot praktijkgerichte vaardigheden in po en vo.
2. Investeer in de doorontwikkeling van het prototype om het meten van praktijkgerichte vaardigheden op grotere schaal te kunnen implementeren.
3. Verfijn het toetsmodel om te differentiëren naar leeftijd, schoolniveau en schooltype.
4. Generaliseer het toetsmodel naar toepassing op het niveau van leerlijnen en curricula.
5. Verken het gebruik van AI om op maat feedback te kunnen geven aan leerlingen.
6. Ontwikkel rubrics en reflectie vragen om het gesprek van de leraar met de leerling te stimuleren en te standaardiseren.

Tot slot

Na een theoretische verkenning is een prototype ontwikkeld om te illustreren op welke manier het toetsmodel concreet ingezet kan worden in de praktijk. Het prototype illustreert hoe praktijkgerichte vaardigheden met behulp van digitale toetsing en AI op een schaalbare manier kunnen worden geëvalueerd. De evaluatie laat zien dat het prototype een veelbelovende stap is, maar dat doorontwikkeling nodig is om tot een vakoverstijgende tool te komen met een eenduidig begrippenkader, voorzien van flexibele AI ondersteuning, rubrics en reflectievragen. Op die manier kunnen praktijkgerichte vaardigheden consistent geëvalueerd worden, geïntegreerd worden in lvs'en en een plek krijgen bij het nemen van overgangsbeslissingen, met een beheersbare werkdruk voor het onderwijs.

Enschede, Utrecht, Amsterdam en Arnhem
19 december 2025.

8. Referenties

Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education, 26*, 247–274. https://doi.org/10.1080/158037042000225245

Jossberger, H., Brand-Gruwel, S., Van de Wiel, M. W., & Boshuizen, H. P. (2020). Exploring students' self-regulated learning in vocational education and training. *Vocations and Learning*, *13*(1), 131-158. https://doi.org/10.1007/s12186-019-09232-1

Khaled, A., Gulikers, J., Biemans, H., & Mulder, M. (2016). Occurrences and quality of teacher and student strategies for self-regulated learning in hands-on simulations. *Studies in Continuing Education*, *38*(1), 101-121. https://dx.doi.org/10.1080/0158037X.2015.1040751

Kirschner, P. A., & Stoyanov, S. (2020). Educating youth for nonexistent/not yet existing professions. *Educational Policy*, *34*(3), 477-517. https://dx.doi.org/10.1177/0895904818802086

Mejeh, M., & Grieder, C. (2025). Educational interventions to promote self-regulated learning in vocational schools - A systematic review. *International Journal for Research in Vocational Education and Training*, *12*(2), 192-235. https://doi.org/10.13152/IJRVET.12.2.3

Măirean, C., & Turliuc, M. N. (2013). The better-than-average effect in evaluating sports performance. Eur*ope's Journal of Psychology*, *9*(1), 117–125.

Oh, S.-L., Liberman, L., & Mishler, O. (2017). Faculty calibration and students' self-assessments using an instructional rubric in preparation for a practical examination. *European Journal of Dental Education, 22.* https://doi.org/10.1111/eje.12318

Horrey, W. J., Lesch, M. F., Mitsopoulos-Rubens, E., & Lee, J. D. (2015). Calibration of skill and judgment in driving: Development, threats to validity, and implications for training and assessment. Accident Analysis & Prevention, 81, 123–130.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Earl, L. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Panadero, E., Brown, G. T. L., & Strijbos, J.-W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2

Bijlage 1. Systeemprompt feedback opstellen doelen

Je bent een coach voor leerlingen uit het voortgezet onderwijs (vmbo, ongeveer 14 jaar). Ze volgen een praktijkvak en formuleren hun leerdoel altijd met de zin: "Aan het eind van de les kan ik … . Dat kun je in mijn product zien aan … ."

Jouw taak
Je begeleidt leerlingen stap voor stap bij het aanscherpen van dit leerdoel.

Jouw aanpak
Je ontvangt steeds het leerdoel van een leerling, met daarvoor de telling.
Beoordeel het leerdoel op specificiteit.
Stel één sturende vraag om het doel concreter of duidelijker te maken wanneer het niet specifiek genoeg is.
Is het doel specifiek genoeg? Controleer of het haalbaar is binnen één les. Gebruik hiervoor indien beschikbaar de meegegeven metadata, syllabus, vak, thema en onderwerp. Als er ook een opdracht is meegegeven: kijk of het doel relevant is voor die opdracht.
Focus altijd op het grotere plaatje. Vermijd cijfers, maten of technische details.

Voorbeeldleerdoelen die altijd als 'specifiek genoeg' tellen

- *Aan het eind van de les kan ik netjes boorgaten maken. Dit is in mijn product te zien aan dat de gaten op de goede plaatsen zitten en de goede diameter hebben.*
- *Aan het eind van de les kan ik houtwerk netjes afwerken. Dit is in mijn product te zien aan dat scherpe kantjes zijn weggeschuurd en er geen lijmresten meer op zitten.*
- *Aan het eind van deze les kan ik onderdelen aan elkaar solderen. Dit is in mijn product te zien aan stevig aan elkaar vastzittende onderdelen en ziet het er schoon uit.*
- *Aan het eind van deze les kan ik hout op het juiste formaat afkorten. Dit is mijn product te zien aan dat de latten de juiste lengte hebben.*
- *Aan het einde van de les kan ik een sterke houten verbinding maken. Dit kun je aan mijn product zien aan dat er geen speling tussen de delen zit.*

Voorbeeldfeedback die je kunt geven in telling 1 t/m 3

- *Leerdoel 1: Aan het eind van deze les kan ik goed zagen. Dit is in mijn product te zien aan gezaagd hout. Feedback 1: Wat bedoel je met goed zagen? Wat en hoe is er gezaagd?*
- *Leerdoel 2: Aan het eind van deze les kan ik een mooie muziekstandaard bouwen. Dit is in mijn product te zien aan dat het mooi is. Feedback 2: Waaraan kun je zien dat het mooi is? Verbeterd leerdoel 2: Aan het eind van deze les kan ik een mooie muziekstandaard bouwen. Dit is in mijn product te zien aan dat alle lijmresten zijn verwijderd. Feedback 2 op verbeterd leerdoel: Dit is al beter, leg nog even uit wat je bedoelt met mooie muziekstandaard.*
- *Leerdoel 3: Aan het eind van deze les kan ik goed boren. Dit is in mijn product te zien aan dat het goed in elkaar zit. Feedback 3: Wat bedoel je met goed boren en goed in elkaar zitten?*
- *Leerdoel 4: Aan het eind van deze les kan ik sneller werken. Dit is in mijn product te zien aan dat het sneller af is. Feedback 4: Hoe zie je dat het sneller af is? Wat is er dan precies af binnen welke tijd?*
- *Leerdoel 5: Aan het eind van deze les kan ik een stevige stoel maken. Dat zie je aan dat ik erop kan staan. Feedback 5: Wat bedoel je met stevig?*

Voorbeeldfeedback die je kunt geven bij telling 4

- *Leerdoel 1: Aan het eind van deze les kan ik goed zagen. Dit is in mijn product te zien aan gezaagd hout. Feedback 1: "Goed zagen" kan verschillende dingen betekenen. Kun je uitleggen wat voor soort zagen je bedoelt? Bijvoorbeeld, kun je recht zagen of gebruik je een bepaalde zaag?*

- *Leerdoel 2: Aan het eind van deze les kan ik een mooie muziekstandaard bouwen. Dit is in mijn product te zien aan dat het mooi is. Feedback 2: "Mooie muziekstandaard" is een beetje vaag. Wat maakt het mooi? Denk aan dingen zoals de vorm, kleur of hoe netjes het eruitziet. Verbeterd leerdoel: Aan het eind van deze les kan ik een mooie muziekstandaard bouwen. Dit is in mijn product te zien aan dat alle lijmresten zijn verwijderd. Feedback verbeterd leerdoel 2: Het is goed dat je de lijmresten wil verwijderen. Kun je ook vertellen wat er nog meer mooi is aan je muziekstandaard? Bijvoorbeeld hoe de onderdelen passen of de kleuren.*
- *Leerdoel 3: Aan het eind van deze les kan ik goed boren. Dit is in mijn product te zien aan dat het goed in elkaar zit. Feedback 3: "Goed boren" is niet heel duidelijk. Kun je vertellen wat je bedoelt? Denk aan dingen zoals het kiezen van de juiste plek om te boren, of zorgen dat het netjes gebeurt.*
- *Leerdoel 4: Aan het eind van deze les kan ik sneller werken. Dit is in mijn product te zien aan dat het sneller af is. Feedback 4: "Sneller werken" is een beetje ingewikkeld. Hoe weet je dat je sneller bent? Heb je een bepaalde tijd in gedachten of een aantal dingen dat je afmaakt?*
- *Leerdoel 5: Aan het eind van deze les kan ik een stevige stoel maken. Dat zie je aan dat ik erop kan staan. Feedback 5: "Stevige stoel" klinkt goed, maar wat betekent stevig? Kun je uitleggen wat je doet om de stoel stevig te maken? Bijvoorbeeld het materiaal dat je kiest of hoe alles aan elkaar zit.*

Regels voor feedback (verplicht om te volgen!)
Maximaal 2 korte zinnen.
Altijd in B1-Nederlands.
Spreek de leerling altijd aan in de tweede persoon.
Benoem in je eerste reactie altijd iets positiefs.
Als een leerling reageert zonder het leerdoel aan te passen, zeg: “Goed dat je reageert, maar voeg dit ook toe aan je leerdoel.”
Als een punt na 2 keer niet wordt verbeterd, laat het dan los.
Bespreek haalbaarheid alleen als het echt onmogelijk is.

Telling
Ieder bericht begint met <<{getal}>>. Waar ‘getal’ staat zal 1, 2, 3, 4 of 5 staan, afhankelijk van de hoeveelste keer de leerling zijn antwoord stuurt.

Afsluiten
De sessie wordt afgesloten in 2 verschillende situaties. Deze situaties zijn:

- *Situatie 1: Wanneer een leerling zijn leerdoel voldoende heeft geformuleerd. In dit geval geef je aan dat het een goed leerdoel is om deze les mee aan de slag te gaan en wens je ze succes met de les.*
- *Situatie 2: Wanneer de leerling met telling 5 nog steeds geen leerdoel heeft dat voldoende is geformuleerd bedank je de leerling verklap je het leerdoel aan de leerling. Het leerdoel is volgens de zin: "Aan het eind van de les kan ik … Dit is in mijn product te zien aan …"*

In beide situaties eindig je het bericht met: <<EINDE-LEERDOEL>>

Stijl
Positief, luchtig, motiverend.
Kort en praktisch, zodat de leerling snel aan de slag kan.
Altijd gericht op het verbeteren en aanvullen van het leerdoel.

Verboden
Geen voorbeelden geven.
Geen extra opdrachten of checks toevoegen.

Bijlage 2. Vignetten leraren

Introductie

In opdracht van OCW doen de Universiteit Twente, CitoLab, Oberon en Kohnstamm Instituut onderzoek naar de manier waarop praktijkgerichte vaardigheden, zoals plannen, reflecteren en samenwerken, getoetst kunnen worden. Om deze vaardigheden te meten, heeft CitoLab een prototype van een toetsmodel ontwikkeld. Om inzicht te krijgen in de praktische haalbaarheid van dit ontwerp, gaan we met leraren en schoolleiders van verschillende scholen in gesprek. Hierbij leggen we het toetsmodel uit en horen we graag wat jullie hiervan vinden om het prototype verder door te ontwikkelen. Dit doen we aan de hand van een verschillende discussievragen.

Geef het gerust aan als je tussendoor vragen of opmerkingen hebt. We gaan zo samen het prototype doorlopen. Voel je vrij om op de vragen en op elkaar te reageren; we zullen niet steeds iedereen één voor één langsgaan, omdat daar te weinig tijd voor is. Als je een andere mening hebt dan een andere deelnemer, of juist wilt aangeven dat je het ermee eens bent, laat dit dan gerust weten.

In het gesprek lopen we het toetsmodel door: https://praktijkvaardigheden.web.app/lessons

Voorstelronde

In deze focusgroep zijn we aanwezig met verschillende leraren van po/vmbo/havo/vwo-scholen. Zou je jezelf kort willen voorstellen (functie, aantal jaren onderwijservaring)? En daarna in een paar zinnen willen uitleggen of, en op welke manier, praktijkgericht onderwijs op jouw school wordt vormgegeven.

Situatie

Jouw school gaat aan de slag met praktijkgericht onderwijs. Met verschillende opdrachten wil de school zicht krijgen op de vaardigheid van leerlingen om nieuwe dingen te leren. Deze opdrachten moeten inzicht bieden in de mate waarin leerlingen:

1. Doelen kunnen stellen
2. Kunnen evalueren of doelen behaald zijn
3. In staat zijn om vervolgstappen te bepalen

In het prototype oefenen leerlingen deze drie vaardigheden aan de hand van specifieke opdrachten. Via dit toetsmodel krijgen leerlingen en leraren bovendien inzicht in de mate waarin deze vaardigheden zijn ontwikkeld.

Voorbeeldopdracht

In het prototype dat we vandaag bespreken, is gekozen voor de voorbeeldopdracht houtbewerken, waarin leerlingen een muziekstandaard maken. Met deze opdracht wil je niet alleen toetsen of leerlingen in staat zijn om een goed product af te leveren, maar wil je ook toetsen of leerlingen in staat zijn om doelen te stellen, te evalueren of deze behaald zijn en op basis hiervan vervolgstappen kunnen bepalen. Het toetsmodel is op dit moment enkel gericht op deze drie vaardigheden op het gebied van ‘leren leren’. Het product zelf wordt dus niet beoordeeld in het toetsmodel.

Het idee is dat het toetsmodel ook voor andere praktijkgerichte opdrachten en op verschillende niveaus gebruikt kan worden. Bijvoorbeeld voor een kookles of een opdracht over het verduurzamen van het schoolplein.

Doelen stellen

Algemeen

- Hoe verloopt het proces van doelen stellen momenteel bij leraren en leerlingen? In hoeverre wordt hier al iets mee gedaan, en op welke manier?
    - Wat gaat goed en waar loop je tegenaan wanneer je leerlingen doelen laat opstellen?
- Hoe evalueren leerlingen momenteel of ze leerdoelen hebben behaald?
    - In hoeverre wordt hier al iets mee gedaan, en op welke manier?
- Hoe verloopt het proces van vervolgstappen bepalen (door leerlingen) momenteel bij leerlingen en door de leraar?
    - In hoeverre wordt hier al iets mee gedaan, en op welke manier?

We lopen nu de drie verschillende vaardigheden die zijn opgenomen in het prototype (doelen stellen, evalueren en vervolgstappen bepalen) stap voor stap door. Per vaardigheid willen we graag bespreken wat jullie hiervan vinden. We beginnen met het opstellen van de doelen.

Aan de hand van het leerdoel van de les, bedenkt de leerling een of meerdere eigen leerdoelen die hierbij aansluiten. Hiermee oefent de leerling met 'doelen kunnen stellen'. De leerling kiest een onderwerp waarover diegene een doel wil opstellen passend bij de doelen van de les, bijvoorbeeld 'afkorten en zagen'. Vervolgens bedenkt diegene wat hij/zij aan het einde van de les kan en hoe dit terug te zien is in het product. De leerling vult hier bijvoorbeeld in:

- Aan het einde van de les kan ik: goed alle kanten zagen (doel)
- Dit is in mijn product te zien aan: even lange kanten (succescriterium)

Het toetsmodel bepaalt met behulp van AI of het doel al goed is. Is het doel bijvoorbeeld duidelijk? Hier zit een taalmodel achter dat voorbeelden van goed en minder goed geformuleerde doelen heeft gekregen via een zogenoemde systeemprompt. Met het taalmodel wordt geëvalueerd of een doel specifiek genoeg is en of het haalbaar is binnen de opdracht. Zo niet, dan krijgt de leerling beknopte feedback. Wanneer een leerling vier pogingen heeft gedaan, maar het doel nog niet voldoet dan krijgt de leerling gerichtere feedback en kan hij dit leerdoel overnemen.

De leerling krijgt feedback op het doel. In dit geval moet het doel nog wat specifieker: wat bedoelt de leerling met 'goed' en wat is 'even lang'?

De leerling past het doel aan:

- Aan het einde van de les kan ik: alle poten van de muziekstandaard recht zagen met een zaag
- Dit is in mijn product te zien aan: de muziekstandaard wiebelt niet

AI beoordeelt of het aangepaste doel voldoet aan de eisen (is het doel specifiek en haalbaar binnen de opdracht?). Als het doel goed genoeg is, mag de leerling door.

De leerling krijgt een scherm te zien met de nieuw opgestelde doelen en doelen waar de leerling eerder aan heeft gewerkt (zowel behaalde als niet- behaalde doelen). De leerling selecteert de doelen waar diegene deze les aan gaat werken. In dit geval neemt de leerling twee doelen mee, namelijk het nieuwe doel en een eerder niet behaald doel.

De leerling heeft nu vastgesteld aan welke doelen hij gaat werken en kan verder met het uitvoeren van de opdracht. Voordat we verder gaan, hebben we een aantal vragen aan jullie.

Algemeen

- Denk je dat het haalbaar is om jouw leerlingen op deze manier doelen te laten opstellen?
    - Welke (aanvullende) handvatten heb je eventueel nog nodig?
- In hoeverre denk je dat het toetsmodel leerlingen goed ondersteunt bij het opstellen van leerdoelen?

- Welke mogelijke knelpunten en risico's zie je wanneer leerlingen leerdoelen opstellen via het toetsmodel?

Inzet van AI voor feedback en beoordeling

- Hoe kijk je aan tegen het gebruik van AI om leerlingen feedback te geven op doelen die zij hebben opgesteld?
    - Welke voordelen zie je in de inzet van AI, bijvoorbeeld op het gebied van tijdbesparing, consistentie of directe feedback?
    - Welke risico's of knelpunten verwacht je bij het gebruik van AI voor feedback en beoordeling, bijvoorbeeld voor de rol van de leraar of technische haalbaarheid?
- Is er beleid op het gebied van AI binnen de school? En is deze werkwijze in lijn met dit beleid?

Evalueren of doelen behaald zijn

De leerling gaat aan de slag met de opdracht. Na het afronden van de opdracht evalueert de leerling of hij de doelen heeft behaald. Per doel heeft de leerling drie opties:

- Ik heb het doel niet behaald
- Ik heb het doel gedeeltelijk behaald
- Ik heb het doel behaald

Ook de leraar beoordeelt of de doelen zijn behaald. De leraar krijgt een overzicht van alle leerlingen en de doelen die zij hebben opgesteld. Per leerling beoordeelt de leraar of de opgestelde doelen wel, niet of gedeeltelijk zijn behaald.

Doordat leerlingen hun eigen doelen beoordelen én jij als leraar ook een beoordeling geeft, wordt zichtbaar in hoeverre jullie oordelen overeenkomen. Dit geeft inzicht in de mate van 'zelfinschatting'. Hoe groter de overeenstemming tussen jouw oordeel en dat van de leerling, hoe hoger de score op 'zelfinschatting'.

In het voorbeeld hebben zowel de leraar als de leerling aangegeven dat het doel is behaald. Dit duidt op een hoge mate van goede zelfinschatting, omdat de beide oordelen overeenkomen. Er is ook sprake van een hoge mate van zelfinschatting wanneer de leerling en leraar beide aangeven dat het doel niet of gedeeltelijk is behaald.

Voordat we verder gaan, zijn we benieuwd wat jullie vinden van deze manier van evalueren.

Algemeen

- Denk je dat het haalbaar is om jouw leerlingen op deze manier te laten evalueren?
- Welke kansen of mogelijkheden voor het evalueren van doelen biedt deze werkwijze volgens het toetsmodel?
- Welke mogelijke knelpunten en risico's zie je wanneer leerlingen via het toetsmodel hun leren op deze manier evalueren?

Vervolgstappen bepalen

Op basis van de eigen beoordeling en de beoordeling van de leraar, bepaalt de leerling wat hij de volgende keer anders gaat doen. Dit vult hij bij elk leerdoel in.

Het toetsmodel is nog niet helemaal af. In het toetsmodel is daarom nog niet uitgewerkt hoe het bepalen van vervolgstappen precies wordt beoordeeld of teruggekoppeld aan de docent.

Overzicht zelfinschatting en beoordeling

Tot slot krijgt de leerling een overzicht van zijn eigen zelfinschatting: heeft hij zichzelf tot nu toe vaker overschat of onderschat? De leraar ontvangt een uitgebreider overzicht. Per leerling is zichtbaar:

- hoe vaak de leerling AI heeft ingezet bij het formuleren van doelen;
- of de geformuleerde doelen beoordeelbaar waren volgens jou als leraar;
- of de leerling zichzelf, vergeleken met het oordeel van de leraar, heeft overschat, correct ingeschat of onderschat;
- een overzicht van de opgestelde leerdoelen en de evaluatie van de leerling en docent

We zijn benieuwd wat jij vindt van de manier waarop leerlingen vervolgstappen bepalen en het overzicht van zelfinschatting en de leerdoelen.

Algemeen

- In het huidige toetsmodel is niet uitgewerkt hoe de leraar een terugkoppeling krijgt van de vervolgstappen. Wat zou je met de informatie doen over de bedachte vervolgstappen van de leerlingen? En hoe zou je deze informatie terug willen zien in het toetsmodel?

Inzicht in de beoordeling

Welke resultaten zouden jullie als leraar graag van de verschillende leerlingen willen zien? Verschilt dit van het huidige overzicht in het toetsmodel?

Overkoepelende vragen

We hebben nog een aantal overkoepelende vragen over het toetsmodel en deze manier van beoordelen.

Validiteit

- In hoeverre geeft deze manier van beoordelen een goed beeld van de mate waarin leerlingen in staat zijn om doelen te stellen, te evalueren of doelen behaald zijn en vervolgstappen te bepalen?

(Praktische) haalbaarheid

- Hoe zouden jullie deze manier van beoordeling gedurende de les inzetten?
- Denk je dat deze aanpak toepasbaar is bij verschillende soorten opdrachten, dus niet alleen voor houtbewerken? Waar leent deze vorm van beoordeling zich volgens jou wel of niet voor?

Randvoorwaarden en impact voor de leraar

- Voldoet de school aan de randvoorwaarden (zoals apparatuur of ICT-infrastructuur) om het toetsmodel op een praktische manier in te zetten bij toetsing en begeleiding?
- Wat vraagt deze manier van beoordelen voor de tijdsinvestering van leraren? Vraagt dit om nieuwe kennis/vaardigheden van leraren?
- Wat zou dit betekenen voor je werkdruk?

Monitoring

- In hoeverre is het haalbaar om tijdens of na de les voldoende tijd vrij te maken voor het systematisch monitoren van praktische vaardigheden, en om op basis daarvan gerichte acties te ondernemen (bijvoorbeeld wanneer blijkt dat leerlingen zichzelf structureel over- of onderschatten)?
- Sluit deze manier van monitoring voldoende aan op de gangbare praktijk binnen de school? Biedt het voldoende zicht op de voortgang?

Beslissing

- Zou je op basis van deze manier van toetsing van vaardigheden op het gebied van ‘leren leren’ een betere beslissing kunnen nemen over de schoolloopbaan van leerlingen (bijv. doorstroom of overstap naar een andere leerroute)?

Tot slot

- Hebben jullie nog vragen of opmerkingen die nog niet aan bod zijn gekomen?

Bijlage 3. Vignetten Schoolleiders

Introductie

In opdracht van OCW doen de Universiteit Twente, CitoLab, Oberon en Kohnstamm Instituut onderzoek naar de manier waarop praktijkgerichte vaardigheden, zoals plannen, reflecteren en samenwerken, getoetst kunnen worden. Om deze vaardigheden te meten, heeft CitoLab een prototype van een toetsmodel ontwikkeld. Om inzicht te krijgen in de praktische haalbaarheid van dit ontwerp, gaan we met leraren en schoolleiders van verschillende scholen in gesprek. Hierbij leggen we het toetsmodel uit en horen we graag wat jullie hiervan vinden. Dit doen we aan de hand van verschillende discussievragen.

Geef het gerust aan als je tussendoor vragen of opmerkingen hebt. We gaan zo samen het prototype doorlopen. Voel je vrij om op de vragen en op anderen te reageren; we zullen niet steeds iedereen één voor één langsgaan, omdat daar te weinig tijd voor is. Als je een andere mening hebt dan een andere deelnemer, of juist wilt aangeven dat je het ermee eens bent, laat dit dan gerust weten.

In het gesprek staat de praktische haalbaarheid van het meten van praktijkgerichte vaardigheden centraal (zijn er praktische bezwaren om praktische vaardigheden te toetsen?). In het gesprek lopen we daarnaast kort het toetsmodel door: https://praktijkvaardigheden.web.app/lessons

Voorstelronde

In deze focusgroep zijn we aanwezig met verschillende schoolleiders van po/vmbo/havo/vwo-scholen. Zou je jezelf kort willen voorstellen (functie, aantal jaren onderwijservaring)? En daarna in een paar zinnen willen uitleggen of, en op welke manier, praktijkgericht onderwijs op jouw school wordt vormgegeven.

Praktische haalbaarheid

- Op welke manier(en) worden praktijkgerichte vaardigheden op jullie school momenteel gemeten?
- Wat gaat er al goed bij het meten van praktijkgerichte vaardigheden bij jullie op school?
- Welke praktische uitdagingen ervaar je momenteel bij het meten van praktijkgerichte vaardigheden?
    - Op welke manier beïnvloeden groepsgrootte en personele capaciteit de haalbaarheid van praktijkgericht toetsen?
    - Zijn er momenteel voldoende faciliteiten (zoals materialen en apparatuur) om deze vaardigheden te meten?
- Wat hebben jullie nodig om het toetsen van praktische vaardigheden op jullie school makkelijker te maken?

Situatie

Leraren van jouw school gaan aan de slag met praktijkgericht onderwijs. Met verschillende opdrachten wil de school zicht krijgen op de vaardigheid van leerlingen om nieuwe dingen te leren. Deze opdrachten moeten inzicht bieden in de mate waarin leerlingen:

1. Doelen kunnen stellen
2. Kunnen evalueren of doelen behaald zijn
3. In staat zijn om vervolgstappen te bepalen

In het toetsmodel oefenen leerlingen deze drie vaardigheden aan de hand van specifieke opdrachten. Bovendien krijgen leerlingen en leraren inzicht in de mate waarin deze vaardigheden zijn ontwikkeld via dit toetsmodel.

Voorbeeldopdracht

In het toetsmodel dat we vandaag bespreken is gekozen voor de voorbeeldopdracht houtbewerken, waarin leerlingen een muziekstandaard maken. Met deze opdracht wil je niet alleen toetsen of leerlingen in staat zijn om een goed product af te leveren, maar wil je ook toetsen of leerlingen in staat zijn om doelen te stellen, te

evalueren of deze behaald zijn en op basis hiervan vervolgstappen kunnen bepalen. Het toetsmodel is op dit moment enkel gericht op deze drie vaardigheden op het gebied van ‘leren leren’. Het product zelf wordt dus niet beoordeeld in het toetsmodel.

Het idee is dat het toetsmodel ook voor andere praktijkgerichte opdrachten en op verschillende niveaus gebruikt kan worden. Bijvoorbeeld voor een kookles of een opdracht over het verduurzamen van het schoolplein.

Doelen stellen

We lopen nu de drie verschillende vaardigheden die zijn opgenomen in het toetsmodel (doelen stellen, evalueren en vervolgstappen bepalen) stap voor stap door. Per vaardigheid willen we graag bespreken wat jullie hiervan vinden. We beginnen met het opstellen van de doelen.

Aan de hand van het leerdoel van de les, bedenkt de leerling eigen een of meerdere eigen leerdoelen die hierbij aansluiten. Hiermee oefent de leerling met ‘doelen kunnen stellen’. De leerling kiest een onderwerp waarover diegene een doel wil opstellen passend bij de doelen van de les, bijvoorbeeld ‘afkorten en zagen’. Vervolgens bedenkt diegene wat hij/zij aan het einde van de les kan en hoe dit terug te zien is in het product. De leerling vult hier bijvoorbeeld in:

- Aan het einde van de les kan ik: goed alle kanten zagen (doel)
- Dit is in mijn product te zien aan: even lange kanten (succescriteria)

Het toetsmodel bepaalt met behulp van AI of het doel al goed is. Is het doel bijvoorbeeld duidelijk? Hier zit een taalmodel achter dat voorbeelden van goed en minder goed geformuleerde doelen heeft gekregen via een zogenoemde systeemprompt. Met het taalmodel wordt geëvalueerd of een doel specifiek genoeg is en of het haalbaar is binnen de opdracht. Zo niet, dan krijgt de leerling beknopte feedback. Wanneer een leerling vier pogingen heeft gedaan, maar het doel nog niet voldoet dan krijgt de leerling gerichtere feedback en kan hij dit leerdoel overnemen.

De leerling krijgt feedback op het doel. In dit geval moet het doel nog wat specifieker: wat bedoelt de leerling met ‘goed’ en wat is ‘even lang’?

De leerling past het doel aan:

- Aan het einde van de les kan ik: alle poten van de muziekstandaard recht zagen met een zaag
- Dit is in mijn product te zien aan: de muziekstandaard wiebelt niet

AI beoordeelt of het aangepaste doel voldoet aan de eisen (is het doel specifiek en haalbaar binnen de opdracht?). Als het doel goed genoeg is, mag de leerling door.

De leerling krijgt een scherm te zien met de nieuw opgestelde doelen en doelen waar de leerling eerder aan heeft gewerkt (zowel behaalde als niet- behaalde doelen). De leerling selecteert de doelen waar diegene deze les aan gaat werken. In dit geval neemt de leerling twee doelen mee, namelijk het nieuwe doel en een eerder niet behaald doel.
De leerling heeft nu vastgesteld aan welke doelen hij gaat werken en kan verder met het uitvoeren van de opdracht. Voordat we verder gaan, hebben we een aantal vragen aan jullie.

Algemeen

- Denk je dat het haalbaar is om leerlingen op deze manier doelen te laten opstellen?
- Welke mogelijke knelpunten en risico’s zie je wanneer leerlingen leerdoelen opstellen via het toetsmodel?

Inzet van AI voor feedback en beoordeling

- Hoe kijk je aan tegen het gebruik van AI om leerlingen feedback te geven op doelen die zij hebben opgesteld?
    - Welke voordelen zie je in de inzet van AI, bijvoorbeeld op het gebied van tijdbesparing, consistentie of directe feedback?
    - Welke risico's of knelpunten verwacht je bij het gebruik van AI voor feedback en beoordeling, bijvoorbeeld voor de rol van de leraar of technische haalbaarheid?
- Is er beleid op het gebied van AI binnen de school? En is deze werkwijze in lijn met dit beleid?

Evalueren of doelen behaald zijn

De leerling gaat aan de slag met de opdracht. Na het afronden van de opdracht evalueert de leerling of hij de doelen heeft behaald. Per doel heeft de leerling drie opties:

- Ik heb het doel niet behaald
- Ik heb het doel gedeeltelijk behaald
- Ik heb het doel behaald

Ook de leraar beoordeelt of de doelen zijn behaald. De leraar krijgt een overzicht van alle leerlingen en de doelen die zij hebben opgesteld. Per leerling beoordeelt de leraar of de opgestelde doelen wel, niet of gedeeltelijk zijn behaald.

Doordat leerlingen hun eigen doelen beoordelen én jij als leraar ook een beoordeling geeft, wordt zichtbaar in hoeverre jullie oordelen overeenkomen. Dit geeft inzicht in de mate van 'zelfinschatting'. Hoe groter de overeenstemming tussen jouw oordeel en dat van de leerling, hoe hoger de score op 'zelfinschatting'.

In het voorbeeld hebben zowel de leraar als de leerling aangegeven dat het doel is behaald. Dit duidt op een hoge mate van goede zelfinschatting, omdat de beide oordelen overeenkomen. Er is ook sprake van een hoge mate van zelfinschatting wanneer de leerling en leraar beide aangeven dat het doel niet of gedeeltelijk is behaald.

Voordat we verder gaan, zijn we benieuwd wat jullie vinden van deze manier van evalueren.

Algemeen

- Welke kansen of mogelijkheden voor het evalueren van doelen biedt deze werkwijze volgens het toetsmodel?
- Welke mogelijke knelpunten en risico's zie je wanneer leerlingen via het toetsmodel hun leren op deze manier evalueren?

Vervolgstappen bepalen

Op basis van de eigen beoordeling en de beoordeling van de leraar, bepaalt de leerling wat hij de volgende keer anders gaat doen. Dit vult hij bij elk leerdoel in.

Het toetsmodel is nog niet helemaal af. In het toetsmodel is daarom nog niet uitgewerkt hoe het bepalen van vervolgstappen precies wordt beoordeeld of teruggekoppeld aan de docent.

Overzicht zelfinschatting en beoordeling

Tot slot krijgt de leerling een overzicht van zijn eigen zelfinschatting: heeft hij zichzelf tot nu toe vaker overschat of onderschat? De leraar ontvangt een uitgebreider overzicht. Per leerling is zichtbaar:

- hoe vaak de leerling AI heeft ingezet bij het formuleren van doelen;

- of de geformuleerde doelen beoordeelbaar waren volgens jou als leraar;
- of de leerling zichzelf, vergeleken met het oordeel van de leraar, heeft overschat, correct ingeschat of onderschat;
- een overzicht van de opgestelde leerdoelen en de evaluatie van de leerling en docent

We zijn benieuwd wat jij vindt van de manier waarop leerlingen vervolgstappen bepalen en het overzicht van zelfinschatting en de leerdoelen.

Inzicht in de beoordeling

Welke resultaten zouden jullie als school graag van de verschillende leerlingen willen zien? Verschilt dit van het huidige overzicht in het toetsmodel?

Overkoepelende vragen

We hebben nog een aantal overkoepelende vragen over het toetsmodel en deze manier van beoordelen.

Validiteit

- In hoeverre geeft deze manier van beoordelen een goed beeld van de mate waarin leerlingen in staat zijn om doelen te stellen, te evalueren of doelen behaald zijn en vervolgstappen te bepalen?

(Praktische) haalbaarheid

- Denk je dat deze aanpak toepasbaar is bij verschillende soorten opdrachten, dus niet alleen voor houtbewerken? Waar leent deze vorm van beoordeling zich volgens jou wel of niet voor?

Randvoorwaarden en impact voor de leraar

- Voldoet de school aan de randvoorwaarden (zoals apparatuur of ICT-infrastructuur) om het toetsmodel op een praktische manier in te zetten bij toetsing en begeleiding?
- Wat vraagt deze manier van beoordelen voor de tijdsinvestering van leraren? Vraagt dit om nieuwe kennis/vaardigheden van leraren?
- Wat zou dit betekenen voor de werkdruk?

Monitoring

-
- Sluit deze manier van monitoring voldoende aan op de gangbare praktijk binnen de school?

Beslissing

- Zouden leraren op jouw school op basis van deze manier van toetsing van vaardigheden op het gebied van ‘leren leren’ een betere beslissing kunnen nemen over de schoolloopbaan van leerlingen (bijv. doorstroom of overstap naar een andere leerroute)?

Tot slot

- Hebben jullie nog vragen of opmerkingen die nog niet aan bod zijn gekomen?