

Monique Bil-Meijer
De Haagse Hogeschool
Contact: a.n.meijer@hhs.nl

De impact van gepersonaliseerde AI-training met ChatGPT op de ontwikkeling van mondelinge gespreksvaardigheden bij hbo-rechtenstudenten. Een gerandomiseerd onderzoek

1. Inleiding

Goede mondelinge communicatie is erg belangrijk in het juridisch werkveld. Adviseren, het toelichten van juridische beslissingen en het terugkoppelen aan cliënten vereisen vaardigheden die, ondanks technologische vooruitgang, nog steeds een menselijk specialisme zijn. Binnen hbo-rechtenopleidingen wordt de ontwikkeling van gespreksvaardigheden vaak vormgegeven via rollenspellen met medestudenten onder begeleiding van docenten. Deze aanpak biedt een veilige leeromgeving (Chernikova e.a. 2020), maar is arbeidsintensief en afhankelijk van de beschikbaarheid van trainers en acteurs. Hierdoor varieert de kwaliteit en consistentie van de feedback.

Generatieve AI (GenAI) en *Large Language Models* (LLM's), zoals ChatGPT, bieden nieuwe kansen. Molenaar (2022) beschrijft het concept van 'hybride *human-AI learning*', waarbij AI als adaptieve leerpartner optreedt en gepersonaliseerde ondersteuning biedt, terwijl de docent de regie behoudt. Dit sluit aan bij de behoefte om vaardigheidstraining schaalbaarder te maken, zonder kwaliteitsverlies.

Binnen het juridisch domein spelen niet alleen gesprekstechniek, maar ook argumentatie een grote rol. Prakken (2021) benadrukt dat AI kan helpen bij het structureren en evalueren van juridische redeneringen, waardoor studenten hun communicatieve en inhoudelijke vaardigheden tegelijk kunnen ontwikkelen. In medische opleidingen zijn al positieve ervaringen opgedaan met AI-gestuurde simulaties, zoals virtuele patiënten, die klinische communicatie trainen (Bowers e.a. 2024). Het gebruik van ChatGPT in niet-medische opleidingen, zoals het juridisch domein, is echter nog maar nauwelijks onderzocht.

Deze studie vergelijkt de effectiviteit van een AI-gebaseerde gespreksvaardigheidstraining met een traditioneel rollenspel in het hbo-rechtenonderwijs.

2. Methode

Het onderzoek is uitgevoerd als een gerandomiseerde gecontroleerde trial (RCT) binnen de opleiding Recht aan De Haagse Hogeschool. De interventie vond plaats in maart 2025 in het tweede studiejaar, tijdens de module 'De Bedrijfsorganisatie'. Studenten oefenen in deze module met professionele communicatie. In dit geval ging het om een slechtnieuwsgesprek over een privacyovertreding.

3. Deelnemers

In totaal namen 156 studenten deel aan het onderzoek. Ze werden willekeurig toegewezen aan de AI-conditie (n = 67) of de controlegroep (n = 89). In de AI-conditie gaven vijftien studenten geen toestemming voor het gebruik van hun gegevens, waardoor 52 studenten overbleven voor de analyse van de vragenlijsten. Van deze groep namen 36 studenten deel aan het summatieve assessment. In de controlegroep namen 57 studenten (64%) deel aan het summatieve assessment. Voor de analyse van de toetsresultaten zijn uitsluitend deze deelnemers meegenomen, wat resulteerde in een analysegroep van 93 studenten (36 in de AI-conditie en 57 in de controlegroep; zie: Tabel 1).

Conditie	Studenten bij start	Geen toestemming	Over voor vragenlijst	Deelnemers assessment	Gebruikt in analyse toetsresultaten
AI-conditie	67	15	52	36	36
Controlegroep	89	-	-	57	57
Totaal	156	15	-	93	93

Tabel 1 – Overzicht deelname en analysepopulatie.

4. Interventies

In de AI-conditie oefenden studenten gespreksvaardigheden met een AI-partner, gebaseerd op ChatGPT, een LLM. De controlegroep volgde reguliere rollenspeltrainingen met medestudenten, onder begeleiding van een docent.

4.1 AI-conditie

Studenten in de AI-conditie kregen gedurende tien weken twee lessen van drie uur, waarin zij diverse gesprekssimulaties van 20-30 minuten uitvoerden. Deze simulaties waren thematisch afgestemd op beroepssituaties uit het juridisch werkveld (slechtnieuwsgesprek). De AI-partner was geprogrammeerd om gespreksstructuur, empathisch responsgedrag en reflectieve prompts te genereren, en gaf direct na afloop automatisch gegenereerde feedback op basis van de indicatoren.

4.2 Controlegroep

Studenten in de controlegroep oefenden dezelfde scenario's in rollenspellen met medestudenten, onder supervisie van een docent. Feedback werd mondeling gegeven op basis van observatie en reflectie met een medestudent. De tijdsinvestering en thematische inhoud waren gematcht aan die van de AI-conditie.

5. Meetinstrumenten

De effectmeting vond plaats aan de hand van een gestandaardiseerde beoordelingsrubric (1-3 Likertschaal) op drie domeinen:

1. *gespreksstructuur*: mate van doelgerichtheid, logische opbouw argumenten en nette afsluiting.
2. *empathisch vermogen*: luistergedrag, erkenning van gevoelens en verbale responsiviteit.
3. *taakgerichte reflectie*: vermogen tot zelfevaluatie en formuleren van ontwikkelpunten.

Beoordelingen werden uitgevoerd door twee onafhankelijke beoordelaars. Daarnaast vulden studenten na afloop een vragenlijst in over motivatie, spanning/stress en zelfvertrouwen, evenals rond enkele achtergrondvariabelen (AI-kennis en algemene motivatie).

6. Data-analyse

Studenten zonder toetsresultaat zijn buiten beschouwing gelaten. Toetsscores zijn vergeleken met een onafhankelijke t-toets (Welch-correctie). Effectgrootte werd berekend met Cohen's d.

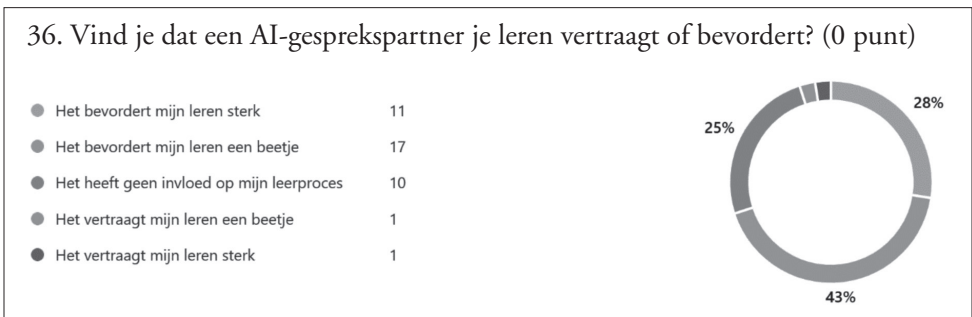
7. Resultaten

De beoordelingen (schaal 1-10) laten zien dat de controlegroep gemiddeld $M = 5,93$ ($SD = 1,83$; bereik 1-9) scoorde, en de experimentele groep $M = 5,69$ ($SD = 1,85$; bereik 1-8). Een onafhankelijke t-toets met Welch-correctie toonde geen significant verschil tussen de groepen, $t(\approx 66) = 0,60$, $p = .55$, $d = 0,13$. De verdelingen waren vergelijkbaar, met een lichte oververtegenwoordiging van hogere scores in de controlegroep.

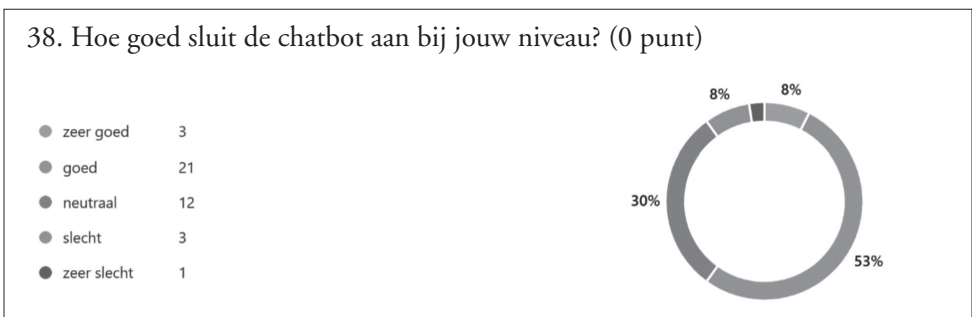
Groep	<i>n</i>	M	SD	Bereik
Controlegroep	57	5,93	1,83	1-9
Experimentele groep	36	5,69	1,85	1-8



Figuur 1 – Motivatie voor het oefenen met een AI-partner.



Figuur 2 – Ervaren spanning en stress tijdens het oefenen met een AI-partner.



Figuur 3 – Zelfvertrouwen tijdens het oefenen.

8. Aanvullende context

Naast de vragen over motivatie, spanning en zelfvertrouwen zijn ook achtergrondvariabelen meegenomen die inzicht geven in het startprofiel van de deelnemers. Meer dan de helft van de studenten (57,1%) gaf aan over gemiddelde kennis te beschikken van *AI-tools*. 28,6% had basiskennis en 14,3% gevorderde kennis.

Wat betreft algemene motivatie om deel te nemen aan onderwijsactiviteiten, gaf de helft van de studenten (50%) aan goed gemotiveerd te zijn. Een derde (33,3%) voelde zich niet gemotiveerd en 16,7% was neutraal. Deze contextvariabelen zijn relevant voor de interpretatie van de bevindingen, aangezien zowel de bekendheid met AI als de algemene motivatie mogelijk van invloed zijn op de waardering van en de reacties op AI-ondersteund oefenen.

9. Conclusie

De studie vergeleek een AI-gebaseerde gespreksvaardigheidstraining met traditionele rollenspeltrainingen met medestudenten binnen een hbo-rechtenopleiding. Er werd geen significant verschil gevonden in de behaalde toetsscores tussen beide groepen, $t(\approx 66) = 0,60$, $p = .55$, $d = 0,13$. Deze bevinding suggereert dat AI-ondersteunde simulaties in deze context een haalbaar alternatief kunnen vormen voor conventionele vaardigheidstraining, zonder dat dit gepaard gaat met lagere eindresultaten.

De resultaten uit de vragenlijst suggereren dat AI-ondersteund oefenen in gespreksvaardigheidstraining door studenten overwegend positief wordt ontvangen. De activiteit werd doorgaans als plezierig en zinvol ervaren, met een hoog gerapporteerd zelfvertrouwen en weinig ervaren spanning. De beperkte bereidheid tot gebruik buiten de onderwijscontext wijst echter op voornamelijk extrinsieke motivatie.

10. Discussie

De bevindingen moeten worden geïnterpreteerd in het licht van de achtergrondkenmerken van de deelnemers. Meer dan de helft beschikte over gemiddelde kennis van *AI-tools*, wat kan hebben bijgedragen aan het vlot kunnen gebruiken van de AI-partner. Studenten met beperkte ervaring met AI zouden mogelijk meer tijd nodig hebben gehad om vertrouwd te raken met de technologie, wat hun waardering voor de activiteit had kunnen beïnvloeden.

Daarnaast gaf een aanzienlijk deel van de studenten aan niet of slechts matig gemotiveerd te zijn om deel te nemen aan onderwijsactiviteiten. Deze algemene motivatie kan van invloed zijn geweest op hun houding tegenover de AI-oefening. Het relatief

hoge aandeel neutrale antwoorden bij sommige motivatie- en ontspanningsitems zou hiermee samen kunnen hangen. Vervolgonderzoek zou kunnen nagaan in hoeverre AI-ondersteunde trainingen juist voor minder gemotiveerde studenten een ingang bieden om actief te oefenen, of dat motivatie eerst via andere interventies versterkt moet worden.

Referenties

- Bowers, P., K. Graydon, T. Ryan, J.H. Lau & D. Tomlin (2024). "Artificial intelligence-driven virtual patients for communication skill development in healthcare students: A scoping review". In: *Australasian Journal of Educational Technology*, 40 (3), p. 39-57.
- Chernikova, O., N. Heitzmann & M. Stadler (2020). "Simulation-based learning in higher education: A meta-analysis". In: *Review of Educational Research*, 90 (4), p. 499-541.
- Molenaar, I. (2022). "Towards hybrid human AI learning technologies". In: *European Journal of Education*, 57 (4), p. 632-645.
- Prakken, H. (2021). 'Kunnen computers juridisch argumenteren?'. Oratie Universiteit Utrecht. Online raadpleegbaar op: <https://webspace.science.uu.nl/~prakk101/pubs/oratieHPdef.pdf>.
-