



Aanpakken voor het ontwerpen van AI-gedreven toetstaken.

Handreiking 3

Aanpakken voor het ontwerpen van AI-gedreven toetstaken.

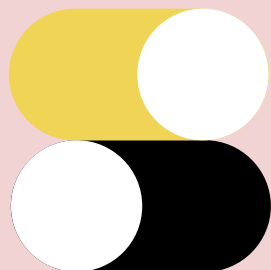
Handreiking 3

Door: K. Beekman (Fontys), S. Draaijer (VU), J. Beckers (Fontys), E. Schagen (VU), I. Hofman (MBO)
(met bijdragen van leden van het schrijfteam, instellingen en andere betrokkenen)

Met dank aan: Frank Verhagen, Bram Enning, Annette Peet, Amber de Wilde, en alle experts, stakeholders en betrokkenen die waardevolle input hebben geleverd.



Op deze uitgave is de Creative Commons Naamsvermelding 4.0-licentie van toepassing. Maak bij gebruik van dit werk vermelding van de volgende referentie: Beekman, K., Draaijer, S., Beckers, J., Schagen, E., & Hofman, I. (2025). *Visie op toetsing en examinering in het tijdperk van AI*. Utrecht. Npuls.



Inhoudsopgave

1. Inleiding	5
2. Ontwerpen van toetsing	9
3. <i>Two-Lane approach</i>	13
4. Naar AI-bewuste toetstaken: Stappenplan	21
5. Conclusie	29
Literatuur	30
Bijlage 1 Sectorspecifieke aanbevelingen en voorbeelden	33

1. Inleiding

De opkomst van generatieve artificiële intelligentie (hierna AI), denk aan tools als ChatGPT, Gemini of DALL-E, schudt het onderwijsveld flink op. Deze revolutionaire ontwikkeling dwingt onderwijsinstellingen om hun toetspraktijken fundamenteel te heroverwegen en *AI-bewuste toetspraktijken* te ontwikkelen, zoals beschreven in de landelijke visie op toetsing en examinering waarop deze handreiking gebaseerd is.

Waar voorheen een schriftelijk eindproduct (essay, ontwerp, verslag) of een klassiek tentamen volstond om studenten te beoordelen, is nu vaak onduidelijk in hoeverre AI heeft bijgedragen aan dat resultaat, terwijl uit recente studies ook blijkt dat ongeveer 80% van de studenten AI-tools inzet tijdens hun studie (Ma, 2025). Niet alleen de authenticiteit, maar ook de validiteit, betrouwbaarheid en inclusiviteit van toetsing komen onder druk te staan naarmate AI breder ingezet wordt (Bearman et al., 2024; Bakker et al., 2024). Tegelijkertijd biedt AI ongekende mogelijkheden voor gepersonaliseerde feedback, adaptieve leeromgevingen en authentieke simulaties (Lim et al., 2025; Mollick & Mollick, 2023). Voor onderwijsinstellingen betekent dit een dubbele opdracht: enerzijds moeten zij studenten voorbereiden op een AI-gedreven arbeidsmarkt door hen kennis, vaardigheden en inzichten bij te brengen over AI-toepassingen, anderzijds moeten zij de kwaliteit en waarde van hun diploma's waarborgen (Bearman et al., 2024). Deze uitdagingen vragen om een systematische aanpak waarbij *constructive alignment* - het op elkaar afstemmen van leerdoelen, onderwijsactiviteiten en toetsing - het fundament vormt (Biggs & Tang, 2011).

Bovendien vraagt de samenleving om AI-geletterde afgestudeerden: professionals die AI op verantwoorde, kritische en ethische wijze kunnen inzetten in hun vakgebied. Onderwijsinstellingen moeten studenten hierop voorbereiden; naast vakkennis en complexe vaardigheden, moeten zij leren hoe ze AI kunnen benutten én de beperkingen en ethische implicaties ervan te begrijpen. AI-geletterdheid is inmiddels een cruciale competentie geworden, waarbij studenten niet alleen technische vaardigheden, maar ook kritisch denkvermogen en ethisch bewustzijn ontwikkelen (Promma et al., 2025; Tsai et al., 2023). Dit omvat het vermogen om AI-output kritisch te evalueren, bias te herkennen, ethische implicaties te overwegen en verantwoorde keuzes te maken over wanneer en hoe AI in te zetten (Ng et al., 2021; Sperling et al., 2024).

AI is en blijft een technologisch hulpmiddel dat ingezet kan worden om leerprocessen te optimaliseren, niet om deze over te nemen. Het ontwikkelen van een stevige basiskennis blijft onverminderd belangrijk, ook in AI-gestuurde leeromgevingen, om output kritisch te kunnen

beredeneren of verbanden te kunnen leggen (Fawns & Schuwirth, 2024). Voor toetsing en examinering betekent dit dat we niet alleen belang hechten aan objectieve, goed onderbouwde beslissingen, maar óók aan het begeleiden en evalueren van leerprocessen gericht op de leerdoelen.

Doel van deze handreiking

De ontwikkelingen rondom AI gaan snel en vragen om een fundamentele herbezinning op wat we verstaan onder waardevol onderwijs en hoe we dit op authentieke wijze kunnen toetsen. De kernboodschap: met een sterke, evidence-informed visie op AI-bewuste toetsing en examinering kunnen we het beste van twee werelden verenigen. We realiseren toetsing en examinering die eerlijk, valide en betrouwbaar is, en omarmen tegelijkertijd innovatie en toekomstgerichte vaardigheden.

Deze handreiking bouwt voort op de landelijke visie op toetsing en examinering en vertaalt de daarin beschreven *Two-Lane approach* naar concrete aanpakken voor het ontwerpen van AI-bewuste toetsen (toetstaken) in mbo, hbo en wo. We baseren ons op recente wetenschappelijke inzichten en praktijkvoorbeelden en streven naar aanpakken die sectorbreed gedeeld kunnen worden. Daarbij erkennen we dat de invulling per sector (mbo, hbo, wo) en per discipline zal verschillen, maar dat de onderliggende principes universeel toepasbaar zijn.

Dat betekent dat we niet simpelweg terugvallen op het verbieden van AI-gebruik, of toetstaken ‘AI-proof’ maken, maar wel ethisch-verantwoorde, betrouwbare toetstaken ontwerpen, waarin AI benut wordt om de beoogde leerdoelen te bereiken, zonder de kwaliteit of betrouwbaarheid van toetsing in diskrediet te brengen.

Kernprincipes

Deze handreiking hanteert dezelfde kernprincipes als de landelijke visie, vertaald naar curriculumontwerp:

- **Mensgericht en transparant:** AI-toepassingen kunnen onderwijs- en toetspraktijken ondersteunen, maar de menselijke regie blijft cruciaal. Dit betekent dat curriculumontwerpers expliciet moeten vastleggen waar, wanneer en hoe AI wordt ingezet, en welke rol dit speelt in het bereiken van leerdoelen.
- **Eerlijkheid en inclusie:** AI mag geen nieuwe barrières opwerpen of bestaande ongelijkheden vergroten. Bij curriculumontwerp moet rekening gehouden worden met verschillen in toegang tot AI-tools, taalbarrières en digitale vaardigheden. Alle studenten moeten gelijke kansen krijgen om AI-geletterdheid te ontwikkelen (Bakker et al., 2024).
- **Validiteit en betrouwbaarheid voorop:** Een curriculum moet zodanig ontworpen zijn dat toetsing valide en betrouwbaar blijft, ongeacht AI-gebruik. Dit vereist een doordachte mix van toetsvormen waarbij soms AI-gebruik wordt toegestaan of zelfs vereist (baan 2), en

soms juist wordt uitgesloten om individuele bekwaamheid vast te stellen (baan 1), conform de *Two-Lane approach*.

- **Constructive alignment als kompas:** Leerdoelen, onderwijsactiviteiten en toetsing moeten in samenhang worden (her)ontworpen, waarbij AI-integratie geen doel op zich is maar een middel om leerdoelen te bereiken (Biggs & Tang, 2011).
- **Flexibiliteit en wendbaarheid:** Curricula moeten kunnen meegroeien met technologische ontwikkelingen zonder de kern van het onderwijs uit het oog te verliezen. Dit vraagt om een flexibele structuur met ruimte voor aanpassingen (Liu et al., 2023; Lodge et al., 2023).

2. (Her)ontwerpen van toetsaken

Idealiter draagt toetsing bij aan twee doelen: enerzijds door procesgerichte toetsing het ondersteunen, stimuleren en versterken bij leren en opbrengsten van studenten. Anderzijds het kunnen vaststellen van kwaliteit en beheersingsniveau door valide, transparante en betrouwbare oordelen over de gerealiseerde prestatie van de student. AI voegt hier een nieuwe dimensie aan toe (Perkins et al., 2023). We moeten AI bewust en verantwoord integreren in toetsing en examinering, door én de benodigde ondersteuning bij leerprocessen in AI-gedreven omgevingen te bieden, én beslissingen over leeropbrengsten kritisch tegen het licht te houden. Uit onderzoek blijkt dat AI-integratie in toetsontwerp vooral effectief verloopt, wanneer dit gekoppeld wordt aan duidelijke leerfasen (Crocco; 2025; Lyanda et al., 2024). Dit kan door onderwijsaanbod te analyseren vanuit leerfasen, om zo gericht keuzes te maken vanuit toetsfuncties en aansluitend op de leerdoelen: acquisition (kennisverwerving), retention (consolidatie in vergelijkbare context), en transfer (toepassing in nieuwe contexten). De integratie van AI onderwijs en toetsen vraagt wellicht ook om een heroverweging van cognitieve niveaus. De herziene Bloom's Taxonomy (Oregon State University, 2024) biedt hiervoor de volgende concrete handvatten:

- Bij lagere cognitieve niveaus (onthouden, begrijpen) kan AI ondersteunen door feitelijke informatie te verstrekken en concepten uit te leggen;
- Bij middelste niveaus (toepassen, analyseren) kan AI als sparringpartner fungeren voor probleemoplossing;
- Bij hogere niveaus (evalueren, creëren) staan menselijke vaardigheden zoals ethische reflectie, creativiteit en oordeelsvorming centraal

Deze indeling helpt toetsontwerpers bij het maken van weloverwogen keuzes over waar AI-integratie waardevol is en waar menselijke bekwaamheid vereist is.

Constructive alignment

Het onderwijskundig principe 'constructive alignment' blijft hierbij onverminderd relevant, en vormt de basis voor (her)ontwerpen van onderwijs en -toetsing richting AI-bewuste onderwijspraktijken. In het mbo betekent dit vaak het afstemmen van kerntaken en werkprocessen op de realiteit van AI-gebruik in de beroepspraktijk, waarbij examencommissies volgens de WEB toezien op de kwaliteit van examinering. In het hbo vraagt dit om herontwerp van competentiegerichte curricula waarbij leeruitkomsten expliciet AI-geletterdheid incorporeren. In het wo vereist dit specifieke aandacht voor academische integriteit en de balans tussen AI-ondersteunde kenniscreatie en individuele onderzoeksvaardigheden. Deze uitgangspunten zorgen voor de realisatie van een samenhangend curriculum, waarin leerdoelen, onderwijsactiviteiten en toetsing consistent, zorgvuldig en gebalanceerd op elkaar zijn afgestemd, passend bij de onderwijs- en toetsvisie.

Leerdoelen bijstellen: De veranderende beroepspraktijk maakt het noodzakelijk dat opleidingen in veel gevallen hun leerdoelen niet alleen herzien maar ook expliciet uitbreiden, zodat afgestudeerden blijven aansluiten bij actuele professionele eisen. Door de opkomst van AI moeten studenten, naast grondige domeinkennis, ook AI-geletterdheid verwerven (Ng et al., 2021). Dat laatste geldt overigens ook voor docenten. AI-geletterdheid is meer dan instrumentele AI-inzet. Recent onderzoek toont dat effectieve AI-geletterdheid bestaat uit vier componenten: AI-bewustzijn, AI-gebruik, AI-evaluatie en AI-ethiek (Laupichler et al., 2023). Dat betekent onder andere leren om AI-uitvoer kritisch te analyseren, data doelgericht te interpreteren en AI-tools verantwoord in te zetten (Sperling et al., 2024).

In aangescherpte leerdoelen staat bijvoorbeeld dat een student ‘effectief en verantwoord AI-toepassingen weet te benutten voor informatievergaring en resultaatverbetering’, ‘kan aantonen hoe hij of zij AI-output op juistheid en mogelijke bias beoordeelt’ én ‘de ethische implicaties en risico’s van AI-gebruik in een specifieke context kent’. Een traditioneel doel als “De student kan een zakelijke brief in het Engels schrijven” ontwikkelt zich daardoor tot de verwachting dat de student met ondersteuning van AI een zakelijke brief opstelt, deze redigeert op correctheid en toon, en de kwaliteit van de AI-gegenereerde tekst zelfstandig kan evalueren. In het mbo vertaalt dit zich naar concrete werkprocessen zoals “De student past AI-tools toe conform brancherichtlijnen”, in het hbo naar competenties als “De student integreert AI-tools kritisch in het beroepsmatig handelen”, en in het wo naar academische vaardigheden zoals “De student evalueert methodologische implicaties van AI-gebruik in onderzoek” (Tan et al., 2025). Op die manier wordt voor studenten duidelijk wat in AI-gedreven omgevingen van hen wordt verwacht, terwijl docenten onderwijsactiviteiten en toetsing gericht kunnen afstemmen op die uitgebreide set competenties.

(Her)ontwerp van toetsstaken: Het constructief afstemmen van toetsing op de leerdoelen en onderwijsactiviteiten, vraagt om beargumenteerde keuzes ten aanzien van de benaderingen van toetsing en de gewenste toepassing. Uit onderzoek (Black & Wiliam, 1998; 2009) blijkt dat toetsing vanuit verschillende toetsfuncties bijdraagt aan het richting geven, bevorderen en realiseren van onderwijs- en toetskwaliteit, op basis van zorgvuldige afstemming met de beoogde leerdoelen en onderwijsactiviteiten:

- **Assessment for learning** (toetsing als middel om van te leren) richt zich op het ondersteunen van het leerproces door middel van formatieve toetsing: leerlingen en leraren krijgen feedback waarmee ze inzicht krijgen in voortgang en leerbehoeften, zodat het leerproces gericht kan worden bijgestuurd. Voorbeelden van AI-gedreven toetsing tijdens het leren in het mbo zijn bijvoorbeeld de AI-ondersteunde praktijksimulaties met real-time feedback (Ifenthaler et al., 2025), in het hbo de AI-verrijkte projectevaluaties waarbij studenten leren reflecteren op AI-gegenereerde suggesties (Katona & Gyonyoru, 2025), en in het wo de NLP-ondersteunde peer review processen die diepere academische discussies faciliteren (Wood, 2022).

- **Assessment as learning** (toetsing als startpunt van leren) benadrukt de actieve rol van de leerling, waarbij toetsing wordt ingezet als een middel voor zelfreflectie en zelfregulatie. Dit helpt leerlingen hun eigen leerproces beter te begrijpen en verantwoordelijkheid te nemen. Nieminen et al. (2025) benadrukken dat digitale self-assessment tools student agency kunnen versterken, mits ontworpen vanuit principes van universeel ontwerp en inclusiviteit.
- **Assessment of learning** (toetsing van leren) is beslissend van aard en bedoeld om vast te stellen wat leerlingen hebben geleerd, bijvoorbeeld voor rapportages, certificering, diplomering of het geven van cijfers. Denk hierbij aan toetsing in gecontroleerde omgevingen, waarbij direct toezicht essentieel is voor het waarborgen van academische integriteit

Deze onderwijskundige bewezen concepten over toetsing als middel om van te leren, als startpunt van leren, of toetsing om kwaliteit van leren vast te stellen, blijven, ongeacht de technologische innovaties en mogelijkheden, onverminderd relevant en bruikbaar. Het doel is namelijk niet om toetsstaken te herzien enkel vanuit het doel AI in te zetten, maar om *AI-bewuste toetspraktijken* te realiseren, waarbij we AI integreren op gecontroleerde wijze waar deze het leerproces verrijkt en tegelijk voorkomen dat AI het leren en de bewijsvoering daarvan dreigt te ondermijnen. Deze holistische benadering zorgt voor samenhang en doorgaande leerlijnen in plaats van losse onderdelen, en voor een consistente leerervaring en onderwijskwaliteit (Boud et al., 2023), ook wanneer AI integraal in het curriculum wordt opgenomen. AI raakt immers alle drie de componenten (leerdoelen, activiteiten, toetsing).

Toekomstgerichte onderwijsactiviteiten: AI opent een breed palet aan mogelijkheden om onderwijs- en leeractiviteiten opnieuw vorm te geven. AI-tutores kunnen geïndividualiseerde oefenvragen genereren, brainstormsessies verrijken of semi-realistische scenario’s simuleren; elders fungeert dezelfde technologie als sparringpartner, docent-assistent, projectbegeleider of feedbackleverancier, en ontstaan samenwerkingsopdrachten waarin studenten met AI-tools een product ontwikkelen (Mollick & Mollick, 2023; Micheni et al., 2024; Lim et al., 2025). Dit potentieel maakt onderwijs flexibeler, adaptiever en inclusiever, maar vraagt om een doordacht ontwerp waarin de rol van AI expliciet wordt geadresseerd en verantwoord gebruik centraal staat (Mollick & Mollick, 2023). Binnen die ontwerpcontext experimenteren docenten al met werkvormen die AI integreren in het leerproces. Studenten laten bijvoorbeeld ChatGPT een eerste versie van een rapport, plan van aanpak of practicumverslag opstellen en bespreken vervolgens in groepen de sterke en zwakke punten ervan, zodat zij zowel de inhoud als de AI-bijdrage kritisch leren beoordelen. In practica verifiëren en corrigeren studenten een door AI gegenereerde data-analyse, een oefening die nauwkeurigheid en analytisch denken bevordert terwijl AI als assistent fungeert. Docenten wisselen deze aanpak bewust af met opdrachten zonder AI om fundamentele kennis en vaardigheden in te slijpen, waarna vergelijkbare taken mét AI worden ingezet om die basis op een hoger niveau toe te passen. Zo ervaren studenten wanneer zelfstandig handelen vereist is en wanneer slim samenwerken met AI

tot betere resultaten leidt. Deze voorbeelden illustreren hoe innovatieve, AI-ondersteunde werkvormen binnen een zorgvuldig ontworpen didactisch kader niet alleen aansluiten bij hedendaagse beroepspraktijken, maar ook bijdragen aan kritische reflectie en verantwoordelijkheidsbesef bij zowel docenten als studenten.

3. *Two-Lane approach*

Het ontwerpen en inrichten van een weloverwogen toetsprogramma is complex, en de impact van AI als complicerende factor heeft de bestaande uitdagingen nog duidelijker zichtbaar gemaakt. Idealiter draagt toetsing bij aan twee doelen: enerzijds de ondersteuning en versterking van het leerproces van de student (ontwikkelingsgerichte toetsing) en anderzijds het nemen van valide, transparante en betrouwbare beslissingen over de leerresultaten van de student (besluitvormende toetsing) (Black & William, 2009; Heeneman et al., 2021)). Er moet zorgvuldig gebalanceerd worden tussen innovatie en borging van toetskwaliteit.

AI beïnvloedt de leerdoelen door veranderende beroepsprofielen en dwingt opleidingen hun huidige leerdoelen aan te passen of onder de loep te nemen. Bovendien moeten onderwijs- en toetsactiviteiten hierop aansluiten (Mollick & Mollick, 2023). Deze onvermijdelijke integratie van AI in onderwijs en toetsing voegt nieuwe dimensies toe aan toetsing en examineren (Bearman, et al., 2024; Perkins et al., 2024; Vargas-Murillo et al., 2023). De meest relevante vraag die iedere onderwijsaanbieder op dit moment dan ook bezighoudt, is ‘waartoe leiden we op nu AI het professionele en persoonlijke leven doordringt, wat betekent dit voor de benodigde kennis, vaardigheden en competenties, en hoe vertalen deze vraagstukken zich naar de toetspraktijk?

Toetsontwerpers hanteren geregeld de benadering waarbij beoogde leerdoelen het uitgangspunt vormen, om vervolgens te bepalen, op basis van (tussentijdse) informatie en/of bewijzen, hoe studenten ondersteund kunnen worden bij het bereiken van leerdoelen of hoe gerealiseerde opbrengsten aangetoond worden (Wiggins & McTighe, 2005). Deze holistische benadering bij het ontwerpen van toetsstaken zorgt voor samenhang in doorgaande leerlijnen, voor een consistente leerervaring en onderwijskwaliteit (Boud et al., 2023). De inhouden van de beoogde leerdoelen (in de vorm van leeruitkomsten en/of eindtermen) variëren tussen opleidingsaanbieders, zowel wat betreft kennis- en toepassingsniveau, discipline als beroepsprofiel, en zullen zo zorgvuldig mogelijk in overleg met het werkveld moeten worden afgestemd. De komst van AI leidt tot veranderende beroepen en noodzakelijke (nieuwe) kennis en vaardigheden, wat zijn weerslag heeft op bestaande leerdoelen. In veel gevallen zijn aanpassingen en uitbreiding van bestaande en nieuwe leerdoelen vereist. En ook onderwijs- en toetsing zullen hierop moeten worden afgestemd, zodat de studenten in staat zijn deze nieuw vormgegeven leerdoelen te behalen. Er zijn allerlei innovatieve mogelijkheden om AI-gedreven onderwijsactiviteiten vorm te geven, en datzelfde geldt voor toetsing (Mollick, & Mollick, 2023). Maar het is de bedoeling dat dit als middel wordt ingezet, ter ondersteuning en versterking van het aanbod, niet om leren en opbrengsten over te nemen. AI zal dus bewust en verantwoord in de gehele basis van het toetsontwerp moeten worden verweven.

Benadering voor AI-integratie in toetstaken

Een veelbelovende benadering om AI in bestaande toetstaken te integreren is de *Two-Lane approach*. Dit raamwerk is geïnspireerd door internationale aanbevelingen en richtlijnen (Liu et al., 2023; TEQSA, 2023; Lodge et al., 2023) en bewezen toetstheorieën (Black & William, 1998; 2009). Het biedt houvast bij het op passende en verantwoorde wijze integreren van AI, in bestaande en nieuwe toetstaken, door de voorgestelde tweebaans-aanpak op grond van twee fundamentele principes. Uitgaande van de twee uitgangsprincipes (AI integreren waar mogelijk, en diploma-kwaliteit borgen) stelt de *Two-Lane approach* twee “banen” voor om AI te integreren in het toetsprogramma, gebaseerd op de **functie** van de toetsing en dekking van de leerdoelen.

Baan 1: Kwaliteitsmeting van leren en opbrengsten (*Assessment of Learning*): Toetsing in baan 1 richt zich op het vaststellen van de kwaliteit van het leren en de leerresultaten van de student in relatie tot de beoogde leerdoelen op basis van valide en betrouwbare, data-geïnformeerde besluitvorming. Deze toetsing vindt plaats onder gecontroleerde omstandigheden en toezicht, en AI-gebruik is hierbij niet of in beperkte mate toegestaan. In het mbo zijn dit bijvoorbeeld praktijkexamens waarbij studenten onder observatie beroepshandelingen uitvoeren conform het kwalificatiedossier, zoals een verpleegkundige student die vitale functies controleert bij een gesimuleerde patiënt (Andreou et al., 2024). In het hbo worden gecontroleerde kennistoetsen en criteriumgerichte interviews ingezet waarbij studenten hun projectwerk mondeling verantwoorden zonder AI-ondersteuning (Schellekens et al., 2023). In het wo valt te denken aan viva voce examens waarbij promovendi hun onderzoek verdedigen, of gesloten boektentamens waarbij conceptueel begrip wordt getoetst zonder externe hulpmiddelen (Evangelista, 2025).

AI mag alleen worden ingezet als de leerdoelen dit vereisen én de individuele bijdrage van de student direct controleerbaar en herleidbaar is. Zo kunnen betrouwbare, valide beslissingen worden genomen over de geleverde prestatie en de individuele studentinspanning hierbij. En kan met voldoende zekerheid worden vastgesteld of en in welke mate de student bepaalde (cruciale) kennis en vaardigheden zelf beheerst, zonder (directe) inmenging van AI of met bewuste, direct-herleidbare AI-inzet op basis van de leerdoelen. Denk aan toetsvormen die zo ontworpen zijn dat AI geen invloed heeft op de te meten prestatie, zoals klassikale tentamens met surveillance, mondelinge toetsvormen, praktische vaardigheidstoetsen of observaties.

Als bij deze besluitvormende toetsing over de kwaliteit van leren en/of opbrengsten van een student AI wél is toegestaan, moet vooraf duidelijk zijn hoe en onder welke condities AI wordt ingezet. Ook moet vastliggen hoe wordt toegezien op validiteit, betrouwbaarheid en meetbaarheid van de resultaten. Voorbeelden van AI-inzet binnen baan 1 zijn: AI-gestuurde simulaties (waarbij een examinerator in een semi-authentieke context op basis van observaties

bepaalt of, en in hoeverre, de student de beoogde kennis en vaardigheden beheerst, bijvoorbeeld het uitvoeren van complexe medische toepassing, of complexe vaardigheden/competenties in gesimuleerde uitdagende contexten of uitzonderlijke omstandigheden (het tonen van klantvriendelijkheid in het hotelwezen bij een veeleisende gast). Een ander voorbeeld is adaptieve toetsing, waarbij AI de antwoorden van de student analyseert tijdens de toets en de moeilijkheidsgraad van volgende vragen dynamisch aanpast. Overigens blijft menselijke kwaliteitscontrole bij geautomatiseerde scoring in adaptieve toetsen cruciaal voor het waarborgen van fairness en validiteit (Jung et al., 2025).

Ongeacht of AI-gebruik in baan 1 is toegestaan (beperkt, onder toezicht en controleerbaar); de toetsing moet de integriteit garanderen en zekerstellen dat de beoordeelde prestatie daadwerkelijk door de student zelf geleverd is.

Baan 2: Ontwikkelingsgerichte toetsing (*Assessment for & as Learning*): Toetsing in baan 2 richt zich op het zichtbaar maken en optimaliseren van het leren van de student gedurende het leerproces, en op de opbrengsten richting de beoogde leerdoelen. Deze toetsing dient als startpunt van en/of (bij)sturing tijdens leren. Toetsing in baan 2 is gericht op data-geïnformeerde ondersteuning, bijsturing en versterking van leerprocessen in de richting en bij realisatie van de beoogde. In deze baan is AI-gebruik geoorloofd, gewenst, of zelfs verplicht (afhankelijk van de leerdoelen en binnen wettelijke kaders en mogelijkheden). De nadruk ligt op het leren van, door en in samenwerking met AI, met expliciete aandacht voor AI-vaardigheden en inzichten die nodig zijn voor kritisch, ethisch en correct gebruik. Studenten moeten bij deze aanpak aantonen wanneer en hoe ze AI hebben ingezet, en hoe dit heeft bijgedragen aan hun leerproces en resultaten. Studenten gebruiken bij AI-feedback vooral cognitieve strategieën (Zhan, & Yan, 2025). Bij toetstaken in baan 1 zullen metacognitieve reflectie en zelfregulatiestrategieën expliciet gestimuleerd moet worden. Zeker als de kwaliteit van deze AI-geïntegreerde ontwikkeling weegt mee in het eindoordeel.

Toetsing in baan 2 heeft een open karakter, aangezien de toetstaken niet noodzakelijkerwijs onder toezicht plaatsvinden of volledig controleerbaar zijn. Denk aan het uitwerken van schriftelijke opdrachten, ontwerp opdrachten, of het verzamelen van bewijslast voor portfolio's. Toetsing in deze baan bereidt studenten voor op situaties waarin AI-gebruik wenselijk of onvermijdelijk is, en stimuleert hen om op een juiste en verantwoorde manier met AI samen te werken. AI wordt ingezet om leerprocessen en resultaten te ondersteunen en te verbeteren. Bij deze toetsing wegen kritische reflectie en verantwoording van het leerproces zwaar mee. Zo moet bijvoorbeeld transparant worden gemaakt welke AI-hulpmiddelen zijn ingezet en hoe AI-gegenereerde output en feedback zijn verwerkt. Hiermee wordt het leren zichtbaar gemaakt en worden kritisch redeneren, probleemoplossend vermogen, creativiteit en AI-geletterdheid in authentieke contexten optimaal ontwikkeld.

Onderstaande tabel zet de kernverschillen op een rij:

Aanpak voor doelmatige, verantwoorde en passende AI-integratie in toetspraktijken		
	Baan 1	Baan 2
Functie van toetsing	Valide, betrouwbare kwaliteitsmeting om vast te stellen of/in welke mate de (individuele) student de beoogde leerdoelen bereikt heeft, op basis van de gerealiseerde opbrengst, met consequenties voor de voortgang.	Cyclische, ontwikkelingsgerichte toetsing waarbij tussentijds (of als startpunt) bewijs verzameld wordt over de studentontwikkeling richting de beoogde leerdoelen; docent, student en peers benutten deze informatie ter ondersteuning, bijsturing en versterking van eigen leerproces en eigenaarschap richting de beoogde leerdoelen.
Mate van toezicht en controle	Toetsing vindt plaats onder direct toezicht van een examiner. In principe individueel en controleerbaar. De geleverde studentprestatie is authentiek en van de student zelf.	Toetsing vindt plaats zonder direct toezicht van een examiner. Individueel en/of in samenwerking met anderen, en de studentprestatie is niet volledig controleerbaar tot een individu te herleiden.
Rol van AI	Afhankelijk van de beoogde leerdoelen kan AI-gebruik beperkt worden toegestaan of uitgesloten (bijvoorbeeld bewust uitsluiten bij taken waar de student eigen vaardigheid zonder AI moet tonen, beperkt toestaan wanneer correct AI-gebruik bij getoetst moet worden).	Afhankelijk van de beoogde leerdoelen wordt AI-gebruik juist toegestaan of zelfs verplicht (bijvoorbeeld om studenten verantwoord AI-gebruik aan te leren en te stimuleren waar dit toegevoegde waarde biedt).
Voorbeelden	Persoonlijke interactieve mondelinge beoordelingen; viva voces; gelijktijdige beoordelingen en vaardigheidsontwikkeling in de les; toetsen en examens. MBO: AI-gestuurde simulaties in de zorg waarbij een student onder observatie van een examiner complexe verzorgende handelingen uitvoert, waarbij de AI verschillende patiëntscenario's genereert maar de handelingen zelf door de student worden uitgevoerd. HBO: Adaptieve kennistoetsen waarbij AI het niveau van vragen aanpast maar alle antwoorden van de student zelf komen WO: Mondelinge verdedigingen waarbij studenten hun scriptie toelichten en verantwoorden. AI-detectie kan worden gebruikt voor screening screenen maar het eindoordeel ligt altijd bij de examiner.	AI om reflectie uit te lokken, kritische analyses te doen op output, structuur te suggereren, ideeën te brainstormen, literatuur samen te vatten, inhoud te maken, tegenargumenten te suggereren, duidelijkheid te verbeteren, formatieve feedback te geven, enz. MBO: Werkplekopdrachten waarbij studenten AI-tools gebruiken zoals in de beroepspraktijk, bijvoorbeeld een administratief medewerker die leert werken met AI-ondersteunde facturatiesystemen. HBO: Projectopdrachten waarbij teams AI inzetten voor onderzoek en ontwikkeling, met verplichte procesverslagen over AI-gebruik. WO: Onderzoeksprojecten waarbij AI wordt ingezet voor literatuuranalyse of data-analyse, met kritische reflectie op methodologische implicaties.

Samenhang & afbakening: De kracht van de *Two-Lane approach* zit in het bewust en beargumenteerd integreren van AI-toepassingen en het (aangeleerde) verantwoord en correct gebruik ervan door studenten gedurende hun opleiding. Het draagt bij aan het realiseren van een AI-bewuste toetspraktijk, afgestemd op onderwijsactiviteiten en leerdoelen. Studenten hoeven deze tweedeling niet expliciet te ervaren; het is de taak van het ontwerpteam om voor een goede samenhang en overgang te zorgen. Ieder toetsprogramma zou idealiter bij elke leerlijn, onderwijseenheid of cursus zowel toetsing volgens baan 1 als baan 2 moeten bevatten (Liu, et al., 2023), in een logische volgorde en afwisseling. De twee banen vullen elkaar aan en zorgen voor constructieve afstemming binnen het onderwijsaanbod (Curtis, 2025). Cruciale leerdoelen worden geborgd in het eerste spoor (waardoor de diplomawaarde gegarandeerd is), terwijl toetsing volgens het tweede spoor ervoor zorgt dat studenten worden voorbereid op een beroepspraktijk waarin AI-gebruik de realiteit is.

In feite bevatten bestaande toetsprogramma's ook nu al verschillende toetsfuncties, gebalanceerd en beargumenteerd, afgestemd op de onderwijsactiviteiten en leerdoelen. De *Two-Lane approach* ondersteunt de integratie van AI in dit reeds weloverwogen aanbod, en maakt tegelijk duidelijk waar de toetsing aanvullende aanpassing of herziening nodig heeft wanneer een huidige toetsvorm niet meer aansluit bij de beoogde functie van de toets.

Ontwerpprincipes: Dit betekent dat bij elke toets expliciet de volgende vragen gesteld moeten worden: Wat is het **doel** van deze toetsing? Kan of mag de student leerdoelen aantonen door opbrengsten die in samenwerking met AI tot stand zijn gekomen, of moet de student dit ook perse zonder AI kunnen? Gaat het om het kunnen nemen van beslissingen over studentprestaties op basis van de vastgestelde kwaliteit van opbrengsten (dan baan 1), of gaat het om het ondersteunen en stimuleren van de studentontwikkeling richting de beoogde leerdoelen (dan baan 2)? Om dit nog concreter te maken: is het leerdoel gericht op kennis-/vaardigheden verwerven, of consolidatie? Dan kan AI-ondersteuning zeker waardevol zijn (Baan 2). Is het leerdoel gericht op complexe toepassingen in nieuwe contexten? Baan 2-toetsing biedt ontwikkelmogelijkheden en bijsturing. Moet kunnen worden vastgesteld in hoeverre de individuele student op basis van aantoonbare inspanning het leerdoel beheerst bij complexe toepassingen of nieuwe contexten? Baan 1, aangezien daar controleerbaarheid en direct toezicht worden meegenomen (Crocco, 2025).

Een richtlijn hierbij is: Als een leerdoel eerlijk, transparant en betrouwbaar beoordeeld kan worden mét AI, ontwerp dan een rijke toets op basis van de Baan 2- benadering. Als AI-gebruik betrouwbare beslissingen over kwaliteit en beheersing niet mag beïnvloeden, dan moet voor dat leerdoel een gecontroleerde AI-beperkte of -vrije toetsvorm ontworpen worden die in gecontroleerde omgeving onder toezicht wordt afgenomen. Wanneer zwaarwegende beslissingen genomen gaan worden over niveaubeheersing en opbrengsten van een student

richting meerdere leerdoelen op basis van een holistisch oordeel, bijvoorbeeld bij portfolio-toetsing, is het van groot belang om ervoor te zorgen dat de samengestelde set van bewijzen bestaat uit informatie en data (waarbij AI geoorloofd en/of uitgesloten is geweest) vanuit beide sporen, met duidelijkheid over de beoordelingscriteria en de wijze van beoordeling. Alleen zo kunnen valide en betrouwbare beslissingen genomen blijven worden.

Een belangrijk aandachtspunt is het voorkomen van een “grijs gebied”: voor iedere toets/examen moet kristalhelder zijn of en hoe AI mag worden ingezet. Onduidelijkheid over AI-gebruik vormt de grootste bedreiging voor toetsintegriteit (Curtis, 2025). Het leidt tot oneerlijke situaties en mogelijk tot fraudegevallen die examencommissies moeten behandelen. Een toetsmatrix of -blueprint helpt om systematisch te bewaken dat alle essentiële leerdoelen op de juiste wijze worden getoetst. In iedere onderwijssector is de *Two-Lane approach* toepasbaar, al zal de invulling en uitvoering ervan verschillen tussen domeinen, disciplines en onderwijssectoren.

Two-Lane approach

Principe 1: Studenten leren om AI doelmatig, bewust en ethischverantwoord in te zetten.

Principe 2: Betrouwbare toetsing afgestemd op leerdoelen zonder afbreuk aan.

Baan 1: Kwaliteitsmeting van leren en opbrengsten

Doel: Toetsing als eindpunt van leren, waarbij de kwaliteit van studentopbrengsten in relatie tot de leerdoelen valide en betrouwbaar wordt vastgesteld (attributie van leerresultaten), in controleerbare omgeving en onder (gedeeltelijk) toezicht: AI-gebruik beperkt of niet geoorloofd.

Baan 2: Ontwikkelingsgerichte toetsing

Doel: Toetsing als startpunt van en/of gedurende het leerproces. De informatie over (tussentijdse) studentprestaties wordt voor procesgerichte ondersteuning en (bij)sturing van leerprocessen in de richting van de beoogde leerdoelen: AI wordt aangemoedigd of verplicht. Niet volledig controleerbaar/onder toezicht.

Waarom is deze aanpak waardevol? De *Two-Lane approach* laat integriteit en innovatie hand in hand gaan. Studenten leren optimaal gebruikmaken van AI-toepassingen op basis van verantwoord en passend gebruik, én de betrouwbaarheid en kwaliteit van toetsing en eindniveau wordt geborgd. In plaats van AI te negeren of te verbieden, omarmen we de technologie op gecontroleerde wijze waar deze het leerproces verrijkt. Tegelijk voorkomen we zoveel mogelijk ongeoorloofd of onjuist gebruik op die punten waar AI het leren en de bewijsvoering dreigt te ondermijnen. Studenten ontwikkelen zo niet alleen vakbekwaamheid, maar ook het inzicht wanneer en hoe ze hulpmiddelen verantwoord kunnen inzetten. Het raamwerk sluit hiermee aan op internationaal breed gedeelde opvattingen dat adviseert om toetsing enerzijds AI-bewust te maken (zorgdragen voor kwaliteit van toetsing dat standhoudt ondanks en dankzij AI) en anderzijds studenten juist stimuleert om (in samenwerking) mét AI relevante kennis en vaardigheden op te doen (TEQSA, 2023). Deze aanpak vergt zorgvuldige communicatie naar alle betrokkenen:

- Studenten moeten begrijpen waarom sommige toetsen AI-vrij zijn en andere juist AI-integratie vereisen;
- Docenten hebben professionalisering nodig om AI-bewuste toetsen te ontwerpen;
- Examencommissies moeten heldere richtlijnen opstellen over toegestaan AI-gebruik;
- Externe stakeholders (werkgevers, accreditatie-instanties) moeten vertrouwen hebben in de waarde van diploma's;

Kort gezegd omarmt de *Two-Lane approach* het motto: “*Leer met en over AI, maar beoordeel essentieel begrip*”. Praktisch betekent dit AI-inzet uitsluiten waar dat nodig is, en AI integreren in onderwijs en toetsing waar dat mogelijk of noodzakelijk is. Zo blijft niet alleen de ontwikkeling van de student centraal staan, maar blijft ook onderwijs- en toetskwaliteit behouden en leert de student navigeren in een AI-gedreven wereld, in plaats van eromheen.

4. Naar AI-bewuste toetstaken: Stappenplan

De *Two-Lane approach* ondersteunt AI-integratie in toetstaken en draagt zo bij aan AI-bewuste toetsprogramma's, doordat het AI inzet als middel om leerprocessen en beoogde opbrengsten te ondersteunen en/of versterken, niet als doel om leertaken/-processen te vervangen of uit handen te geven. AI-inzet mag in geen geval de bewijsvoering over beheersing van leerdoelen of eindtermen in zijn geheel overnemen. Het raamwerk sluit hiermee aan op internationaal breed gedeelde opvattingen dat adviseert om toetsing enerzijds AI-bewust te maken (zorgdragen voor kwaliteit van toetsing dat standhoudt ondanks en dankzij AI) en anderzijds studenten juist stimuleert om (in samenwerking) mét AI relevante kennis en vaardigheden op te doen (TEQSA, 2023). De kracht zit in het combineren van beide opvattingen.

Bij de inzet van de *Two-Lane approach* wordt ervan uitgegaan dat het zinloos is om AI te willen beperken op momenten waar er sowieso geen toezicht kan plaatsvinden. Cotton et al. (2024) bevestigen dit uitgangspunt en stellen dat effectieve strategieën zich moeten richten op doorgaande toetsing, transparante communicatie over AI-gebruik, en het ontwikkelen van academische integriteit in plaats van enkel op detectie en verbod. Daarom wordt voorgesteld om enerzijds *bewust* toetsactiviteiten onder toezicht in gecontroleerde omgevingen uit te voeren, wanneer de kwaliteit van leren en opbrengsten moet worden vastgesteld en men zicht wil hebben op de eigen inspanning van de student bij de totstandkoming van de te beoordelen opbrengst. Daarbij kan AI ook worden toegestaan als dat volgt uit de geformuleerde leerdoelen. Tegelijkertijd wordt voorgesteld om het gebruik van AI **vrij** toe te staan of te verplichten. De toetsactiviteiten die niet in controleerbare contexten plaatsvinden of waarbij geen/ minder toezicht mogelijk is, biedt hiervoor mogelijkheden. Beide toetsvormen, op basis van de verschillende toetsfuncties, leiden tot informatie en data, die meegenomen kunnen worden in beoordelingen, afhankelijk van de beoogde leerdoelen.

Huidige toetstaken

Bij het (her)ontwerpen van toetstaken is het belangrijk om eerst kritisch te bekijken waar zich mogelijke kwetsbaarheden bevinden:

- Waar hebben we twijfels of het leren bij de student blijft?
- Waar hebben we twijfels over de betrouwbaarheid van toetstaken over leeropbrengsten?
- Waar leren we studenten bewust en verantwoord AI te gebruiken bij de ondersteuning van leren, en niet het ondermijnen?
- Hoe leren we studenten AI ethisch en verantwoord te gebruiken (focus zowel op technische aspecten, als op pedagogisch-didactische implicaties)?

Herontwerp van toetstaken

1. Beoordeel de leerdoelen, herformuleer of vul aan: Een toetstaak ontwerpen begint bij de leerdoelen. Bepaal welke kennis, vaardigheden en houding getoetst gaan worden. Zijn de leerdoelen nog passend, duidelijk, relevant of is herformulering en/of aanvulling nodig? Moeten er AI-vaardigheden toegevoegd worden bij het aangescherpte leerdoel? (Bijvoorbeeld: “kan een betoog schrijven” blijft een leerdoel, maar wellicht voegen we toe “kan effectief tools als AI inzetten bij het schrijfproces en de beperkingen daarvan toelichten”. Formuleer waar mogelijk ook leerdoelen rond procesvaardigheden (zoals kritisch redeneren, analyseren, verbanden leggen) omdat AI vooral invloed heeft op productgerichte doelen).
2. Ga na welke invloed AI-gebruik mogelijk heeft op de huidige toetstaak, ten aanzien van het ontwikkelen en/of behalen van de leerdoelen door de student. Gebruik hierbij het competency framework voor feedbackgeletterde docenten (Boud & Dawson, 2023) om te bepalen welke aspecten van de toetstaak mogelijk door AI kunnen worden overgenomen en waar menselijke expertise onmisbaar blijft.
3. Evalueer de huidige toetsvorm op geschiktheid: Check of de bestaande toets, op basis van toetsfunctie, nog aansluit op de leerdoelen, gegeven de AI-mogelijkheden. Een traditionele take-home essaytentamen is mogelijk minder geschikt geworden (omdat AI makkelijk kan meehelpen), terwijl een presentatie of een essay geschreven onder toezicht wel de student zelf dwingt tot presteren. Zo nee, overweeg aanpassingen. Dit kan variëren van kleine wijzigingen (bv. extra mondelinge vragen toevoegen om begrip te controleren) tot een geheel andere toetsvorm kiezen. In het mbo kan dit betekenen dat traditionele kennistoetsen worden aangevuld met praktijksimulaties waarbij Mixed Reality-technologie wordt ingezet (Bödding et al., 2025), in het hbo dat portfolioassessment wordt verrijkt met verplichte reflecties op AI-gebruik, en in het wo dat peer review processen worden ondersteund door tools om academische discussies te faciliteren (Wood, 2022). Documenteer deze afwegingen, zeker in het mbo/hbo, zodat de examencommissie of opleidingscommissie ze kan meenemen in haar vaststelling van de toets.
4. Bepaal de functie van de toets, en hanteer bij het herzien of ontwerpen van de toetstaak op grond van de *Two-Lane approach*, met oog voor de kaders, mogelijkheden en richtlijnen bij baan 1, baan 2 of een combinatie daarvan.
5. Analyseer de rol van AI in de toetstaak en gewenste integratie: Onderzoek of – en zo ja, hoe – AI ingezet kan worden om het leren richting de toets te ondersteunen. Deze stap gaat over didactische integratie van AI. Kan AI dienen als oefentool, feedbackgever of creatieve sparringpartner voor studenten terwijl ze toewerken naar de toets? Zo ja, moedig dat dan aan met duidelijke richtlijnen. AI-gegenereerde feedback kan even effectief zijn

als menselijke feedback voor schrijfvaardigheden, maar studenten hebben behoefte hebben aan duidelijke begeleiding over hoe deze feedback te interpreteren en toe te passen (Escalante et al., 2023). Een voorbeeld: “Studenten mogen voor dit verslag ChatGPT gebruiken om feedback op taal en structuur te krijgen, maar niet om inhoudelijke paragrafen te laten schrijven.” Bedenk ook welke AI-vaardigheden studenten nodig hebben om hier verstandig mee om te gaan (bijv. kritisch beoordelen van AI-output, juiste prompts schrijven). Neem eventueel zulke vaardigheden op in het leerdoel. Als AI inzet niet of beperkt wenselijk is bij het leren (bijv. omdat het de kernvaardigheid ondermijnt), bedenk dan of en hoe je dat gaat communiceren en handhaven.

6. Beoordeel de risico's van AI-gebruik bij deze toets: Breng in kaart waar en hoe studenten bij de toetstaken in de verleiding zouden kunnen komen ongewenste AI-hulp te gebruiken. Vraag jezelf af: Kan een student AI-inzetten waar dat niet gewenst of geoorloofd is? Zo nee, dan kan de toetstaak behouden blijven. Als het risico onacceptabel is, en uit te leggen is waarom en voor wie, dan moet je maatregelen treffen: de toetsvorm aanpassen (zie stap 2) of beheersmaatregelen invoeren (zie stap 6 en 7 hieronder). Houd hierbij rekening met de EU AI Act die vanaf 2025 strenge eisen stelt aan AI-systemen in educatieve contexten, vooral wanneer deze beslissingen nemen over toegang tot onderwijs of beoordeling van studenten (high-risk categorie). Dit vereist transparantie, menselijke controle en toezicht (European Commission, 2023). Houd ook rekening met technisch-ethische risico's: Als studenten AI gebruiken, delen ze mogelijk data – is dat veilig en toegestaan? Dit hoort ook bij de risicoanalyse.
7. Zorg voor transparantie in verwachtingen en regels: Stel duidelijke richtlijnen op voor AI-gebruik, zowel richting studenten als betrokken docenten. Studenten moeten vóór ze aan de toets beginnen weten of ze AI mogen gebruiken of niet, wanneer, waarom en in welke vorm. Bijvoorbeeld: “Voor deze opdracht mag je generatieve AI gebruiken om ideeën te brainstormen. Vermeld in een bijlage hoe en welke AI-tools je hebt gebruikt.” Transparantie betekent ook dat je als instelling en opleiding consistent bent: leg vast in toetsafspraken of protocol wat de algemene lijn is (met ruimte voor vak/specifieke invulling). Dit draagt bij aan eerlijkheid: iedereen weet waar hij/zij aan toe is, en het voorkomt verwarring (“Ik dacht dat ik het wel mocht gebruiken want bij vak X mocht het ook”).
8. Waarborg validiteit en betrouwbaarheid van de beoordeling: Ontwikkel concrete maatregelen om ervoor te zorgen dat de beoordeling van het studentwerk eerlijk en inhoudelijk kloppend is, ondanks (of juist met dank aan) eventuele AI-inzet. Voorbeelden van maatregelen: extra mondelinge checks (bijv. een kort oral exam naast een ingeleverd essay, om te toetsen of de student het geschrevene werkelijk beheerst), versleutelde of gecontroleerde toetsafnames (bij digitale examens: surveillance, lockdown browsers indien AI

verboden is), of juist AI-inzet onderdeel maken van de beoordeling (zie stap 4 en voorbeeld hieronder). Menselijke kwaliteitscontrole blijft cruciaal, vooral bij zwaarwegende beslissingen. Jung et al. (2025) adviseren een hybride model waarbij AI-scoring altijd wordt aangevuld met steekproefsgewijze menselijke verificatie.

9. Vaststelling: In het mbo is het gebruikelijk dat exameninstrumenten door een examencommissie worden vastgesteld op kwaliteit (validiteit, betrouwbaarheid, etc.). Indien AI nieuwe risico's introduceert, moeten daar in de vaststellingsprocedure expliciet criteria voor zijn: is de toetsvorm passend en zijn afdoende maatregelen genomen zodat de examencommissie kan verantwoorden dat de toetsresultaten dekken wat ze moeten dekken (conform bijv. WEB artikel 7.4.5a in het mbo)? In het hbo dienen examencommissies conform artikel 7.12b WHW toe te zien op de kwaliteit van tentamens en examens, waarbij AI-gebruik expliciet in de onderwijs- en examenregeling (OER) moet worden opgenomen. In het wo geldt aanvullend dat promotiereglementen moeten worden aangepast wanneer AI-tools worden ingezet bij onderzoek (VSNU, 2024).] Deze stap kan per sector verschillen in invulling, maar het principe is overal van belang.
10. Integreer AI in toetsstaken waar relevant en mogelijk: Deze stap is feitelijk het omkeren van de risico-benadering: kijk naar kansen om AI een positieve plek te geven in de toets. Kun je AI inzetten om authentiek te toetsen of hogere denkvaardigheden uit te lokken? Door AI in de opdracht te verweven, vergroot je ethisch-verantwoord gebruik en omzeil je een gebrek aan transparantie doordat het onderdeel van het leren en/of de opbrengst is. Uiteraard moet dit wel in lijn zijn met je leerdoelen en eerlijk blijven voor iedereen (dus alleen doen als alle studenten toegang hebben tot dezelfde tool). Enkele ideeën uit de praktijk: laat studenten AI gebruiken om tegenargumenten te genereren die ze vervolgens in een debat moeten pareren (toetsing van argumentatievaardigheid); geef een opdracht waarbij studenten een door AI geschreven tekst moeten herkennen en verbeteren (beoordeel kritisch denkvermogen); of laat studenten tijdens een practicum een AI-assistent raadplegen zoals ze dat in het werkveld ook zouden kunnen doen, en beoordeel hen op hoe ze die assistent benutten.
11. Ontwikkel een diversiteit aan toetsvormen: Baseer je niet op één enkel type toetstaak voor een belangrijk leerdoel, maar combineer verschillende vormen zodat zwakke plekken van de een gecompenseerd worden door de ander. Bijvoorbeeld: in plaats van alleen een schriftelijk examen (waarbij AI eventueel zou kunnen helpen), organiseer je een mix: een praktijkopdracht, een presentatie en een schriftelijke toets. AI zal niet op al die fronten even effectief of toegestaan zijn, waardoor de student meerdere vaardigheden moet laten zien. Deze mix aan toetsvormen verhoogt de kans dat je een volledig beeld krijgt van iemands competentie en maakt het moeilijker om met één trucje (AI of anderszins) het systeem

te omzeilen. Bovendien maakt variatie de toetsing vaak interessanter en meer verbonden met de praktijk. Let er wel op dat de belasting voor student en docent haalbaar blijft; diversificatie moet niet ontaarden in veel méér toetsdruk, maar slim gekozen onderdelen.

12. Evalueer en optimaliseer het toetsbeleid continu: De ontwikkeling rond AI gaat razendsnel. Wat dit jaar nog toekomstmuziek leek, kan volgend jaar werkelijkheid zijn. Instellingen dienen daarom een lerende aanpak te hanteren voor hun toetsbeleid met AI. Plan na elke toetscyclus een evaluatie: zijn er incidenten geweest (fraude of andere problemen)? Hoe hebben studenten de regels ervaren? Zijn de leerdoelen behaald of zagen we onverwachte effecten van AI-gebruik? Gebruik deze feedback om het beleid en de toetsstaken bij te stellen. Dit kan formaliseren via bijv. de examencommissie of kwaliteitszorg: maak "inzet van AI" een vast agendapunt in nabesprekingen. Daarnaast is het belangrijk om onderwijsprofessionals bij te scholen. Professionalisering moet zich hierbij niet alleen richten op technische vaardigheden maar ook op het versterken van vertrouwen en begrip. Zo blijft de AI-bewuste toetsing niet alleen toekomstbestendig, maar kan men er ook proactief op anticiperen.

Aanbevelingen: betrouwbare en inclusieve toetsstaken

In beide sporen is het belangrijk oog te hebben voor betrouwbaarheid (consistente, objectieve beoordeling) en inclusiviteit (toetsing werkt voor alle studenten eerlijk). Enkele aandachtspunten:

- Combineer verschillende toetsvormen: Vertrouw niet op één enkel toetsinstrument. Een mix zorgt voor een completer beeld. Bijvoorbeeld: beoordeel schrijfvaardigheid niet alleen via één geschreven examen, maar ook via observatie van schrijfproces of een mondelinge toelichting. Voor praktische vaardigheden: combineer een praktijkproef met een theoretische reflectie.
- Duidelijke beoordelingscriteria: Bij toetsing moet voor zowel beoordelaars als studenten helder zijn wat er beoordeeld wordt, wanneer en door wie. Maak hierbij duidelijk hoe met AI-inzet omgegaan wordt.
- Professionalisering en vergroten van toetsbekwaamheid: In alle onderwijssectoren is het cruciaal dat beoordelaars toetsbekwaam zijn en bij de uitvoering van toetsing op één lijn zitten, zeker nu AI een factor van invloed kan zijn. Investeer in professionalisering van onderwijsaanbieders en beoordelaars ten aanzien van toetsbekwaamheid met expliciet aandacht voor AI-geletterdheid bij docenten, kennis en inzicht in de impact van AI op toetsing, het benutten van kansen en mogelijkheden, en proactief anticiperen en/of beheersen van risico's en ethische en juridische implicaties. Kohnke et al. (2025) benadrukken dat comprehensive AI literacy curricula voor docenten moeten omvatten: technische vaardigheden, pedagogische integratie, ethisch begrip, en praktische ervaring met AI-tools.
- Toegankelijkheid: AI kan voor sommige studenten juist een ondersteunend hulpmiddel zijn (denk aan studenten met dyslexie die een taalmodel gebruiken voor spellingscontrole). Wees je bewust van voorzieningen en redelijke aanpassingen. Bakker et al. (2024) waar-

schuwen echter dat AI ook nieuwe vormen van ongelijkheid kan creëren wanneer niet alle studenten gelijke toegang hebben tot premium AI-tools of de digitale vaardigheden om deze effectief te gebruiken.

- Tenslotte: privacy en ethiek horen ook bij betrouwbaarheid. Forceer studenten niet om privacy-gevoelige dingen te doen met AI (zoals persoonlijke data uploaden naar een externe dienst) en houd je aan AVG-richtlijnen. Kies waar mogelijk institutioneel goedgekeurde AI-tools die veilig zijn. Dit alles draagt bij aan een toetsprogramma dat passend, relevant en verantwoord is.

Checklist: Toetstaken met AI-integratie

1. **Uitnodiging tot eigen werk? In hoeverre** nodigen de leeractiviteiten/toetstaken de student uit om hier actief en op basis van eigen inspanningen aan bij te dragen, of kan het volledig of deels uit handen gegeven worden (of dit nu geoorloofd, gewenst, of verplicht is, of juist niet)?

2. Rolverdeling AI versus student?

Is het gebruik van AI bewust onderdeel van het leerdoel?

Indien ja: Integreer het expliciet. Leg vast hoe en waar student en AI samenwerken in de opdracht, en hoe dat een onderdeel vormt van de beoordeling. Bijvoorbeeld: “Student gebruikt AI om minimaal 3 suggesties voor verbetering te krijgen op zijn ontwerp en reageert daarop in een logboek.”

Indien nee: Overweeg of je dit expliciet wilt vermelden in het leerdoel of de beoordelingscriteria. Bijvoorbeeld: “De student demonstreert individuele schrijfvaardigheid zonder hulp van AI.” In dat geval moeten de toetsvorm en -setting dat mogelijk maken (bv. klassikaal schrijven zonder hulpmiddelen). Communiceer duidelijk dat voor deze opdracht geen AI gebruikt mag worden en waarom. Richt de omgeving waarin getoetst gaat worden zo in dat dit haalbaar en controleerbaar is.

3. Aanpassing bij negatieve impact?

Stel dat AI-gebruik de leerervaring toch ondermijnt of ethisch ongewenst is in een bepaalde opdracht, pas aan.

Actie: Wees flexibel en creatief. Voeg bijvoorbeeld extra toetstaken toe. Of hervorm de opdracht zodat AI inzetten minder direct de oplossing geeft (bijv. maak de opdracht complexer of meer gepersonaliseerd). Het curriculum mag evolueren: als docenten signaleren dat studenten door AI minder leren, grijp in.

4. Authenticiteitsrisico: Hebben we voldoende zekerheid dat de toetsresultaten echt door de student zelf zijn behaald? Met andere woorden, is het risico afgedekt dat AI of andere bronnen onterecht hebben bijgedragen?

Actie: Bij twijfel of nee? Tref aanvullende maatregelen.

Enkele opties:

- Ander toezichtniveau: Verplaats de toets naar een setting met direct toezicht
- Directe verificatie: Voeg een observatie, mondelinge verdediging of een criteriumgericht interview toe.
- Documentatieplicht: Laat de stappen/handelingen in de richting van de gerealiseerde opbrengst zichtbaar maken. Vereis bijvoorbeeld prompt-logging, de student houdt bij welke prompts aan AI gegeven zijn, welke output kwam en hoe die is gebruikt. Dit bewijsmateriaal kan worden bekeken om te zien hoe, of in welke mate de student zijn samenwerking met AI benut heeft, of passend gebruikt heeft
- Geobserveerd leerproces: Laat (delen van) het maakproces onder toezicht oog plaatsvinden. Bijvoorbeeld: belangrijke vraagstukken in tussentijdse toetsing in de les uitwerken om te zien hoe ieder individueel denkt.
- Duidelijke beperkingen vooraf: Bepaal en communiceer expliciet welke AI-hulpmiddelen wel of niet mogen, en waarvoor.

5. Herijking leerdoelen^a:

Moeten we onze leerdoelen aanpassen vanwege AI?

Actie: In verschillende gevallen, ja. Als bepaalde vaardigheden door AI minder relevant worden in het werkveld, kun je overwegen het leerdoel bij te stellen. Als bepaalde bestaande vaardigheden door AI relevanter worden, of nieuwe vaardigheden noodzakelijk zijn in het werkveld, kun je het leerdoel uitbreiden of toevoegen. Deze afwegingen verschillen per discipline, domein en toekomstig beroepenveld. Betrek het werkveld en de actuele beroepscontext hierbij. Mbo en hbo moeten vooral naar de praktijkrelevantie kijken: welke competenties wil de werkgever zien, en hoe verhoudt AI zich daartoe? WO moet kijken naar academische vaardigheden: bijvoorbeeld, blijft het belangrijk om literatuuronderzoek handmatig te doen, of mag AI daarbij assisteren zolang kritisch denken behouden blijft? Pas leerdoelen aan als ze echt niet langer passen bij wat een afgestudeerde nodig heeft – maar borg altijd de kern van de opleiding.

Bevorderen van bewuste en ethisch-verantwoorde AI-inzet door studenten

Als AI een plek krijgt in het onderwijs, moeten studenten leren hoe ze er eerlijk en verantwoord mee omgaan. Dit betekent dat we afspraken maken over transparantie in AI-gebruik:

- Erkenning/vermelding van AI-gebruik: Spreek af dat studenten altijd aangeven als en hoe ze generatieve AI hebben gebruikt bij het maken van een opdracht. Bijvoorbeeld: voeg een standaardzin toe aan elke uitwerking: “Ik heb voor deze opdracht gebruikgemaakt van [AI-

^a Met leerdoelen worden leerresultaten of leeruitkomsten bedoeld. In het mbo zijn dit de generieke en specifieke kwalificatie-eisen zoals beschreven in de kwalificatiedossiers, evenals de eisen van de keuzedelen.

tool] voor [doel]. Ik heb de output verwerkt door ...”. Bij gering gebruik kan dit een korte noot zijn, bij intensiever gebruik vraag je wellicht om een apart AI-gebruikslogboek of appendix. Bijvoorbeeld: laat studenten hun prompts en de gegenereerde teksten als bijlage meesturen, of een reflectie schrijven op wat de AI bijdroeg en wat zij zelf deden. Dit maakt AI-gebruik bespreekbaar en zichtbaar.

- Begeleiding in correct gebruik en citatie: Leer studenten dat AI geen klassieke bron is (het is geen auteur van ideeën). Ze hoeven ChatGPT niet als auteur in de literatuurlijst te zetten, maar wel te vermelden dat ze het gebruikt hebben en hoe (zoals bovenstaand). Het APA Style Guide (2024) adviseert om AI-gebruik te documenteren in de methodesectie of als voetnoot, waarbij de gebruikte tool, versie, datum en specifieke prompts worden vermeld. Bespreek dat content van AI niet klakkeloos vertrouwd kan worden, en dat ze altijd kritisch moeten valideren. Ethisch gebruik is ook: geen gevoelige data invoeren, biases herkennen, AI niet gebruiken waar het niet mag.
- Detectiesoftware: AI-detectie tools zijn niet altijd volledig betrouwbaar. Het advies is om detectie alleen als startpunt voor gesprek te gebruiken, nooit als definitief bewijs. We focussen op bewustwording en afspraken, in plaats van straffen op basis van niet volledige detectie. Bij vermoeden van fraude leunen we op degelijker methoden (zoals een gerichte mondeling of vergelijking met eerder werk van de student).
- Begeleiding en ondersteuning: Integreer korte training voor studenten in hoe ze AI kunnen gebruiken voor leren. Bijvoorbeeld workshops “omgaan met ChatGPT”, waarin ook de grenzen besproken worden. Zorg dat studenten weten dat ze bij vragen altijd hun docent kunnen raadplegen zonder bang te hoeven zijn meteen beschuldigd te worden. Een open dialoog is de beste manier om te zorgen dat studenten juist met AI omgaan (Lodge et al., 2023; zie ook AAIN-guidelines 2023 in Australië die dit benadrukken).

Door expliciet te maken wat wel en niet mag, en studenten hierin te begeleiden, normaliseren we AI als hulpmiddel op dezelfde manier zoals we bijvoorbeeld boeken, rekenmachines of internetbronnen behandelen. Het wordt onderdeel van de benodigde vaardigheden van de student om te weten wanneer AI gepast is en hoe dit te verantwoorden. Dit draagt bij aan een cultuur van integriteit en verantwoord gebruik in AI-gedreven onderwijs.

Conclusie

Deze handreiking heeft geschetst hoe onderwijsinstellingen in mbo, hbo en wo hun toetsen kunnen herzien in het tijdperk van AI. Centraal staat de balans tussen twee doelstellingen uit de landelijke AI-visie: het ondersteunen van leren (met AI waar mogelijk) en het valideren van leerresultaten (met zekerheid over de eigen prestatie). De voorgestelde *Two-Lane approach* biedt een helder kader om dit te operationaliseren. Ontwikkelingsgerichte toetsing en validerende toetsing zijn beide onmisbaar; door ze bewust in te zetten versterken ze elkaar.

Concreet betekent dit:

- Onderwijsprofessionals ontwerpen toetsen die relevant zijn voor de AI-gedreven wereld waarin studenten zullen werken, en nemen AI-vaardigheden mee als leerdoel waar passend.
- Tegelijk bewaken ze de validiteit en betrouwbaarheid van toetsen waarmee besloten wordt over leren, zodat een diploma blijft staan voor daadwerkelijk behaalde competenties.
- Via stappen van analyse, (her)ontwerp, duidelijke afspraken, begeleiding en continue evaluatie (analyseren – diagnosticeren – concluderen – bijstellen) wordt een toekomstbestendig toetsprogramma opgebouwd. Dit is geen eenmalige exercitie maar een voortdurend proces van kwaliteitszorg.
- Alle stakeholders, management, docenten, examencommissies, studenten, worden betrokken en geïnformeerd, zodat er een gedeeld begrip is van wat wel en niet mag rond AI in toetsing. Heldere communicatie en regelgeving hierover voorkomt misstanden en schept ruimte voor innovatie.

Ten slotte is het van belang te onderstrepen dat AI geen bedreiging hoeft te zijn voor toetsing, mits we onze aanpak erop aanpassen. Het dwingt ons vooral om nóg explicieter te maken waarom we toetsing kunnen en willen inzetten, wat we willen toetsen en hoe. Het is hierbij van belang dat mbo, hbo en wo van elkaar leren in deze snelle ontwikkelingen. Zoek elkaar op, werk samen, en deel *best practices*. AI in toetsing en examinering is nieuw terrein voor iedereen; een cultuur van kennisdeling en open dialoog met alle betrokkenen zorgt ervoor dat we hier als onderwijsaanbieders coherent en effectief mee om kunnen gaan.

Literatuur

- Andreou, E., Tiemens, K., Draaisma, H., & Kusurkar, R. (2024). Identifying stakeholder needs for workplace-based assessment in postgraduate medical education: A focus group study. *BMC Medical Education*, 24(1), 659. doi.org/10.1186/s12909-024-04572-9
- APA Style Guide. (2024). *How to cite ChatGPT and other AI tools*. American Psychological Association. apastyle.apa.org/blog/how-to-cite-chatgpt
- Bakker, T. C., De Kraker, C. C., Eegdeman, I. M., Saçan, E., & Van Leeuwen, A. (2024). *AI en kansengelijkheid. Visie op de impact van AI op kansengelijkheid in het vervolgonderwijs* (pp. 1-56). Utrecht: Npuls.
- Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893-905.
- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (4th ed.). Maidenhead: Open University Press.
- Black, P., & Wiliam, D. (1998). *Assessment and classroom learning*. *Assessment in Education*, 5(1), 7-74.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Bödding, R., Schriek, S. A., & Maier, G. W. (2025). A systematic review and meta-analysis of mixed reality in vocational education and training: Examining behavioral, cognitive, and affective training outcomes and possible moderators. *Computers & Education*, 212, 105098.
- Boud, D., & Dawson, P. (2023). What feedback literate teachers do: An empirically-derived competency framework. *Assessment & Evaluation in Higher Education*, 48(2), 158-171.
- Boud, D., Dawson, P., Tai, J., & Ajjawi, R. (2023). Creating an agenda for developing students' evaluative judgement. In D. Carless & N. Winstone (Eds.), *Feedback for learning: Closing the assessment loop* (pp. 45-63). Routledge.
- Chua, Y., & Habib, M. (2023). Plagiarism in the Digital Age: A Case for AI-powered Detection Systems. *Educational Technology Journal*, 15(2), 105-117.
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239.
- Crocco, M. S. (2025). *Assessment in the age of AI: Three phases of learning*. Educational Researcher.
- Cukurova, M. (2025). The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence. *British Journal of Educational Technology*, 56(1), 23-45.
- Curtis, G. J. (2025). The two-lane road to hell is paved with good intentions: Why an all-or-none approach to generative AI, integrity, and assessment is insupportable. *Higher Education Research & Development*, 1-8.
- Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity*, 19(1), 23.
- European Commission. (2023). *Proposal for a Regulation laying down harmonised rules on Artificial Intelligence (AI Act)*. Brussels: EC.
- Evangelista, E. D. L. (2025). Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education. *Contemporary Educational Technology*, 17(1), ep559.
- Fawns, T., & Schuwirth, L. (2024). Rethinking the value proposition of assessment at a time of rapid development in generative artificial intelligence. *Medical Education*, 58(1), 14-16.
- Guan, L., Zhang, Y., & Gu, M. M. (2025). Pre-service teachers preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teachers' identity changes. *Computers and Education: Artificial Intelligence*, 6, 100341.
- Ifenthaler, D., Sahin, M., Boo, I., Rajasegeran, D. D., & Yuh, A. S. (2025). Adaptive serious games assessment: The case of the blood transfusion game in nursing education. *Computers in Human Behavior*, 150, 108234.
- Jung, J. Y., Tyack, L., & Von Davier, M. (2025). Towards the implementation of automated scoring in international large-scale assessments: Scalability and quality control. *Educational Measurement: Issues and Practice*, 44(1), 23-39.
- Katona, J., & Gyonyoru, K. I. K. (2025). Integrating AI-based adaptive learning into the flipped classroom model to enhance engagement and learning outcomes. *Interactive Learning Environments*, 1-18.
- Kennispunt MBO Onderwijs & Examinering. (2024). *Lerend kwalificeren in het mbo*. Utrecht: Kennispunt.
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614.
- Laupichler, M. C., Aster, A., Perschewski, J.-O., & Schleiss, J. (2023). Evaluating AI Courses: A Valid and Reliable Instrument for Assessing Artificial-Intelligence Learning through Comparative Self-Assessment. *Computers and Education: Artificial Intelligence*, 4, 100123.
- Lim, J., Lee, U., Koh, J., Jeong, Y., Lee, Y., Byun, G., Jung, H., Jang, Y., Lee, S., & Moon, J. (2025). Development and Implementation of a Generative Artificial Intelligence-enhanced Simulation to Enhance Problem-Solving Skills for Pre-service Teachers. *Computers & Education*, 232, 105306.
- Liu, D., Fawns, T., Cowling, M., & Bridgeman, A. (2023). *Responding to generative AI in Australian higher education*. doi: 10.35542/osf.io/9wa8p
- Lodge, J. M., Thompson, K., & Corrin, L. (2023a). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1), 1-8.
- Lodge, J. M., de Barba, P., & Broadbent, J. (2023b). Learning with generative artificial intelligence within a network of co-regulation. *Journal of University Teaching and Learning Practice*, 20(7), 1-10.
- Lyanda, J. N., Owidi, S. O., & Simiyu, A. M. (2024). Rethinking higher education teaching and assessment in-line with AI innovations: A systematic review and meta-analysis. *African Journal of Empirical Research*, 5(3), 325-335.
- Ma, J. (2025). *Student AI usage patterns in higher education: A global survey*. *International Journal of Educational Technology*.
- Masry Herzallah, A., & Makaldy, R. (2025). Technological self-efficacy and sense of coherence: Key drivers in teachers' AI acceptance and adoption. *Computers & Education*, 213, 105123.
- MBO Digitaal. (2024). *Datadinsdag bij SintLucas: Hoe maak je beleid rond AI en examinering, toetsing en afsluiting?* Verslag, MBO Digitaal/NPULS.
- Micheni, E., Machii, J., & Murumba, J. (2024). The Role of Artificial Intelligence in Education. *Open Journal for Information Technology*, 7, 43-54.
- Ministerie van BZK. (2024). *Overheidsbrede visie generatieve AI*. Den Haag: Rijksoverheid.
- Moccozet, L., Camacho, A., Bowman, A., Roth, P., & Benkacem, O. (2024). Exploring teachers' future digital assessment practices in higher education: Instrument and model development. *Computers & Education*, 198, 104897.
- Mollick, E., & Mollick, L. (2023). Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts. *SSRN Electronic Journal*. doi.org/10.2139/ssrn.4391243
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- Nieminen, J. H., Yan, Z., & Boud, D. (2025). Self-assessment design in a digital world: Centring student agency. *Assessment & Evaluation in Higher Education*, 50(1), 89-104.
- Oregon State University. (2024). *Bloom's Taxonomy Revisited*. Oregon State University Ecampus.
- Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2), 1-24.
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice*, 21(4), 1-24.
- Promma, W., Imjai, N., Usman, B., & Aujirapongpan, S. (2025). The influence of AI literacy on complex problem-solving skills through systematic thinking skills and intuition thinking skills: An empirical study in Thai gen Z accounting students. *Thinking Skills and Creativity*, 51, 101456.
- Schellekens, L. H., Kremer, W. D. J., Van Der Schaaf, M. F., Van Der Vleuten, C. P. M., & Bok, H. G. J. (2023). Between theory and practice: Educators' perceptions on assessment quality criteria and its impact on student learning. *Assessment & Evaluation in Higher Education*, 48(7), 987-1001.

- Sperling, K., Stenberg, C.-J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, 6, 100169.
- Sydney Assessment Framework. (2025). *The Sydney Assessment Framework*. University of Sydney.
- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Teaching and Teacher Education*, 138, 104456.
- TEQSA. (2023). *Standards for ethical AI in higher education: A consultation paper*. Tertiary Education Quality and Standards Agency & University of Technology Sydney.
- Tsai, Y. S., Perrotta, C., & Gašević, D. (2023). Critical perspectives on AI literacy in teacher education: Power, inclusivity, and pedagogical responsibility. *Teaching and Teacher Education*, 121, 103903.
- U.S. Department of Education. (2023). *Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments*. Washington, DC.
- Vargas-Murillo, A. R., de la Asuncion, I. N. M., & de Jesús Guevara-Soto, F. (2023). Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research*, 22(7), 122-135.
- Vereniging Hogescholen. (2021). *Handreiking examencommissies* (editie 2021).
- VSNU. (2024). *Richtlijnen AI-gebruik in universitaire onderzoek*. Vereniging van Universiteiten.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Upper Saddle River, NJ: Pearson Education.
- Wood, J. (2022). Making peer feedback work: The contribution of technology-mediated dialogic peer feedback to feedback uptake and literacy. *Assessment & Evaluation in Higher Education*, 47(8), 1234-1248.
- Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 50(1), 145-162.

Bijlage 1

Sectorspecifieke aanbevelingen en voorbeelden

De impact en toepassing van AI in toetsing verschillen per onderwijssector. Hoewel de grote lijnen (zoals hierboven geschetst) voor iedereen gelden, hebben mbo, hbo en wo hun eigen context, regels en uitdagingen. Hier volgen enkele concrete aanbevelingen en voorbeelden per onderwijssector.

MBO

Context: In het mbo ligt de focus op competenties voor een beroep. Toetsing bestaat vaak uit praktijkexamens, proeven van bekwaamheid en kennisexamens, vaak volgens landelijke kwalificatiedossiers. In het mbo richt de examencommissie zich enkel op het bewaken van toetskwaliteit en rechtmatigheid over examinering, zoals gedefinieerd in de wet WEB. Sinds 2016 gelden sectorale afspraken waarbij exameninstrumenten in principe worden ingekocht bij gecertificeerde examenleveranciers, met de producteisen van Stichting Valide Examens mbo (VEMBO) als norm.

Aanbevelingen voor MBO:

- **Bespreek AI-inzet met examencommissies én examenleveranciers.**: Gecertificeerde examenleveranciers zijn een belangrijke partij in de constructieve afstemming van validerende (mbo: kwalificerende) toetsing op kwalificatiedossier en beroepsuitoefening in een wereld met AI. Dit betekent dat examencommissies actief de dialoog moeten aangaan met leveranciers over AI-bewuste examens. Bij de vaststelling van examens dient de commissie na te gaan of een toetsvorm fraudegevoelig is door AI en of aanpassingen nodig zijn. MBO Digitaal rapporteert dat SintLucas al beleid heeft ontwikkeld waarbij studenten in creatieve opleidingen Adobe Firefly gebruiken voor ontwerp-inspiratie, maar bij beoordeling de eigen creatieve inbreng centraal staat (MBO Digitaal, 2024).
- **Benadruk beroepsauthenticiteit**: Leg bij toetsing de nadruk op situaties die direct uit de beroepspraktijk komen, waar studenten vaardigheden in persoon moeten tonen. Het uitvoeren van medische handelingen of het tonen van klantvriendelijkheid in het hotelwezen kunnen in AI-gestuurde simulaties voorkomen, maar de handelingen moeten door de student zelf uitgevoerd worden. Gebruik AI bij ontwikkelingsgerichte (formatieve) onderwijsactiviteiten en -toetsing door docenten aan te moedigen om in lessen en oefentoetsen AI te (laten) gebruiken als leer-

middel. Onderzoek toont aan dat promptontwikkeling als vaardigheid essentieel is voor AI-geletterdheid (Ng et al., 2021). Laat studenten bijvoorbeeld een AI goede vragen stellen in hun vakgebied en beoordelen hoe nuttig de antwoorden zijn. Of laat ze een door AI gegenereerd scenario oefenen. Zo leren studenten de beperkingen en mogelijkheden kennen. Maak duidelijk dat bij het échte examen wellicht andere regels gelden, maar dat dit hun voorbereiding dient.

Concreet voorbeeld (mbo): In een opleiding Marketing & Communicatie krijgen studenten de opdracht een communicatieplan te maken voor een lokale ondernemer. AI-gebruik: Studenten mogen een tool als ChatGPT gebruiken om bijvoorbeeld een SWOT-analyse te genereren op basis van ingevoerde gegevens. Echter, bij de beoordeling moeten ze mondeling verdedigen waarom ze bepaalde strategische keuzes maken in het plan en wat de betrouwbaarheid is van de AI-output. De docent vraagt bijvoorbeeld: “Hoe kwam je aan deze doelgroepomschrijving?” Waarop de student uitlegt dat hij eerst een ruwe versie door AI heeft laten opstellen en die daarna heeft aangepast op regionale context. Op deze manier wordt AI geïntegreerd, maar met de toets wordt beoordeeld of de student de tool zelf kritisch kan toepassen. De examencommissie heeft dit goedgekeurd, omdat het past bij het leerdoel “een marketingplan kritisch kunnen opstellen en onderbouwen” en AI hier niet het beoordelingscriterium zelf is, maar een tool waaraan de student sturing moest geven.

HBO

Context: In het hbo combineren opleidingen theoretische kennis met toepassing in complexe beroepssituaties. HBO-studenten worden voorbereid op een beroepspraktijk waarin beschikken over digitale vaardigheden steeds belangrijker is. De toetsing varieert van schriftelijke kennistoetsen tot projectverslagen, presentaties, portfolio-assessment en stage-evaluaties. Examencommissies in het hbo bewaken en borgen de kwaliteit van toetsing en diplomering op grond van de Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek (WHW).

Aanbevelingen voor HBO:

- Ontwikkel een instellingsbreed AI-toetsbeleid, omdat hbo-instellingen veel autonomie hebben per opleiding. In het toetsbeleid kunnen richtlijnen opgesteld worden voor AI in toetsing als onderdeel van de examenregeling. Bijvoorbeeld: een beleidsdocument dat kaders geeft (“AI mag geen auteur zijn van ingeleverd werk, maar kan wel als hulpmiddel dienen mits verantwoord vermeld”) en voorbeelden van goede praktijken. De examencommissies kunnen erop toezien dat opleidingen het toepassen.
- Stimuleer authenticiteit en professionaliteit door gebruik te maken van de realiteit dat in de beroepspraktijk professionals ook AI-tools zullen inzetten. In veel beroepen (van marketing tot journalistiek, van logistiek tot HR) wordt al gewerkt met AI. Dit biedt kansen om studenten voor te bereiden op deze realiteit door bijvoorbeeld opdrachten te geven waarbij AI bewust wordt ingezet, gecombineerd met kritische reflectie.

- Werk met portfolioassessments en besteed hierin aandacht aan procesbeoordeling. Toetsing in het hbo zal niet alleen gericht moeten zijn op het beoordelen van een eindproduct op niveau en kwaliteit, maar ook het proces daarnaartoe in de beoordeling mee moeten nemen. Dit sluit aan bij de herziene Bloom’s Taxonomy waarbij hogere cognitieve niveaus zoals evalueren en creëren menselijke vaardigheden vereisen die AI niet kan vervangen (Oregon State University, 2024).

Concreet voorbeeld (hbo): In een HBO-Rechten opleiding krijgen studenten een casus waarin zij een juridisch adviesrapport moeten schrijven voor een fictieve cliënt. AI-gebruik: Ze mogen een AI-tool gebruiken om jurisprudentie op te zoeken of een eerste opzet van het advies te laten formuleren, mits zij alle door AI gegeven content controleren en herformuleren in eigen woorden. Bij inlevering voegen studenten een bijlage toe met de prompt(s) die ze hebben gebruikt en de ruwe AI-output. De beoordeling kijkt naar het uiteindelijke advies (inhoudelijk juist, goed onderbouwd, correcte taal) én naar de kritische houding van de student in het gebruik van AI (bijv. heeft de student onjuistheden van de AI herkend en gecorrigeerd? Is de AI-bron correct vermeld indien gebruikt?). Op deze wijze leert de student professioneel om te gaan met AI, een skill die in de rechtspraak ook van pas komt (denk aan het gebruik van jurisprudentiezoekmachines), terwijl de toetsing valide blijft omdat de docent kan verifiëren dat de student het werk begrijpt en eigen maakt.

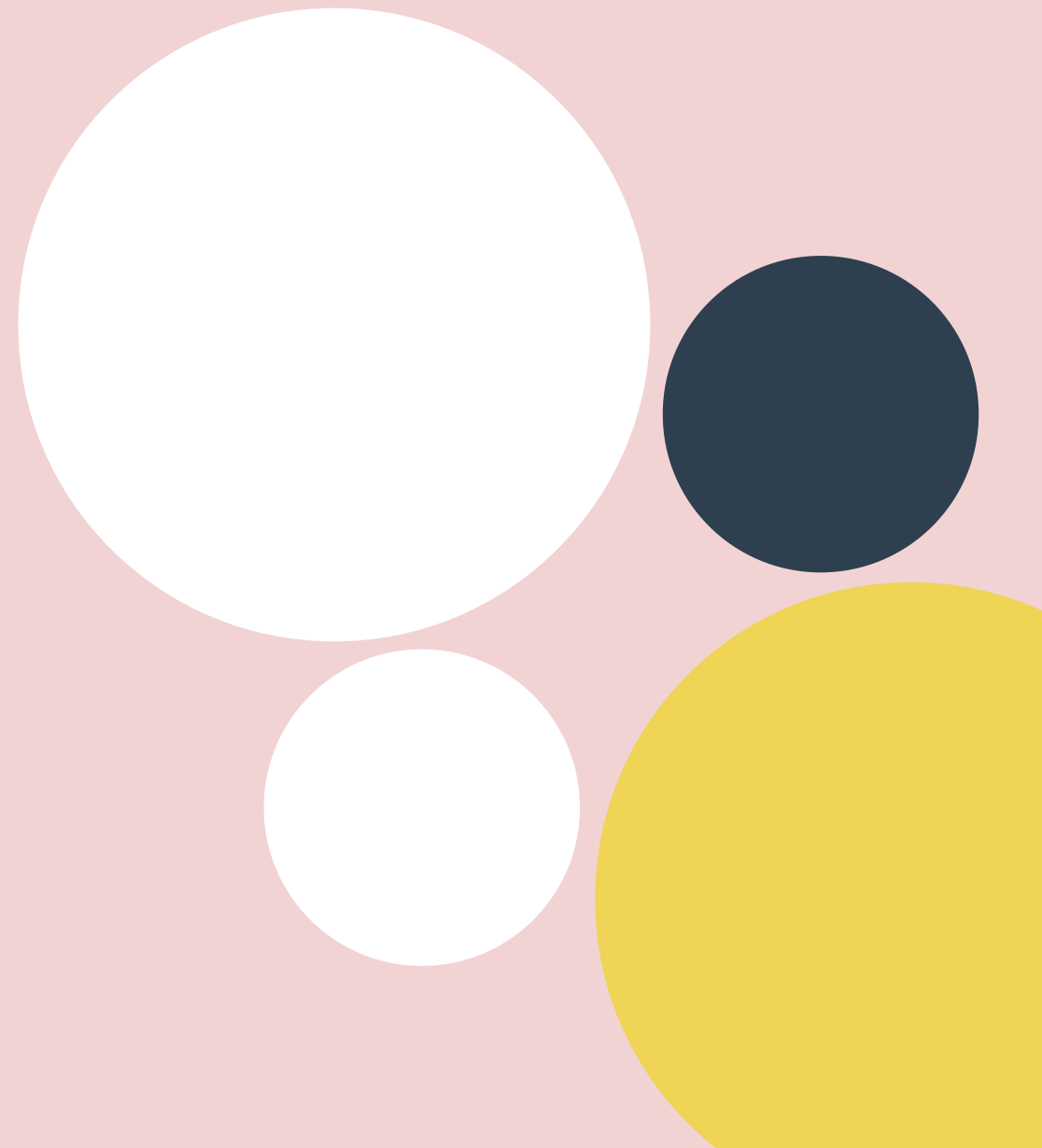
WO

Context: In het wetenschappelijk onderwijs ligt de nadruk op kennisontwikkeling, -creatie, en op academische integriteit, waarbij studenten worden opgeleid tot kritische denkers die zelfstandig complexe informatie kunnen verwerken. Toetsing omvat vaak essays, tentamens, research papers, scripties en mondelinge tentamens. Academische integriteit is een hoeksteen: originaliteit, correcte bronvermelding en zelfwerkzaamheid krijgen veel nadruk.

Aanbevelingen voor WO:

- Herzie toetsvormen door opdrachten te verschuiven naar hogere cognitieve niveaus. Evangelista (2025) pleit voor “rethinking exam design” waarbij traditionele reproductievragen worden vervangen door opdrachten die kritisch denken en creativiteit vereisen.
- Integriteit en ethiek vereisen expliciete aandacht in het curriculum. Er zijn zorgen over het behouden van schriftelijke vaardigheden, en dat AI tot wetenschappelijke onjuistheden of fraude kan leiden, bijvoorbeeld fictieve bronnen of studenten die een scriptie door AI laten schrijven. Universiteiten moeten daarom niet alleen beleid ontwikkelen, maar ook studenten actief onderwijzen over de ethische implicaties van AI-gebruik in academische context.
- Gebruik mondelinge toetsing strategisch als verificatie-instrument. De wetenschapstraditie kent dit principe al eeuwen (colloquia, mondelinge examens), en het blijkt een zeer effectieve manier om te verzekeren dat het ingeleverde werk begrepen is door de student zelf.

Concreet voorbeeld (wo): In een universitaire psychologiecursus moeten studenten een onderzoeksvoorstel schrijven. Studenten mogen AI gebruiken om literatuur te zoeken en outline voor het theoretisch kader te laten genereren. Inhoudelijke keuzes, theoretische uitgangspunten en beargumenteerde methodologische keuzes moet de student mondeling kunnen verantwoorden. Bij inlevering geeft de student aan welke delen met AI-assistentie tot stand zijn gekomen. Vervolgens vindt een mondeling colloquium plaats: de docent stelt vragen als “Waarom heb je voor deze onderzoeksmethode gekozen?” of “Kun je deze theoretische achtergrond nader toelichten?” Hierbij zal duidelijk worden of de student de stof zelf voldoende beheerst of enkel AI-gegenereerde teksten heeft ingeleverd.





Onderwijs
bewegen.