



Uit de praktijk

■ Karin J. Gerritsen-van Leeuwenkamp

Dr. K.J. Gerritsen-van Leeuwenkamp is adviseur toetsing bij de dienst Onderwijs en Student Support van hogeschool Saxion. E-mail: k.j.vanleeuwenkamp@saxion.nl

Illustratie van ChatGPT bij toetsing en aandachtspunten voor weloverwogen gebruik

ChatGPT, een Artificial Intelligence (AI)-model, wordt ook ingezet bij toetsing. Het model reageert op een specifieke instructie, ook wel prompt genoemd, die wordt gegeven om een reactie uit te lokken. In dit artikel worden diverse prompts gedemonstreerd, om de toepassingsmogelijkheden van ChatGPT te illustreren bij toetsing.

Inleiding

ChatGPT is een AI-model dat gebaseerd is op de GPT-3.5-architectuur¹ en getraind is op diverse tekstbronnen. ChatGPT maakt mensachtige interacties mogelijk door te reageren op prompts. Dit artikel verkent, ter illustratie, diverse toepassingsmogelijkheden van ChatGPT bij toetsing, specifiek gericht op de constructie, uitvoering en evaluatie van toetsen. In online-bronnen zijn diverse toepassingsmogelijkheden

gevonden en deze worden in dit artikel in een eigen context en in het Nederlands toegepast, waarbij de gegenereerde reacties van ChatGPT worden weergegeven. Ook zijn eigen prompts uitgetoetst. De focus van dit artikel ligt op de praktische toepassing van ChatGPT bij toetsing, waarbij er geen uitgebreide kritische beschouwing wordt gegeven van de toepassing van AI in het onderwijs. Niettemin worden aan het einde wel enkele aandachtspunten be-

¹In dit artikel is gebruikgemaakt van de gratis variant van ChatGPT, die gebaseerd is op GPT-3.5. Er is ook een betaalde variant beschikbaar ChatGPT Plus, die gebaseerd is op GPT-4. GPT staat voor Generative Pre-trained Transformer.

noemd die relevant zijn bij het weloverwogen gebruik van dergelijke prompts in de dagelijkse onderwijspraktijk. Ook bij het schrijven van dit artikel is gebruikgemaakt van ChatGPT om onder meer de tekst te verbeteren, voorbeelden te genereren en aanvullende prompts te bedenken (OpenAI, 2023).

ChatGPT bij de constructie van toetsen

Hieronder worden vier toepassingsmogelijkheden bij de constructie van toetsen beschreven te weten de constructie van toetsvragen (Tabel 1), opdrachten (Tabel 2), beoordelingsinstrumenten (Tabel 3) en toetsprocedures (Tabel 4).

Tabel 1. Toepassingsmogelijkheid: Constructie van toetsvragen (V. Kalis, persoonlijke communicatie, 21 juni 2023; OpenAI, 2023; Patel, 2023; Dr. Alex Young, 2023)	
Prompt	Reactie ChatGPT
a. Maak een meerkeuzevraag over inflatie met 1 correct antwoord en 2 afleiders. De antwoorden moeten gelabeld worden van A tot C. Elk antwoordalternatief moet een uitleg bevatten van maximaal 1 zin.	Welke van de volgende verklaringen beschrijft inflatie het meest accuraat? A) Inflatie is de geleidelijke stijging van prijzen van goederen en diensten in een economie. B) Inflatie is de vermindering van de beschikbare banen in een land. C) Inflatie is de afname van de waarde van de nationale munteenheid. Uitleg bij elk antwoordalternatief: A) Dit is het juiste antwoord, aangezien inflatie verwijst naar de geleidelijke stijging van prijzen van goederen en diensten in een economie. B) Dit beschrijft werkgelegenheidskwesties, niet specifiek inflatie. C) Dit beschrijft valutadevaluatie, niet specifiek inflatie.
b. Formuleer drie antwoordalternatieven bij de vraag: "Wat is de definitie van ziekte?" met 1 correct antwoord en 2 afleiders.	A) Ziekte wordt gedefinieerd als een abnormale fysieke of mentale toestand die de normale functies van het lichaam verstoort. B) Ziekte wordt gedefinieerd als een tijdelijke toestand van ongemak of malaise veroorzaakt door externe factoren. C) Ziekte wordt gedefinieerd als een natuurlijk proces van veroudering en slijtage van het lichaam. Uitleg bij elk antwoordalternatief: A) Dit is het juiste antwoord, omdat het de juiste definitie van ziekte weergeeft als een abnormale fysieke of mentale toestand die de normale functies van het lichaam verstoort. B) Dit is een afleider, omdat het de tijdelijke toestand van ongemak benadrukt, maar geen volledige definitie van ziekte geeft. C) Dit is een afleider, omdat het de veroudering en slijtage van het lichaam benoemt, maar geen brede definitie van ziekte biedt.

>>

>>

c. De volgende meerkeuzevraag gaat over een experiment van Ivan Pavlov m.b.t. de klassieke conditionering dat hij in 1890 uitvoerde. Genereer 5 plausible afleiders op basis van de volgende stam en het correcte antwoordalternatief. Uit de afleiders moet blijken wat voor soort apparaat Pavlov gebruikte en hoe. Stam: Een experiment van Ivan Pavlov dat bekendstaat onder de Hond van Pavlov is een bekend voorbeeld van klassieke conditionering. Hoe mat Pavlov de speekselafscheiding? Correct antwoordalternatief: Pavlov plaatste een buisje in de mond van de hond en mat de hoeveelheid speekseldruppels via een opnameapparaat.	Hier zijn vijf plausible afleiders voor de meerkeuzevraag: 1. Pavlov gebruikte een stopwatch om de duur van de speekselafscheiding bij de hond te meten. 2. Pavlov gebruikte een weegschaal om het gewicht van de speekseldruppels te meten. 3. Pavlov gebruikte een temperatuurmeter om de temperatuurverandering in de mond van de hond tijdens de speekselafscheiding te meten. 4. Pavlov gebruikte een bloeddrukmeter om de bloeddruk van de hond tijdens de speekselafscheiding te meten. 5. Pavlov gebruikte een geluidsmeter om het geluid van de speekseldruppels te meten tijdens de speekselafscheiding. Let op: Deze afleiders zijn bedoeld om plausibel te klinken en kunnen de studenten in verwarring brengen als ze niet de juiste kennis hebben van het specifieke experiment van Pavlov. Het is belangrijk om ervoor te zorgen dat de studenten de informatie over Pavlov's experiment en de gebruikte apparatuur goed begrijpen om de juiste keuze te kunnen maken.
d. Maak een invuloefening voor studenten over C1 Engels. Deze invuloefening bevat een lijst met zinnen, elk met een lege ruimte waar een woord mist. De taak van de student is om deze lege ruimte te vullen met een correct woord van een vooraf gegeven lijst met opties. De zinnen moeten grammaticaal correct zijn. De invuloefening mag geen uitleg of aanvullende instructies bevatten, alleen de zinnen en de opties met woorden.	Zeker! Hier is een invuloefening voor studenten op C1-niveau Engels: Vul de lege ruimtes in de zinnen in met het juiste woord uit de opties hieronder: Opties: 1. although 2. furthermore 3. despite 4. whereas 5. nevertheless Zinnen: 1. I enjoy swimming, _____ I'm not very good at it. 2. Sarah is an excellent pianist. _____, she can also play the guitar and violin. 3. _____ the bad weather, they decided to go hiking in the mountains. 4. _____ studying late at night, she managed to get enough sleep and perform well in the exam. 5. He didn't win the race. _____, he broke the previous record.
e. Stel een discussievraag voor over inflatie die studenten uitnodigt om hun mening te geven en hun standpunten te onderbouwen.	Discussievraag: Wat zijn volgens jou de mogelijke gevolgen van langdurige inflatie op de economie en de samenleving, en hoe kunnen deze gevolgen worden aangepakt? Deze discussievraag nodigt studenten uit om hun mening te geven over de gevolgen van inflatie en om hun standpunten te onderbouwen. Ze kunnen verschillende aspecten van inflatie bespreken, zoals de impact op de koopkracht, de investeringen, de werkgelegenheid en de inkomensongelijkheid. Daarnaast kunnen ze mogelijke strategieën en beleidsmaatregelen verkennen om de negatieve gevolgen van inflatie te beperken of te voorkomen. Dit stimuleert studenten om kritisch na te denken over de complexiteit van het onderwerp en om hun argumenten te ondersteunen met relevante economische theorieën, voorbeelden of empirisch bewijs.

Tabel 2. Toepassingsmogelijkheid: Constructie van opdrachten (OpenAI, 2023)	
Prompt	Reactie ChatGPT
a. Genereer een korte interdisciplinaire opdracht waarbij hbo-studenten samen moeten werken om een probleem op te lossen dat verband houdt met het klimaat.	<i>Ten behoeve van dit artikel is de reactie door ChatGPT (261 woorden) samengevat in 50 woorden.</i> Interdisciplinaire opdracht: Klimaatbewuste Gemeenschap Werk in teamverband aan een plan om een lokale gemeenschap klimaatbewuster en veerkrachtiger te maken. Voer een milieuanalyse uit, onderzoek de sociale en economische impact, ontwikkel duurzame oplossingen, creëer een communicatiestrategie en maak een financieel plan. Presenteer het plan aan docenten en belanghebbenden, met nadruk op multidisciplinaire samenwerking en haalbaarheid.

Tabel 3. Toepassingsmogelijkheid: Constructie van beoordelingsinstrumenten (OpenAI, 2023; Ramin, 2023; Scott, z.d.)	
Prompt	Reactie ChatGPT
a. Ontwerp een vierpuntsrubric over het onderwerp 'NLQF niveau 4 probleemoplossende vaardigheden'.	<i>Eén criterium van de Rubric is hier opgenomen.</i> Hier is een vierpuntsrubric voor het beoordelen van probleemoplossende vaardigheden op NLQF niveau 4: Criterium Analyse van het probleem: 1 punt: De student vertoont beperkte vaardigheden bij het analyseren van het probleem en kan de belangrijkste elementen niet identificeren. 2 punten: De student toont enig begrip van het probleem, maar mist diepgang in de analyse en kan belangrijke elementen over het hoofd zien. 3 punten: De student voert een degelijke analyse uit van het probleem en identificeert de belangrijkste elementen en relaties. 4 punten: De student voert een grondige en systematische analyse uit van het probleem, waarbij alle relevante aspecten worden geïdentificeerd en begrepen.

Tabel 4. Toepassingsmogelijkheid: Constructie van toetsprocedures (OpenAI, 2023)	
Prompt	Reactie ChatGPT
a. Schrijf een gedetailleerde instructie aan de studenten over het gebruik van rekenmachines of andere technische hulpmiddelen tijdens de toets, in 100 woorden en een opsomming van maximaal drie punten.	Beste studenten, tijdens de toets is het gebruik van rekenmachines of andere technische hulpmiddelen toegestaan. Houd echter rekening met de volgende instructies: 1. Zorg ervoor dat de rekenmachine correct is ingesteld en goed werkt voordat je de toets begint. 2. Gebruik de rekenmachine alleen voor berekeningen en vermijd het opslaan van formules of andere niet-toegestane informatie. 3. Controleer je antwoorden met de rekenmachine, maar zorg ervoor dat je de tussenstappen handmatig uitvoert om je begrip van het probleem te tonen. Onthoud dat de toets voornamelijk gericht is op het begrijpen en toepassen van concepten, dus gebruik technische hulpmiddelen verstandig en als ondersteuning bij je berekeningen. Veel succes!

ChatGPT bij de uitvoering van toetsing

Hieronder worden vier toepassingsmogelijkheden bij de uitvoering van toetsen beschreven te weten tijdens het leren voor toetsen door

studenten (Tabel 5), genereren van producten (Tabel 6), formatieve toetsvormen in de les (Tabel 7) en het geven van feedback (Tabel 8).

Tabel 5. Toepassingsmogelijkheid: Leren voor toetsen door studenten (Scott, z.d.; Dr. Alex Young, 2023)	
Prompt	Reactie ChatGPT
a. Creëer een mindmap over het onderwerp inflatie met hoofdtakken en subtakken.	<pre> graph TD Inflatie --> Oorzaken Inflatie --> Gevolgen_en_effecten[Gevolgen en effecten] Oorzaken --> Monetair_beleid[Monetair beleid] Oorzaken --> Kosten Gevolgen_en_effecten --> Beleggers Gevolgen_en_effecten --> Consumenten_en_spaarders[Consumenten en spaarders] Monetair_beleid --> Geldhoeveelheid Kosten --> Prijsstijgingen_en_lonen[Prijsstijgingen en lonen] Beleggers --> Koopkrachtverlies </pre>

>>

>>

b. Wat zijn drie centrale concepten die ik moet weten als ik het onderwerp hypertensie bestudeer?	Als je het onderwerp hypertensie (hoge bloeddruk) bestudeert, zijn hier drie centrale concepten die je moet kennen: - Bloeddruk: Bloeddruk verwijst naar de kracht waarmee het bloed tegen de wanden van de bloedvaten drukt. Het wordt gemeten met twee waarden: de systolische druk (de bovendruk) en de diastolische druk (de onderdruk). Hoge bloeddruk treedt op wanneer de bloeddruk consistent hoger is dan normaal, wat kan leiden tot gezondheidsproblemen. - Risicofactoren: Verschillende factoren kunnen bijdragen aan het ontwikkelen van hypertensie. Deze omvatten levensstijlfactoren zoals slechte voeding, gebrek aan lichaamsbeweging, overgewicht, roken en overmatig alcoholgebruik. Andere risicofactoren zijn leeftijd, genetica, familiegeschiedenis, stress en bepaalde medische aandoeningen. - Gevolgen en complicaties: Ongecontroleerde hoge bloeddruk kan leiden tot verschillende ernstige complicaties, zoals hart- en vaatziekten, beroertes, nierproblemen, oogproblemen en schade aan bloedvaten. Het is belangrijk om hypertensie tijdig te diagnosticeren en te behandelen om deze complicaties te voorkomen of te verminderen. Door een goed begrip van deze centrale concepten kun je een stevige basis leggen voor verdere studie en begrip van hypertensie.						
c. Maak een twee-kolom spreadsheet met vragen en antwoorden over democratie.	Twee vragen en antwoorden <table><tr><th>Vraag</th><th>Antwoord</th></tr><tr><td>Wat is de definitie van democratie?</td><td>Democratie is een politiek systeem waarin de macht wordt uitgeoefend door het volk, meestal via verkiezingen, en waarin individuele vrijheden en mensenrechten worden gerespecteerd.</td></tr><tr><td>Welke zijn de belangrijkste kenmerken van democratie?</td><td>Belangrijke kenmerken van democratie zijn politieke participatie, vrije en eerlijke verkiezingen, respect voor de rechtsstaat, bescherming van individuele rechten en vrijheden.</td></tr></table>	Vraag	Antwoord	Wat is de definitie van democratie?	Democratie is een politiek systeem waarin de macht wordt uitgeoefend door het volk, meestal via verkiezingen, en waarin individuele vrijheden en mensenrechten worden gerespecteerd.	Welke zijn de belangrijkste kenmerken van democratie?	Belangrijke kenmerken van democratie zijn politieke participatie, vrije en eerlijke verkiezingen, respect voor de rechtsstaat, bescherming van individuele rechten en vrijheden.
Vraag	Antwoord						
Wat is de definitie van democratie?	Democratie is een politiek systeem waarin de macht wordt uitgeoefend door het volk, meestal via verkiezingen, en waarin individuele vrijheden en mensenrechten worden gerespecteerd.						
Welke zijn de belangrijkste kenmerken van democratie?	Belangrijke kenmerken van democratie zijn politieke participatie, vrije en eerlijke verkiezingen, respect voor de rechtsstaat, bescherming van individuele rechten en vrijheden.						
d. Prompt 1: Stel me een vraag over democratie. Je wacht op mijn antwoord. Nadat ik een antwoord heb gegeven, geef je aan of mijn antwoord juist of onjuist is en geef je daarbij een uitleg. ----- Prompt 2: Inspraak	Reactie op prompt 1: Natuurlijk, ik kan je een vraag stellen over democratie. Hier is de vraag: Wat is het belangrijkste principe van een democratie? ----- Reactie op prompt 2: Bedankt voor je antwoord! Je antwoord "inspraak" is juist. In een democratie is inspraak inderdaad een van de belangrijkste principes. Het houdt in dat alle burgers het recht hebben om deel te nemen aan het besluitvormingsproces van de overheid. Dit kan gebeuren door middel van verkiezingen, waarbij burgers hun stem kunnen uitbrengen om hun vertegenwoordigers te kiezen. Inspraak kan ook plaatsvinden via vrije meningsuiting, vrijheid van vergadering, het recht om petities in te dienen en andere mechanismen die burgerparticipatie bevorderen. In een democratie worden beslissingen genomen op basis van de wil van de meerderheid, maar het respecteren van de rechten van minderheden is ook een belangrijk aspect. Daarom zijn er vaak mechanismen zoals grondwetten en mensenrechtenverdragen die de rechten van individuen beschermen tegen de tirannie van de meerderheid. Inspraak speelt een cruciale rol bij het waarborgen van een eerlijke en inclusieve samenleving, waarin iedereen de mogelijkheid heeft om zijn of haar stem te laten horen.						

Tabel 6. Toepassingsmogelijkheid: Genereren van producten (OpenAI, 2023; Ramin, 2023; Dr. Alex Young, 2023)	
Prompt	Reactie ChatGPT
a. Schrijf drie probleemstellingen voor een afstudeeronderzoek waarin de volgende concepten centraal staan: decubitus, voeding, ouderen.	1. Wat is het verband tussen de voedingstoestand en het risico op decubitus bij oudere patiënten in verpleeghuizen? 2. Welke voedingsinterventies zijn effectief in het voorkomen van decubitus bij ouderen en het bevorderen van wondgenezing? 3. Hoe beïnvloedt de voedingstoestand van oudere patiënten het herstel en de prognose van decubitus?
b. Schrijf, in het Nederlands, een beleefde, niet opdringerige mondelinge verkooppitch waarom mensen het Tijdschrift Examens (www.e-xamens.nl) zouden moeten lezen die 30 seconden duurt.	"Goedemiddag! Bent u geïnteresseerd in toetsing en examens? Dan wil ik u graag wijzen op een waardevolle bron van kennis: het Tijdschrift Examens. Op www.e-xamens.nl vindt u diepgaande artikelen, interviews en praktische casestudies die uw professionele ontwikkeling kunnen verrijken. Of u nu een docent, toetsontwikkelaar of onderwijsprofessional bent, dit tijdschrift biedt waardevolle inzichten die u direct kunt toepassen. Neem gerust een kijkje en registreer u om toegang te krijgen tot deze waardevolle bron. Dank u wel voor uw tijd!"
c. Schrijf een voorbeeldessay van 50 woorden over ziekte. De doelgroep zijn hbo-studenten. Geef daarbij drie sterke punten en verbeterpunten van dit essay. Eerstejaars hbo-studenten gaan de kwaliteit van dit essay analyseren.	"Ziekte is een complex fenomeen dat zowel individuen als samenlevingen beïnvloedt. Het begrijpen van ziekte vereist kennis van biologische, psychologische en sociale factoren. Door de impact van ziekte te bestuderen, kunnen hbo-studenten een holistisch perspectief ontwikkelen en bijdragen aan gezondheidszorg en welzijn." Sterke punten: 1. Holistisch perspectief. 2. Relevantie voor de samenleving. 3. Beknopte en heldere presentatie. Verbeterpunten: 1. Diepgang en onderbouwing. 2. Structuur en samenhang. 3. Bronvermelding en academische referenties. <i>Ten behoeve van de lengte van dit artikel zijn alle argumenten van ChatGPT per punt achterwege gelaten.</i>
Noot. Er kunnen ook uitgebreidere producten gegenereerd worden zoals een essay, betoog, script, code, creatieve teksten of blogposts.	

Tabel 7. Toepassingsmogelijkheid: Formatieve toetsvormen in de les (Ramin, 2023)	
Prompt	Reactie ChatGPT
a. Ik wil een formatieve evaluatievorm hanteren waarmee ik kan nagaan of het leerdoel van mijn les is behaald. Welke vorm kan ik hiervoor gebruiken en welke instructie kan ik de studenten concreet geven bij deze vorm?	Een effectieve vorm van formatieve evaluatie om de behaalde leerdoelen te beoordelen, is het gebruik van exit tickets. Geef de studenten de volgende instructie: "Schrijf op een klein kaartje of papiertje één belangrijk concept, feit of vaardigheid die je vandaag hebt geleerd en waarvan je denkt dat het aansluit bij de leerdoelen van deze les. Leg beknopt uit waarom dit belangrijk is. Lever het kaartje aan het einde van de les in." Deze exit tickets geven jou inzicht in de begrip en het behalen van de leerdoelen, zodat je gerichte feedback kunt geven en eventuele lacunes kunt identificeren.

Tabel 8. Toepassingsmogelijkheid: Het geven van feedback
(D.E.E.P Teaching, 2023; Scott, z.d.)

Prompt	Reactie ChatGPT
a. De student heeft zelf een onderzoeksverslag geschreven. In het verslag zit een duidelijk betooglijn. Het verslag bevat weinig onderbouwing op basis van wetenschappelijke literatuur. Ook wordt de onderzoeksvraag niet volledig beantwoord. Kan je de hierbovenstaande feedback aan de student kort en bondig weer-geven gericht aan de student. Wil je starten met de sterke punten en daarna de verbeterpunten benoemen. Wil je advies geven aan de student hoe hij zich kan verbeteren	Beste student, Sterke punten: • Je onderzoeksverslag heeft een duidelijke betooglijn. Verbeterpunten: • Er is weinig onderbouwing met wetenschappelijke literatuur. • De onderzoeksvraag wordt niet volledig beantwoord. Advies voor verbetering: • Versterk je verslag door meer gebruik te maken van wetenschappelijke literatuur om je argumenten te ondersteunen en je verslag geloofwaardiger te maken. • Breid je analyse en conclusies uit om de onderzoeksvraag volledig te beantwoorden en een completer beeld te geven. • Vraag feedback aan docenten, experts of medestudenten om je verslag verder te verbeteren. Door deze adviezen op te volgen, kun je je verslag versterken door een stevigere onderbouwing en een volledige beantwoording van de onderzoeksvraag. Maak gebruik van wetenschappelijke literatuur en zoek externe input om je verslag te verbeteren. Veel succes met je verdere onderzoeksproces! Met vriendelijke groet, [Je naam]
b. Het redactioneel van het vorige nummer van Examens is voorgelegd aan ChatGPT. Beoordeel de bovenstaande tekst op leesbaarheid, bondigheid en inhoudelijke juistheid. Geef de feedback in 100 woorden weer.	De tekst is goed leesbaar en trekt de aandacht van de lezer door de introductie van het Examens-tijdschrift en de nadruk op examinering. Het biedt een overzicht van de inhoud en bijdragen van verschillende auteurs. Hoewel het informatief is, kan de tekst iets bondiger zijn door minder herhaling van zinsconstructies. Inhoudelijk lijkt de tekst juist en bevat het relevante informatie over de onderwerpen en bijdragen in het tijdschrift. Over het geheel genomen is het een goede introductie voor de lezer om inzicht te krijgen in het nummer en aanverwante onderwerpen. <i>Alternatief: Heb je een beoordelingsinstrument (bijvoorbeeld een rubric) dan is het mogelijk om ChatGPT de opdracht te geven om de tekst te beoordelen aan de hand van deze rubric en feedback te vragen per onderdeel. In D.E.E.P Teaching (2023) wordt aan studenten gevraagd om voordat ze een verslag inleveren eerst feedback te vragen aan ChatGPT.</i>

>>

>> c. Vier regels JAVA-code zijn voorgelegd aan ChatGPT. <pre>public static void main(String args[]) { int x = 5; int y = 10; int sum = x + y; System.out.println("De som is: " + sum); }</pre> Geef feedback op deze code	De code bevat een kleine fout die moet worden gecorrigeerd. De juiste methode om tekst naar de console te printen is System.out.println() in plaats van System.ot.println(). Hier is de aangepaste code: <pre>java Copy code public static void main(String args[]) { int x = 5; int y = 10; int sum = x + y; System.out.println("De som is: " + sum); }</pre> Feedback voor de student: • Variabelen: Goed gedaan met het definiëren van de variabelen x en y om respectievelijk de waarden 5 en 10 op te slaan. Dit is een juiste manier om waarden toe te wijzen aan variabelen. • Berekening: Correct gebruik van de operator '+' om de waarden van x en y op te tellen en de som op te slaan in de variabele sum. • Uitvoer: Er is echter een typfout in de code bij System.out.println(). Het moet System.out.println() zijn om de som correct naar de console af te drukken. Zorg ervoor dat je de code controleert op typfouten en syntactische correctheid voordat je deze uitvoert.
---	---

ChatGPT bij de evaluatie van toetsing

Hieronder worden drie toepassingsmogelijkheden bij de evaluatie van toetsen beschreven te weten ondersteuning bij: de evaluatie van

toetsitems (Tabel 9); het proces van valideren (Tabel 10); en het opstellen van evaluatieformulieren (Tabel 11).

Tabel 9. Toepassingsmogelijkheid: Evaluatie van toetsitems	
Prompt	Reactie ChatGPT
a. In de toets-itemanalyse heeft een toetsvraag een lage p-waarde. Wat kan hiervan de oorzaak zijn?	Natuurlijk! Hier zijn twee mogelijke oorzaken voor een lage p-waarde in de toets-itemanalyse: • Moeilijkheidsgraad: De vraag kan te moeilijk zijn voor de studenten. Dit betekent dat een groot aantal studenten de vraag niet correct beantwoordt, wat resulteert in een lage p-waarde. Mogelijke redenen voor de moeilijkheid kunnen zijn dat de vraag te complex is gesteld, dat er onvoldoende context of aanwijzingen worden gegeven, of dat er een gebrek is aan voorkennis die nodig is om de vraag te beantwoorden. • Plafondeffect: Als de vraag te gemakkelijk is en vrijwel alle studenten deze correct beantwoorden, kan dit leiden tot een lage p-waarde. In zo'n geval kan de vraag geen onderscheid maken tussen studenten met verschillende niveaus van kennis of vaardigheden, omdat vrijwel iedereen de vraag correct beantwoordt. Het identificeren van de specifieke oorzaken van een lage p-waarde kan helpen bij het verbeteren van de vraag. Bij een te moeilijke vraag kan overwogen worden om de formulering te verduidelijken, context toe te voegen of de vereiste voorkennis beter aan te geven. Bij een te gemakkelijke vraag kan de vraag herzien worden om meer uitdagend te zijn of vervangen worden door een andere vraag die beter onderscheid maakt tussen studenten met verschillend niveau van bekwaamheid.

>>

>>

b. Evalueer de kwaliteit van deze toetsvraag: Hoe wordt het oppervlak van een raam gemeten? A. De dagmaat van het kozijn. B. De maten van het glasoppervlak plus de glaslatten. C. De nettomaten van het glasoppervlak (Molkenboer, 2015; p. 197). Leg uit wat er niet goed is aan deze toetsvraag en hoe ik deze het beste kan verbeteren.	Deze toetsvraag bevat een probleem met betrekking tot de kwaliteit. Het beste zou zijn om duidelijk te specificeren wat er precies wordt bedoeld met “het oppervlak van een raam” om mogelijke verwarring te voorkomen. Daarnaast zijn de antwoordopties in dit geval niet goed gestructureerd, omdat ze verschillende aspecten van het raamoppervlak behandelen. Om de vraag te verbeteren, zou je de vraag kunnen herformuleren en de antwoordopties consistent maken. Een voorbeeld van een verbeterde vraag zou kunnen zijn: “Hoe wordt het totale oppervlak van een raam gemeten?” A. De totale buitenafmetingen van het raamkozijn. B. De totale afmetingen van het glasoppervlak inclusief de glaslatten. C. De totale binnenafmetingen van het raamkozijn. <i>Alternatief: Je zou ook meerdere toetsvragen (mits toegestaan) aan ChatGPT kunnen voorleggen. Zo heeft V. Kalis, persoonlijke communicatie, 19 juni 2023 de volgende prompt uitgeprobeerd: Uit een itemanalyse blijkt dat sommige afleiders van meerkeuzevragen nauwelijks zijn gekozen door tentamenkandidaten. Ik geef je vijf meerkeuzevragen van het tentamen en jij legt me uit wat er mogelijk mis is met de formulering van de afleiders.</i>
--	--

Tabel 10. Toepassingsmogelijkheid: Proces van valideren	
Prompt	Reactie ChatGPT
a. Ik ben bewijzen aan het verzamelen om de mate van validiteit van een performance assessment aan te tonen. Ik heb al een toetsmatrijs gemaakt. Welke drie aanvullende validiteitsbewijzen zou ik nog kunnen gebruiken?	Als aanvulling op de toetsmatrijs zijn hier drie aanvullende validiteitsbewijzen die je kunt gebruiken om de mate van validiteit van een performance assessment verder aan te tonen: - Concurrente validiteit: Vergelijk de resultaten van het performance assessment met die van een al bestaand assessment of een andere geaccepteerde maatstaf die hetzelfde construct meet. Als de scores op beide metingen sterk met elkaar correleren, duidt dit op concurrente validiteit. - Predictieve validiteit: Onderzoek of de scores op het performance assessment voorspellend zijn voor toekomstige prestaties of succes in het desbetreffende domein. Als er een aantoonbaar verband is tussen de scores op het assessment en de latere prestaties, toont dit de predictieve validiteit aan. - Constructvaliditeit: Verzamel bewijs dat aantoont dat het performance assessment het beoogde construct daadwerkelijk meet. Dit kan worden gedaan door middel van literatuuronderzoek, raadplegen van experts en het uitvoeren van cognitieve interviews met deelnemers om na te gaan of het assessment de juiste vaardigheden en kennis meet die gerelateerd zijn aan het construct. Door het verzamelen van gegevens en bewijs uit deze aanvullende validiteitsbewijzen naast de toetsmatrijs, kun je een sterker geval maken voor de validiteit van het performance assessment. Het is belangrijk om verschillende vormen van validiteitsbewijzen te gebruiken om een uitgebreide en overtuigende validiteitsargumentatie op te bouwen.

Tabel 11. Toepassingsmogelijkheid: Opstellen van evaluatieformulieren										
Prompt		Reactie ChatGPT								
a. Formuleer drie gesloten vragen, genummerd van 1 tot 3, die ik aan mbo-studenten kan stellen om hun tevredenheid over toetsing te meten. Plaats de vragen in een matrix met de antwoordmogelijkheden helemaal mee oneens tot helemaal mee eens.		Hier is de gewenste matrix met drie gesloten vragen (genummerd van 1 tot 3) om de tevredenheid van mbo-studenten over toetsing te meten, inclusief de antwoordmogelijkheden “helemaal mee oneens” tot “helemaal mee eens”:								
		<table><tr><td>Helemaal mee oneens</td><td>Mee oneens</td><td>Neutraal</td><td>Mee eens</td><td>Helemaal mee eens</td></tr></table>				Helemaal mee oneens	Mee oneens	Neutraal	Mee eens	Helemaal mee eens
		Helemaal mee oneens	Mee oneens	Neutraal	Mee eens	Helemaal mee eens				
		1. De toetsinstructies waren duidelijk en begrijpelijk.								
		2. Ik voelde me eerlijk beoordeeld na het maken van de toets.								
3. De toetsvragen waren representatief voor de leerstof.										
		Studenten kunnen het bijbehorende nummer invullen in de kolom die het beste hun tevredenheid weergeeft voor elke vraag.								
b. Schrijf een enthousiasmerende oproep van 50 woorden, gericht aan studenten, waarom het belangrijk is dat zij de vragenlijst invullen over hun tevredenheid over toetsing.		Beste studenten, jouw mening telt! Maak een verschil en laat je stem horen over de toetsing. Jouw feedback is essentieel voor het verbeteren van het onderwijsproces en het optimaliseren van toetsmethoden. Jouw tevredenheid is van groot belang. Vul de vragenlijst in en help ons samen de toetservaring te verbeteren!								

Aandachtspunten bij het gebruik van ChatGPT bij toetsen

De bovenstaande tabellen geven een aantal toepassingsmogelijkheden hoe ChatGPT bij de constructie, uitvoering en evaluatie van toetsing ingezet wordt of kan gaan worden. Het gebruik van de resultaten van deze prompts bij toetsing heeft implicaties voor de dagelijkse onderwijspraktijk. Enkele aandachtspunten voor een weloverwogen gebruik van ChatGPT zijn:

- 1. Evalueer de kwaliteit.** Hoewel het een sterk punt van ChatGPT is dat het antwoorden genereert en dat het zichzelf verbetert (Farrokhnia et al., 2023), kan het ook onjuiste of onnauwkeurige informatie genereren (OpenAI, 2023). Zoals ook blijkt uit enkele van de hierboven weergegeven voorbeelden. ChatGPT mist een diep begrip van de informatie en heeft moeilijkheden met het evalueren van de kwaliteit van de reacties (Farrokhnia et al., 2023). Ook maakt het gebruik van een beperkte verzameling aan websites waarvan niet altijd de kwaliteit bij voorbaat gewaarborgd is (Hofman & Veerbeek, 2023). Daarbij schrijft ChatGPT goed, maar niet foutloos, Nederlands (Salemans, 2023). Ook dit is

terug te zien in de voorbeelden. Beoordeel en controleer dus kritisch de gegenereerde reacties en verifieer deze op basis van andere bronnen en expertises (OpenAI, 2023).

- 2. Kijk naar de toepassing van ChatGPT in de context.** Ga kritisch na in welke context ChatGPT wordt gebruikt, wat de (ethische) implicaties hiervan zijn en of dit wenselijk is. Zo is er een risico van discriminatie door ChatGPT (Farrokhnia et al., 2023). Van Rooij (2022) geeft aan, op basis van diverse bronnen, dat studenten geleerd wordt om niet te plagiëren en goed, onbevooroordeeld literatuuronderzoek te doen en literatuurverwijzingen te hanteren. Het gebruik van ChatGPT kan bestaande biases versterken in plaats van ze uit te dagen, zoals verwacht mag worden van wetenschappers.
- 3. Formuleer je prompt duidelijk en concreet.** Het formuleren van duidelijke en goed gestructureerde prompts is essentieel om nauwkeurige en relevante antwoorden van ChatGPT te verkrijgen. Je kan dit volgens OpenAI (2023) en Scott (z.d.) doen door onder andere:
 - De prompt duidelijk en beknopt te

formuleren, zodat de vraag helder is.

- In de prompt het gewenste format te vragen van het antwoord, zodat ChatGPT het antwoord alvast kan opmaken in het gewenste format. Denk hierbij aan een lijst, tabel of een mindmap.
- De toon te beschrijven. Denk hierbij aan vriendelijk, grappig of professioneel.
- De prompt specifiek te formuleren en hierbij context te geven, zodat ChatGPT een beter begrip heeft van de gestelde vraag.
- Waar nodig achtergrondinformatie te geven zodat het antwoord meer diepgang krijgt.
- ChatGPT waar nodig bij te sturen. Dit kan door bijvoorbeeld te vragen: “Kun je me een voorbeeld geven?”.



**ChatGPT maakt
mensachtige interacties
mogelijk door te reageren
op prompts.**

- De vraag te herhalen aan het einde van de prompt. Dit kan nuttig zijn, omdat het model de neiging heeft om de meest recente informatie in de prompt te prioriteren bij het genereren van een respons. Hierdoor kan de vraag beter begrepen worden.

4. **Kies kritisch en bewust welke informatie je wel of niet deelt.** Wees je ervan bewust dat door ChatGPT te gebruiken je informatie deelt van jezelf, de onderwijsinstelling en eventueel van studenten met ChatGPT. Hieraan zitten veiligheids- en privacyrisico's verbonden (Ray, 2023). Zoals OpenAI (2023) zelf aangeeft wordt de informatie die je invoert gebruikt om een respons te genereren. Ook kan het gebruikt worden door de makers om de service te verbeteren. Ga hier dus zorgvuldig mee om door te voldoen aan wetten en regelgeving met betrekking tot het auteursrecht, vertrouwelijkheid en gegevensbescherming (OpenAI, 2023). Lees de actuele privacy policy van OpenAI om op de hoogte te zijn en te blijven hoe met de data wordt omgegaan.
5. **Inventariseer en ken de huidige en toekomstige kansen en risico's van AI-tools.** ChatGPT biedt kansen maar brengt ook risico's met zich mee bij toetsen. Het is van belang dat deze kansen worden benut, dat de risico's nu en in de toekomst worden geanalyseerd. Een aantal kansen voor het onderwijs die worden genoemd door Farrokhnia et al. (2023) zijn onder andere: eenvoudige toegang tot informatie, het faciliteren van gepersonaliseerd leren en het verminderen van de werkdruk van docenten. Genoemde risico's voor het onderwijs zijn onder andere: een verminderd begrip van de context, een bedreiging voor de academische integriteit en de democratisering van plagiaat (Cotton et al., 2023; Farrokhnia et al., 2023; Van Rooij, 2023). Op basis van de analyse van kansen en risico's kunnen richtlijnen

worden opgesteld hoe om te gaan met AI-tools in het onderwijs (Cotton et al., 2023). Van onderwijsprofessionals vraagt dit AI-geletterdheid. Dit betekent een bewustzijn voor onderwijsprofessionals van de alomtegenwoordigheid van AI, hun vermogen om het te gebruiken en om kritisch te denken over AI (Anders, 2023). Nu en in de toekomst, ook gezien het feit dat er steeds meer op ChatGPT gebaseerde tools en plugins gelanceerd worden.

6. **Wees transparant over het gebruik van ChatGPT.** Heb je zelf gebruikgemaakt van ChatGPT zorg dan voor een literatuurverwijzing naar de gebruikte tool (McAdoo, 2023; Scott, z.d.). Het is belangrijk om erbij te vermelden hoe ChatGPT is gebruikt (de prompt) en wat de reactie is, omdat de resultaten niet reproduceerbaar zijn voor andere gebruikers (McAdoo, 2023). Dus dezelfde prompt genereert op verschillende momenten verschillende reacties (Bosshardt, 2023a). De literatuurverwijzing refereert hiermee aan het gebruik van de tool, maar het maakt niet duidelijk waar de resultaten vandaan komen die gegenereerd worden door ChatGPT. Van Rooij (2022) legt dit uit als een vorm van geautomatiseerd plagiëren, dat schade kan veroorzaken aan jezelf en anderen.
7. **Gebruik het als hulpmiddel.** ChatGPT kan worden gebruikt als een hulpmiddel bij toetsgerelateerde activiteiten, maar het kan (uiteraard) geen vervanging zijn voor de toetsexpertise van docenten en toetsexperts (OpenAI, 2023). Ook bij het gebruik van ChatGPT door studenten mag het niet zo zijn dat het voor docenten moeilijk of zelfs onmogelijk wordt om de daadwerkelijke mate van begrip van de student van de leerstof adequaat te beoordelen (Cotton et al., 2023). AI bij het ontwerpen van toetsing is het van belang om hiermee rekening te houden (Cotton et al., 2023). Voor studenten kan ChatGPT echter wel een effectief hulpmiddel zijn

bij onder andere brainstormen en het opdoen van nieuwe kennis (Bosshardt, 2023b).

Tot slot

Dit artikel illustreert diverse toepassingsmogelijkheden van ChatGPT bij toetsing en geeft enkele aandachtspunten die relevant zijn voor een weloverwogen gebruik daarvan. Beoordeel zelf of, in welke mate en onder welke voorwaarden deze voorbeelden daadwerkelijk toepasbaar zijn in de dagelijkse onderwijspraktijk.

Speciale dank gaat uit naar Vincent Kalis (Tilburg University), Annemieke Bosshardt (Saxion) en Engeline van Lokven voor hun feedback op dit artikel. ■

Referenties

- Anders, B. A. (2023). Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking? *Patterns*. 4(3), 1-2. <https://doi.org/10.1016/j.patter.2023.100694>
- Bosshardt, A. (2023a). Handreiking ChatGPT bij Saxion. Saxion.
- Bosshardt, A. (2023b). Handreiking AI/ChatGPT voor studenten. Saxion
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. 1-12. <https://doi.org/10.1080/14703297.2023.2190148>
- D.E.E.P Teaching. (2023, 30 januari). Marking and feedback using ChatGPT [Video]. YouTube. Geraadpleegd op 2 juni 2023, van <https://youtu.be/FNw-rsPI-yo>
- Dr. Alex Young. (2023, 5 februari). Learn Anything FASTER with ChatGPT (13 ChatGPT Prompts For Studying) [Video]. YouTube. Geraadpleegd op 7 juni 2023, van <https://youtu.be/TEq5chOb4is>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: implications for educational practice and research. *Innovations*

- in Education and Teaching International. 1-15. <https://doi.org/10.1080/14703297.2023.2195846>
- Hofman, E., & Veerbeek, J. (2023, 7 juni). Onderzoek. De bronnen van ChatGPT. 'Dat zijn toch gewoon ál onze artikelen?'. De Groene Amsterdammer. Geraadpleegd op 12 juni 2023, van <https://www.groene.nl/artikel/dat-zijn-toch-gewoon-al-onze-artikelen>
 - McAdoo, T. (2023, 7 april). How to cite ChatGPT. Geraadpleegd op 13 juni 2023 van <https://apastyle.apa.org/blog/how-to-cite-chatgpt>
 - Molkenboer, H. (2015). Toetsen volgens de toetscyclus. Bureau voor toetsen en beoordelen.
 - OpenAI. (2023). ChatGPT-3.5 (May 24 version) [Large language model]. <https://chat.openai.com/chat>
 - Patel, N. (2023, 22 mei). Using ChatGPT prompts to generate assessment questions (with Examples). Geraadpleegd op 7 juni 2023, van <https://www.playpowerlabs.com/post/using-chatgpt-prompts-to-generate-effective-formative-assessment-questions>
 - Ramin, S. (2023, 1 maart). 50 time-saving ChatGPT prompts for teachers. Geraadpleegd op 6 juni 2023, van <https://www.helloteacherlady.com/blog/2023/3/50-time-saving-chatgpt-prompts-for-teachers>
 - Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems. 3. 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
 - Salemans, B. (2023, 7 januari). ChatGPT: de rapportcijfers. Neerlandistiek. Online tijdschrift voor taal- en letterkunde. Geraadpleegd op 4 juli 2023, van <https://neerlandistiek.nl/2023/01/chatgpt-de-rapportcijfers/>
 - Scott, M. (z.d.). Generative AI is a game-changer for education & business. Geraadpleegd op 12 juni 2023, van <https://scottybreaksitdown.com/ai/>
 - Van Rooij, I. (2022, 29 december). Against automated plagiarism. Geraadpleegd op 15 juni 2023, van <https://irisvanrooijcogsci.com/2022/12/29/against-automated-plagiarism/>
 - Van Rooij, I. (2023, 14 januari). Stop feeding the hype and start resisting. Geraadpleegd op 12 juni 2023, van <https://irisvanrooijcogsci.com/2023/01/14/stop-feeding-the-hype-and-start-resisting/>

Dit boekje bevat een verzameling annotaties en artikelen over onderwijsrechtelijke thema's.

De afzonderlijke bijdragen zijn alle reeds eerder gepubliceerd in 'Examens, tijdschrift voor de toetspraktijk'. Deze verzameling is bij elkaar gebracht omdat ze een overzicht geeft van de rechtsontwikkeling met betrekking tot onderwijsrechtelijke onderwerpen zoals onder andere het bindend afwijzend studieadvies, tentamenfraude, de waarde van buitenlandse diploma's, recht op inzage van gemaakt tentamenwerk, bijzondere persoonlijke omstandigheden, toelating tot opleidingen en de wijze van toetsing in de onderwijsrechtspraak.

Ton Lamers publiceert regelmatig over handelsrechtelijke-, arbeidsrechtelijke en onderwijsrechtelijke thema's.

17.⁹⁵

Terecht of Niet?

Onderwijsrechtelijke publicaties en annotaties in 'Examens'



mr. dr. A.H. Lamers

Bestel nu bij: www.bol.com