

Lieve De Wachter & Lotte Uyttenbroeck

ILT KU Leuven

Contact: [lieve.dewachter@kuleuven.be](mailto:lieve.dewachter@kuleuven.be)

[Lotte.uyttenbroeck@student.kuleuven.be](mailto:Lotte.uyttenbroeck@student.kuleuven.be)

## Hoe gebruiken studenten Generatieve AI bij schrijfp opdrachten in het academisch hoger onderwijs? Een onderzoek door middel van *think aloud protocols* naar de inzet van ChatGPT in het schrijfproces

4

### 1. Inleiding

De komst van ChatGPT in november 2022 heeft in het hoger onderwijs zowel student als docent geconfronteerd met een reeks vragen. Regelmatig duiken persberichten op over dalende schrijfkwaliteit bij studenten en de ‘dood’ van bachelor- en masterproeven in hun huidige vorm. Hoe behendig zijn studenten echt met generatieve artificiële intelligentie (GenAI) en hoe kritisch stellen zij zich op tegenover deze nieuwe uitdaging voor academische integriteit?

### 2. (Academisch) schrijven en technologie?

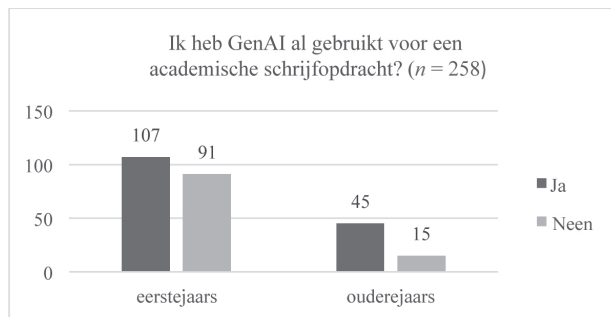
Schrijven is een cognitieve vaardigheid die niet alleen een diepgaand begrip van taal en grammatica vraagt, maar ook het vermogen om ideeën helder over te brengen. Binnen de cognitieve en ontwikkelingspsychologie bestuderen wetenschappers al decennialang schrijfmodellen om dat proces te schetsen (Flower & Hayes 1980; Bereiter & Scardamalia 1987; Zimmerman & Risemberg 1997). Technologische innovaties veranderen de aard van het schrijfproces en zorgen steeds voor conceptuele verschuivingen (Cummings 2023).

Binnen de academische gemeenschap neemt schrijfvaardigheid een essentiële plaats in, aangezien veel kennis schriftelijk geëvalueerd en gepresenteerd wordt. Er is bijgevolg een specifiek academisch register met eigen stijlkenmerken. Ondanks het belang van deze vaardigheid, signaleren onderzoeken dat die daalt bij studenten (de Bakker e.a. 2015). Het is daarom essentieel om te onderzoeken welke digitale ondersteuningsmechanismen beschikbaar zijn voor studenten en welke rol GenAI daarin kan spelen.

Deze tekst focust op hoe GenAI ervaren en gepercipieerd wordt door universiteitsstudenten aan de KU Leuven. Aan de hand van stappen uit het schrijfproces van Flower & Hayes (1987) bestudeert het onderzoek of studenten zich in het eerste en het derde bachelorjaar anders verhouden tegenover GenAI in academische schrijfoopdrachten. In het voorjaar van 2024 heeft een *mixed-methods* onderzoek die vraag verkend. Een online enquête met stellingen (studie A) werd verdiept aan de hand van schrijfoopdrachten waarbij elf studenten ChatGPT gebruikten en hun gedachten simultaan uitten met *think aloud protocols* (studie B).

### 3. Resultaten vragenlijst

De resultaten van de anonieme enquête, afgenomen bij 258 studenten Sociale Wetenschappen of Letteren aan de KU Leuven, tonen aan dat de grote meerderheid vooral bekend is met ChatGPT en ongeveer 60% het inzet voor academische schrijfoopdrachten. Opvallend is dat procentueel gezien bijna de helft (46,46%) van de eerstejaarsstudenten nooit GenAI gebruikt, terwijl dit bij de ouderejaars slechts 25% van de groep beslaat (zie: Figuur 1).



Figuur 1 – Gerapporteerd gebruik GenAI in schrijfoopdrachten ( $n = 258$ ).

Redenen om geen GenAI te gebruiken variëren: angst voor plagiaat of fraude, een hoog vertrouwen in de eigen schrijfvaardigheid, er onvoldoende mee vertrouwd zijn, slechte ervaringen die vooral duiden op ontbrekende bronvermeldingen en simpelweg een gevoel van onbehagen en angst.

De fervente gebruikers geven aan het vooral tijdens de *pre-writing* en *writing*-fasen in te schakelen om concepten uit te leggen, te brainstormen en teksten te verrijken. Ervaren schrijvers (Adv,  $n = 188$ ) rapporteren een frequenter gebruik dan onervaren schrijvers (Nov,  $n = 60$ ) voor alle doeleinden, behalve bij het zoeken naar bronnen. Tabel 1 geeft een gedetailleerd overzicht van scores op 5 voor een aantal stellingen waarbij een 0 overeenkomt met “doe ik nooit” en een 5 met “doe ik altijd”. Daarbij wordt telkens het gemiddelde (M), de standaarddeviatie (SD), de mediaan (Mdn) en modus (Mo) voor

de groepen getoond, evenals de statistische resultaten van een *Mann Whitney U-test* (W) met p-waarde en effectgrootte ( $r_{rb}$ ).

Stelling?	Groep	M	SD	Mdn	Mo	W	p	$r_{rb}$
Brainstormen over een onderwerp en inspiratie krijgen	Adv ( $n = 45$ )	2.444	1.078	3	3	3142	0.002	0.305
	Nov ( $n = 107$ )	1.850	1.148	2	2			
	Tot ( $n = 152$ )	2.026	1.156	2	2			
Moeilijke concepten uitleggen of definiëren	Adv ( $n = 45$ )	2.444	1.341	3	3	2940	0.024	0.221
	Nov ( $n = 107$ )	1.972	1.292	2	3			
	Tot ( $n = 152$ )	2.112	1.320	3	3			
Op gang komen en een leeg blad vermijden	Adv ( $n = 45$ )	1.711	1.236	2	2	3222	< 0.001	0.338
	Nov ( $n = 107$ )	0.981	1.149	1	0			
	Tot ( $n = 152$ )	1.197	1.218	1	0			
Genereren voorbeeldzinnen	Adv ( $n = 45$ )	1.711	1.141	2	3	3295.5	< .001	0.369
	Nov ( $n = 107$ )	0.963	1.009	1	0			
	Tot ( $n = 152$ )	1.184	1.100	1	0			
Herformuleren van eigen zinnen	Adv ( $n = 45$ )	2.444	1.056	3	3	3322.5	< .001	0.380
	Nov ( $n = 107$ )	1.636	1.177	2	2			
	Tot ( $n = 152$ )	1.875	1.198	2	2			
Eigen tekst samenvatten of lengte aanpassen	Adv ( $n = 45$ )	1.578	1.288	2	2	3189.5	< .001	0.325
	Nov ( $n = 107$ )	0.832	1.023	0	0			
	Tot ( $n = 152$ )	1.053	1.155	1	0			

Tabel 1 – Gerapporteerd omgang met GenAI bij academisch schrijven ( $n = 152$ ).

De participanten vinden GenAI over het algemeen een handig hulpmiddel, maar zijn ook sceptisch over de betrouwbaarheid en de academische integriteit van de gegenereerde inhoud. Ervaren schrijvers zijn globaal genomen positiever ingesteld dan onervaren schrijvers tenzij over het leerverlies en de duidelijkheid van de richtlijnen. De richtlijnen zijn opmerkelijk duidelijker voor de eerstejaarsstudenten ( $M = 2.664$ ,  $SD = 1.205$ ,  $Mdn = 3$ ) dan voor de ouderejaarsstudenten ( $M = 1.987$ ,  $SD = 1.158$ ,  $Mdn = 2$ ) met een statistisch significant verschil ( $W = 1643$ ,  $p = .001$ ,  $r_{rb} = -0.318$ ). De overige verschillen in perceptie zijn minder beduidend.

#### 4. Resultaten schrijfpdracht

In Studie B krijgen deelnemers een uur de tijd om op hun eigen laptop (met toegang tot het internet en ChatGPT), zonder bronnenstudie, een academische tekst van maximaal 500 woorden te produceren over een grafiek. Daarbij wordt de spontane interactie met ChatGPT vastgelegd aan de hand van de chatgeschiedenis en niet-gemedieerde simultane *think aloud protocols* of geverbaliseerde gedachtenuitingen (Ericsson & Simon 1996).

Deelnemer	Groep	Duur opdracht	Aantal woorden	Aantal prompts
Student 1	Novice	45:32	490	5
Student 2	Novice	36:36	314	6
Student 3	Novice	56:01	365	15
Student 4	Advanced	24:08	395	2
Student 5	Advanced	69:35	353	8
Student 6	Advanced	26:10	464	0
Student 7	Advanced	46:10	341	7
Student 8	Advanced	61:01	343	3
Student 9	Novice	51:24	454	9
Student 10	Advanced	47:41	476	4
Student 11	Advanced	53:10	487	7

Tabel 2 – Schrijfpdracht (duur, aantal woorden, aantal prompts) ( $n = 11$ ).

Tabel 2 laat zien dat het schrijfgedrag en de interactie met ChatGPT (aantal prompts) sterk varieert per individu en geen link met academische ervaring lijkt te vertonen. Vooral tijdens het brainstormen en om de stijl academischer te maken en eigen zinnen te herformuleren wordt een beroep gedaan op ChatGPT. De *think alouds* tonen verder aan dat deelnemers ook synoniemen opzoeken en hele teksten door GenAI laten genereren om die vervolgens te reviseren en in eigen woorden om te zetten. Ze gebruiken GenAI op verschillende momenten, als aanvulling in een arsenaal van hulpmiddelen, wat om een herziening van gangbare theoretische schrijfmodellen vraagt. Ook valt op dat sterke schrijfvaardigheden essentieel blijven bij het beoordelen van de output en dat sommige deelnemers verkiezen om nog steeds alles zelf te schrijven. Het formuleren van prompts blijkt een vaardigheid op zich te zijn die tekortkomingen vertoont. Frustraties wegens de ontoereikende en repetitieve antwoorden van GenAI leiden bij sommige studenten tot het volledig afschrijven en opgeven van de hulpmiddelen.

## 5. Conclusie

Het hoger onderwijs moet zich dus richten op het aanleren van *prompt engineering* en een betere kadering bieden voor toegestaan gebruik van GenAI in schrijfpdrachten. Om een schrijfpdracht goed te interpreteren en aan te pakken, blijken elementaire cognitieve vaardigheden en schrijfvaardigheden essentieel. Het kritisch inzicht overstijgt voorlopig nog de digitale hulp.

## Referenties

- Bereiter, C. & M. Scardamalia (1987). *The psychology of written composition*. Mahwah, NJ: Erlbaum.
- Cummings, L. (2023). 'Writing Processes in the Digital Age: A Networked Interpretation'. In: O. Kruse, C. Rapp, M. Anson, K. Benetos, E. Cotos, A. Devitt & A. Shibani (red.). *Digital Writing Technologies in Higher Education*. Midtown Manhattan, NYC: Springer International Publishing, p. 485-497.
- De Bakker, I., B. Meuffels, M. Muris & E. Rietstap (2015). "Diagnose en remediëring van schrijfvaardigheidsproblemen bij eerstejaarsstudenten Geesteswetenschappen". In: *Tijdschrift voor Taalbeheersing*, 37 (2), p. 127-157.
- Ericsson, K. & H. Simon (1996). *Protocol analysis: verbal reports as data*. Cambridge, Massachusetts: MIT press.
- Flower, L., J. Hayes, K. Schriver, J. Stratman & L. Carey (1987). 'Cognitive processes in revision'. In: S. Rosenerg (red.). *Advances in Applied Psycholinguistics II: reading, writing and language processes*. Cambridge: Cambridge University Press, p. 176-240.
- Zimmerman, B. & R. Risemberg (1997). "Becoming a Self-Regulated Writer: A Social Cognitive Perspective". In: *Contemporary Educational Psychology*, 22 (1), p. 73-101.