

Van pen tot prompt: academisch schrijven met ChatGPT

Een exploratief *mixed methods*-onderzoek naar de inzet van generatieve artificiële intelligentie in het schrijfproces van universiteitsstudenten

Lotte Uyttenbroeck

Masterproef aangeboden binnen de opleiding
master in de taal- en letterkunde

Promotor: prof. dr. Lieve De Wachter
Co-promotor: dr. Jordi Heeren

Academiejaar 2023-2024

141.871 tekens



Ik verklaar me akkoord met de code of conduct van de faculteit Letteren voor geloofwaardig auteurschap.

Dankwoord

Deze masterproef luidt het einde in van vier boeiende en leerrijke jaren taal- en letterkunde en zou nooit tot stand zijn gekomen zonder de hulp van een aantal belangrijke figuren. Daarom wil ik in dit dankwoord ten eerste mijn promotor en copromotor, Lieve De Wachter en Jordi Heeren, uit de grond van mijn hart bedanken. Zij reikten me dit interessante onderwerp aan, hebben uren mee gebrainstormd over het onderzoek en stonden tot vlak voor de deadline klaar met hun bemoedigende en constructieve feedback. Het is dankzij hun suggesties, oplettend oog en ideeën dat deze masterproef de huidige vorm gekregen heeft.

Ten tweede wil ik graag iedereen bedanken die bij de studie betrokken was. Alle 258 respondenten die de enquête invulden en de 11 deelnemers die een uur vrijmaakten om deel te nemen aan de schrijfo opdracht. Verder ook Lotte Malchair, Simen Verberckmoes, Lieve De Wachter en Jordi Heeren bij wie ik in bepaalde lessen de enquête mocht verspreiden.

Verder wil ik mijn ouders en broer bedanken om me de kans te geven zo'n mooie opleiding zorgeloos te voltooien. Zij stonden elke weekend klaar om een luisterend oor te bieden als ik gefrustreerd was, een licht beschuldigende blik als ik uitstelgedrag vertoonde of een warme knuffel als ik het even niet meer zag zitten. Mijn ouders voor hun onvoorwaardelijke steun en mijn broer in het bijzonder om als proefkonijn enquêtes uit te testen.

Ten slotte wil ik ook al mijn vrienden, kotgenoten en overige familieleden bedanken om me te motiveren en om alle interessante krantenartikelen die ze lazen over ChatGPT met mij te delen. Jullie zorgden voor titelinspiratie, de welkome ontspanning en de humoristische toets die ik soms nodig had om het plezier in het onderzoek te herontdekken. Bedankt!

Inhoudsopgave

Dankwoord	3
Samenvatting.....	6
Lijst met figuren en tabellen	7
Inleiding	8
Deel 1: Theoretisch kader	9
1.1 Schrijfprocessen	9
1.1.1 cognitieve modellen	9
1.1.1.1 Flower & Hayes (1980)	9
1.1.1.2 Bereiter & Scardamalia (1987)	11
1.1.2 sociaal-cognitieve modellen	12
1.1.2.1 Zimmerman en Risemberg (1997)	12
1.1.2.2 Hayes (1996)	13
1.1.3 socioculturele modellen	15
1.2 Academisch schrijven.....	17
1.3 Schrijven en technologie	18
1.3.1 Technologische evoluties.....	18
1.3.2 Conceptuele evoluties.....	19
1.3.3 Elektronische schrijfondersteuning	19
1.3.4 Generatieve Artificial Intelligence	21
1.3.5 KU Leuven en generatieve Artificial Intelligence	23
Deel 2: Studie (A) vragenlijst	24
2.1. Respondenten.....	24
2.2 Materiaal en methode	26
2.3 Analyse en resultaten	28
2.3.1 Omgang met generatieve AI: algemeen	28
2.3.2 Omgang met generatieve AI: per schrijffase	30
2.3.2.1 Voor het schrijven	30
2.3.2.2 Tijdens het schrijven.....	32
2.3.2.3 Na het schrijven.....	33
2.3.2.4 Tussentijdse conclusie: novice en advanced writers.....	34
2.3.3 Houding tegenover generatieve AI.....	34

Deel 3: Studie (B) schrijfpdracht met <i>think aloud protocols</i>	37
3.1 Respondenten.....	37
3.2 Materiaal en methode	38
3.3 Analyse en resultaten	41
3.3.1 Omgang met ChatGPT	41
3.3.1.1 Algemeen overzicht.....	41
3.3.1.2 Voor het schrijven	43
3.3.1.3 Tijdens het schrijven.....	46
3.3.1.4 Na het schrijven.....	49
3.3.2 Houding tegenover ChatGPT	50
3.3.3 Terugkoppeling naar het schrijfmodel	53
Deel 4: Discussie.....	55
4.1 Beperkingen.....	56
4.2 Gebruik van GenAI (ChatGPT) in schrijfpdrachten	56
4.3 Omgang met ChatGPT bij (on)ervaren schrijvers	57
4.4 Houding tegenover ChatGPT bij (on)ervaren schrijvers	59
4.5 ChatGPT en het schrijfproces	60
Deel 5: Conclusie	61
Bibliografie	63
Bijlagen	67
Bijlage 1: Enquête uit studie (A)	67
Bijlage 2: Schrijfpdracht uit studie (B)	69
Bijlage 3: Alle resultaten enquête uit studie (A)	70
Bijlage 4: teksten, chatgeschiedenis en think alouds studie (B)	70

Samenvatting

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; Engeström 1987; Zimmerman & Risemberg 1997). Technologische innovaties veranderen de aard van het schrijfproces echter en zorgen 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 ook een specifiek academisch register met eigen stijkenmerken (Van Kalsbeek & Kuiken 2014). Ondanks het belang van deze vaardigheid, signaleren onderzoeken dat die daalt bij studenten (Berckmoes & Rombout 2009; De Bakker et al 2015). Het is daarom essentieel om te onderzoeken welke digitale ondersteuningsmechanismen beschikbaar zijn voor studenten (Allen et al. 2016; Strobl 2019) en welke rol generatieve artificiële intelligentie (GenAI) daarin kan spelen sinds november 2022 (Su et al. 2023). De onderzoeksvragen luiden dan ook “In welke mate en voor welke doeleinden schakelen (on)ervaren academische schrijvers GenAI in voor academische schrijfp opdrachten?” en “Wat is de houding van (on)ervaren academische schrijvers ten opzichte van GenAI?”

Om een antwoord te bieden op die onderzoeksvragen, heeft deze masterproef kwantitatief en kwalitatief onderzoek gecombineerd. Een online enquête met stellingen in matrixtabellen (studie A) wordt verdiept aan de hand van schrijfp opdrachten waarbij 11 studenten ChatGPT gebruiken en in real time geobserveerd worden met *think aloud protocols* (studie B). De enquête, afgenomen bij 258 studenten sociale wetenschappen of letteren aan de KU Leuven, toont dat de grote meerderheid vooral bekend is met ChatGPT en ongeveer 60% het inzet voor academische schrijfp opdrachten. Ze geven aan het vooral te gebruiken tijdens de *pre-writing* en *writing*-fasen om concepten uit te leggen, te brainstormen en teksten te verrijken. Ze vinden het een handig hulpmiddel maar zijn ook sceptisch over de betrouwbaarheid en academische integriteit van de gegenereerde inhoud. Ervaren schrijvers ($n = 188$) rapporteren een frequenter gebruik dan onervaren schrijvers ($n = 60$) voor alle doeleinden behalve het zoeken van bronnen en staan positiever ingesteld tenzij over het leerverlies en de duidelijkheid van de richtlijnen. Studie B laat zien dat studenten ook synoniemen opzoeken en hele teksten door GenAI laten genereren om die vervolgens te reviseren, waarbij gedrag sterk varieert per individu en geen link met de academische ervaring lijkt te vertonen. Ze gebruiken GenAI als aanvulling in een arsenaal van hulpmiddelen en de resultaten tonen aan dat de gangbare theoretische schrijfmodellen herzien moeten worden. Ook valt op dat sterke schrijfvaardigheden essentieel blijven bij het beoordelen van de output en sommige deelnemers hier voldoening uithalen en verkiezen alles zelf te schrijven. Uit studie B blijkt dat het formuleren van prompts een vaardigheid op zich is die tekortkomingen vertoont. Het hoger onderwijs moet zich zeker ook richten op het aanleren van *prompt engineering* en een betere kadering bieden van toegestaan gebruik van GenAI in schrijfp opdrachten. (3402 tekens)

Lijst met figuren en tabellen

Figuur 1: Schrijfmodel volgens Flower & Hayes (1980: 370)	10
Figuur 2: Knowledge transforming process van Bereiter en Scardamalia (1987: 12)	11
Figuur 3: Herziene schrijfmodel van Hayes (1996: 4).....	13
Figuur 4: Schrijfmodel volgens Lei gebaseerd op Engeström (2016: 106)	16
Figuur 5: Leeftijd respondenten (n = 258).....	24
Figuur 6: Studierichting en fase respondenten (n = 258).....	25
Figuur 7: Studiefase respondenten (n = 258).....	26
Figuur 8: Gerapporteerde frequentie gebruik GenAI (n = 258)	29
Figuur 9: Gerapporteerd gebruik GenAI in academische schrijfoopdrachten (n = 258)	29
Figuur 10: Beïnvloedende factoren studiekeuze (in absolute aantallen)	70
Tabel 1: Geslacht, nationaliteit en moedertaal respondenten (n = 258).....	24
Tabel 2: Gebruik van GenAI voor het schrijven (n = 152).....	31
Tabel 3: Gebruik GenAI tijdens het schrijven (n = 152).....	32
Tabel 4: Gebruik GenAI na het schrijven (n = 152).....	33
Tabel 5: Houding ten opzichte van GenAI (n = 152).....	35
Tabel 6: Studie, geslacht, leeftijd, moedertaal participanten (n = 11).....	38
Tabel 7: Overzicht schrijfoopdrachten: duur, woorden, prompts (n = 11)	41
Tabel 8: Overzicht schrijfoopdracht: codering strategieën (n = 11).....	43
Tabel 9: Gebruik GenAI volgens enquête en opdrachten (n = 11).....	51
Tabel 10: Retorisch probleem: nakijken aantal woorden (n = 11).....	54

Inleiding

Schrijven is een complexe cognitieve vaardigheid waarvoor niet alleen een diepgaand begrip van taal en grammatica nodig is maar ook het vermogen om ideeën helder en gestructureerd over te brengen. Binnen de cognitieve psychologie en ontwikkelingspsychologie buigen wetenschappers zich al decennialang over het uittekenen van modellen die de cognitieve processen en componenten daarvan vatten (Flower & Hayes 1980; Bereiter & Scardamalia 1987; Engeström 1987; Hayes 1996; Zimmerman & Reisemberg 1997). Technologische innovaties, zoals de komst van word processors, het internet of Natural Language Processing zorgen voor veranderingen in de aard van het schrijfproces (Cummings 2023). Schrijfvaardigheid vormt bovendien een essentieel onderdeel van de academische gemeenschap, waar studenten worden uitgedaagd om inzichten te synthetiseren en presenteren in een specifiek academisch register (Van Kalsbeek & Kuiken 2014; Turner 2012). Ondanks het cruciale belang van academische schrijfvaardigheid, wordt steeds vaker gerapporteerd dat die de laatste jaren daalt bij studenten (Berckmoes & Rombout 2009; Bonset 2010; De Bakker et al 2015). In deze context is het essentieel om te onderzoeken hoe de recentste technologieën en digitale innovaties kunnen worden ingezet om de schrijfvaardigheden van studenten te verbeteren (Allen et al. 2026; Strobl et al. 2019). De komst van generatieve Artificiële Intelligentie (GenAI) zoals ChatGPT hoort thuis in een reeks technologische innovaties en biedt zowel mogelijkheden als valkuilen voor het schrijfproces (Beck & Levine 2023; Su et al 2023; Barrot 2023). Dit onderzoek richt zich specifiek op de manier waarop (on)ervaren studenten aan de KU Leuven omgaan met GenAI in academische schrijfopdrachten en hoe zij hiertegenover staan.

Na een schetsing van het theoretisch kader (deel 1), voert deze masterthesis twee studies uit bij studenten sociale wetenschappen en letteren aan de KU Leuven. Studie A (deel 2) bestaat uit een enquête afgenomen bij 258 respondenten die onderzoekt hoe zij omgaan met GenAI en wat hun percepties zijn ten aanzien van deze technologie. Deze enquête bevat vragen over de frequentie van het gebruik, de doeleinden waarvoor GenAI wordt ingezet en de mogelijke voor- en nadelen. Studie B (deel 3) gaat een stap verder en observeert het schrijfproces met ondersteuning van ChatGPT bij 11 studenten in real-time. Tijdens het schrijven van een tekst moeten zij hardop nadenken (*think aloud protocols*) en wordt de interactie met ChatGPT vastgelegd. Zo kunnen gehanteerde strategieën geëxpliciteerd worden. In elke studie is er aandacht voor de respondenten, materiaal en methode en tot slot de analyse en resultaten. Daarop volgt een discussie (deel 4) die opvallende inzichten uit beide studies belicht en becommentarieert. Het afsluitend gedeelte (deel 5) formuleert een conclusie en suggesties voor toekomstig onderzoek.

Deel 1: Theoretisch kader

1.1 Schrijfprocessen

Schrijven is een complexe cognitieve vaardigheid die informatie verwerkt en synthetiseert. Deze sectie bespreekt opeenvolgende invloedrijke modellen over het schrijfproces die zich in een cognitief paradigma (1.1.1), sociaal-cognitief paradigma (1.1.2) of sociocultureel paradigma (1.1.3) situeren. Met aandacht voor onderlinge verschillen en overeenkomsten hoopt dit hoofdstuk te verduidelijken op welke aspecten van het schrijfproces academici de afgelopen decennia gefocust hebben.

1.1.1 cognitieve modellen

1.1.1.1 Flower & Hayes (1980)

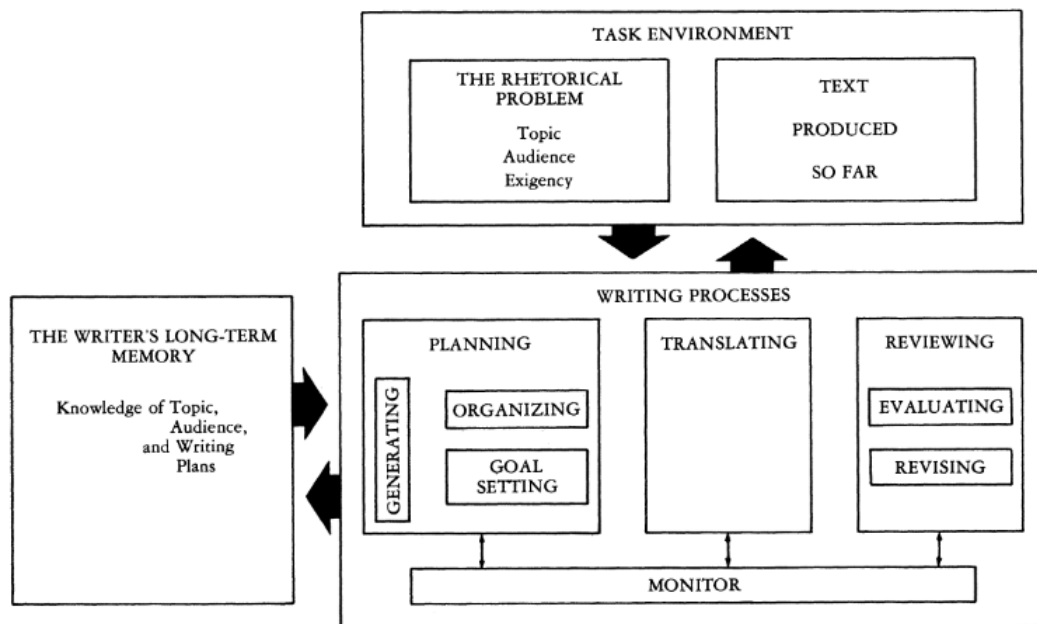
Het cognitief model dat Flower & Hayes in hun artikel *A cognitive process theory of writing* (1980) voorstelden ging in tegen de destijds heersende sequentiële modellen zoals uitgewerkt door Rohman (1965). Een sequentieel model deelt het schrijfproces op in drie lineaire fases die na elkaar plaatsvinden: *Pre-writing*, *Writing* en *Re-writing*. *Pre-writing* wijst op alles dat zich afspeelt voor er woorden op papier verschijnen, de *Writing* fase is die waarin woorden geproduceerd worden en de *Re-writing* fase herwerkt het bekomen product en voert correcties door. Voor een dergelijk model is alleen het eindproduct of de geproduceerde tekst belangrijk (Flower & Hayes 1980: 367).

Aan de hand van *protocol analyses* waren de Amerikaanse linguïst Flower en cognitief psycholoog Hayes echter in staat om niet alleen het eindproduct maar ook hiërarchisch gestructureerde mentale processen die actief zijn tijdens het schrijven te bestuderen. Om die denkprocessen bloot te leggen, vragen zij aan deelnemers in experimenten om hun losse, fragmentarische gedachtegangen en bedenkingen luidop uit te spreken tijdens het uitvoeren van een schrijfoopdracht. Het doel is niet om analyserend commentaar of introspectieve bedenkingen te uiten maar simpelweg alles wat door hun hoofd gaat te verbaliseren. Al deze *think aloud* uitingen worden opgenomen en vervolgens getranscribeerd in wat als een *protocol* benoemd wordt. Aan de hand daarvan abstraheren de onderzoekers algemene mentale processen (Flower & Hayes 1980: 368).

Deze experimentele methode resulteerde in een conceptueel model bestaande uit drie componenten dat ontzettend invloedrijk geweest is en vandaag nog vaak als basis dient voor studies rond schrijfvaardigheid. Een schematische voorstelling (zie Figuur 1) verduidelijkt dit model ook al waarschuwen Flower en Hayes zelf voor een mogelijke misleidende interpretatie wat betreft de pijlen

(Flower en Hayes 1980: 386) en verplaatst Hayes in een latere herziening de kaders om de onderlinge relaties beter in beeld te brengen (Hayes 1996: 3).

Figuur 1: Schrijfmodel volgens Flower & Hayes (1980: 370)



De eerste component is de taakomgeving die zowel het retorisch probleem of de opdracht (onderwerp, publiek, schrijfdoelen) als de reeds geproduceerde tekst bevat. Het definiëren van het retorisch probleem is een essentieel onderdeel van het schrijfproces aangezien dit de schrijfdoelen en uiteindelijke stijl vormgeeft.

De tweede component uit het model is het langetermijngeheugen. Dat bevat kennis over het onderwerp, het doelpubliek en schrijfplannen. Schrijfplannen wijzen bijvoorbeeld op de structuur van een tekst met een inleiding-midden-slot of op kennis van genres. Deze component is als vergaarde kennis op de achtergrond aanwezig en vormt de basis waar een schrijver op steunt tijdens alle processen.

De cognitieve processen vormen de derde component in het model van Flower & Hayes. Deze zijn op hun beurt opgedeeld in drie deelprocessen: plannen, vertalen en reviseren. Plannen zien zij als het vormen van een interne representatie van de kennis die in het schrijven gebruikt zal worden en kan zich uiten in de vorm van een sleutelwoord of een visueel beeld. Tijdens het plannen worden ideeën gegenereerd, georganiseerd en doelen vastgezet. Die doelen zijn beslissingen van de schrijver zelf over het begin van een tekst of de betrokken argumenten. Vertalen wijst naar de handeling van het transcriberen of neerschrijven. Reviseren bestaat ten slotte uit een constante herhaling van lezen en aanpassen, zowel van de geschreven tekst als van gevormde ideeën. De verschillende cognitieve

processen worden bestuurd door een monitor. De monitor is geen overheersend controlemechanisme maar dient individualiteit aan te duiden en verklaart waarom sommige schrijvers bijvoorbeeld veel plannen en anderen net onmiddellijk vertalen (Hayes 2012: 373).

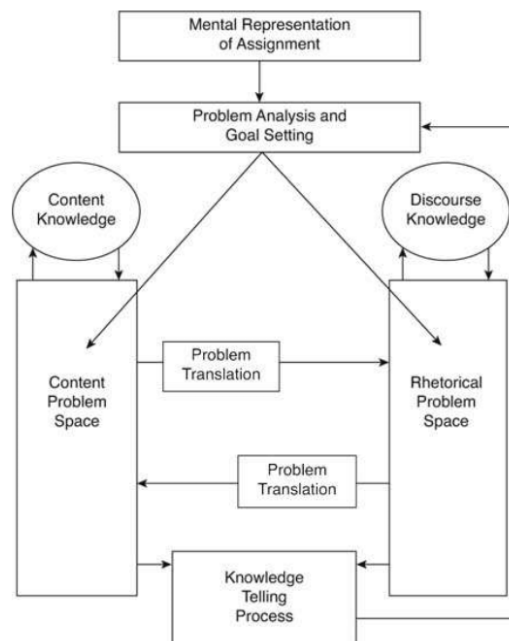
Deze drie dimensies maken duidelijk dat Flower en Hayes schrijven als een doelgericht proces zagen dat constant verschillende vaardigheden combineert die tegelijk aandacht vragen (McArthur & Graham 2016: 26).

1.1.1.2 Bereiter & Scardamalia (1987)

De Canadese ontwikkelingspsychologen Bereiter en Scardamalia stelden in 1987 ook een cognitief schrijfmodel op waarbij ze zich vooral bogen op het verschil tussen ervaren en onervaren schrijvers. Hun onderzoek had zich de jaren hiervoor namelijk toegespitst op schrijfvaardigheid bij kinderen.

Zij onderscheiden twee soorten schrijfmodellen, het *knowledge-telling* en *knowledge-transforming* model. *Knowledge telling* is een simpele opeenvolging van stappen aansluitend bij de natuurlijke dimensie van schrijven als iets vanzelfsprekend terwijl *knowledge transforming* (zie Figuur 2) een complex proces is dat de problematische definitie van schrijven als een zware cognitieve taak verklaart.

Figuur 2: Knowledge transforming process van Bereiter en Scardamalia (1987: 12)



Het *knowledge telling* proces draait rond het op schrift stellen van gedachten die reeds aanwezig zijn in het hoofd van een lezer en is daarom vrij analoog met een conversatie waarbij je luidop uitspreekt wat je associeert met een bepaald onderwerp. Het mentale proces steunt op drie externe bronnen: de tekst voor zover die geproduceerd is, de kennis over het onderwerp en de retorische kennis over de verwachte literaire vorm. Een mentale voorstelling van de opdracht genereert ideeën of *cues* die intern

beoordeeld worden en als ze thematisch relevant en retorisch gepast lijken, neergeschreven worden. De reeds geproduceerde zinnen roepen nieuwe *cues* op en initiëren het proces opnieuw. In het *knowledge transforming* proces (zie Figuur 2) daarentegen is er sprake van het actief verwerken van zowel gedachten als de reeds geschreven tekst. Ervaren schrijvers beoordelen namelijk of hun schrijfproducten wel bevatten wat ze willen uitdrukken en of ze zelf geloven wat die schrijfproducten zeggen (Bereiter en Scardamalia 1987: 11). Het denkproces wordt dus ingebed in een ruimere probleemanalyse die zowel beroep doet op een inhoudelijke als een retorische ruimte. Retorische twijfels over de duidelijkheid van een paragraaf kunnen bijvoorbeeld leiden tot het inhoudelijk toevoegen van een definitie en zo veranderen de schrijfdoelen constant. Het uiteindelijk neerschrijven van al deze reflecties verloopt volgens het *knowledge telling* proces zoals hierboven uitgelegd en deze input kan opnieuw een *knowledge transforming* proces stimuleren.

Verscheidene elementen uit dit model zoals de reeds geproduceerde tekst, inhoudelijke kennis, genrekennis, beoordelingscomponenten, ... waren ook in het model van Flower en Hayes (1980) aanwezig. Bereiter en Scardamalia leggen de nadruk echter meer op interactie tussen deze elementen en de mentale transformaties die plaatsvinden. De opvatting van schrijven als een spontaan *vertellen* dat minder beroep doet op cognitieve processen bij onervaren schrijvers en kinderen is ook nieuw.

1.1.2 sociaal-cognitieve modellen

1.1.2.1 Zimmerman en Risemberg (1997)

In het sociaal-cognitief model van Zimmerman en Risemberg ligt de nadruk op schrijven als een sociaal gesitueerde *performance* in plaats van een cognitieve *competence* (Zimmerman & Risemberg 1997: 75). Ze verwijzen hierbij onder andere naar het feit dat schrijvers hun schrijfproces zelf inplannen, initiëren en volbrengen en daarbij voldoende motivatie moeten opbrengen om hun tekst retorisch en communicatief te perfectioneren. Dit gebeuren wordt volgens hen gestuurd door zelfregulerende processen die via een cyclische feedbackloop bijdragen tot een gevoel van *self-efficacy* en intrinsieke motivatie. Als een schrijver een bepaald proces als succesvol beschouwt, zal hij meer geneigd zijn dit proces te herhalen en zo zijn uiteindelijke schrijfvaardigheid te bevorderen.

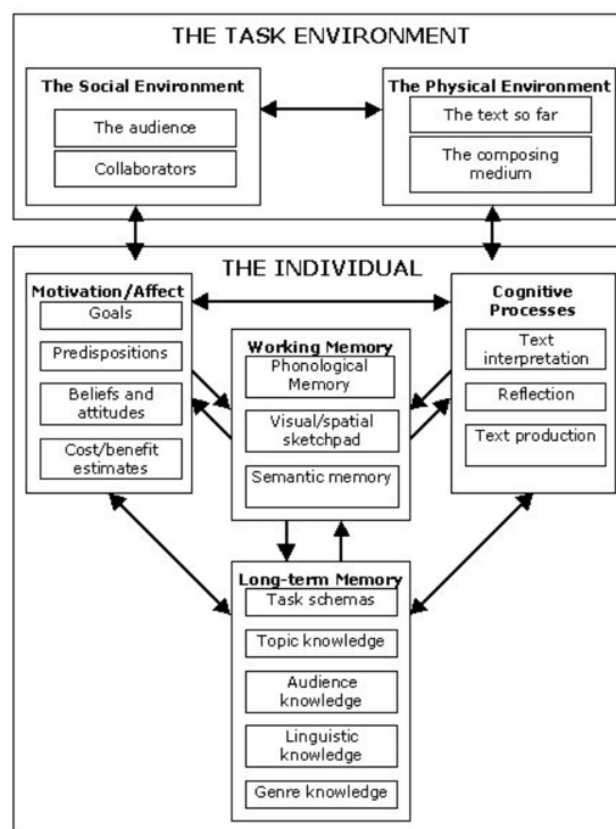
Deze zelfregulerende processen ontwikkelen zich op informele wijze en kunnen zowel gedachten, gevoelens als acties zijn. Zimmerman en Risemberg stellen een driedelige classificatie op waarin ze verschillende processen groeperen. In de eerste plaats de omgevingsprocessen of *environmental processes* die bestaan uit het creëren van een bevorderlijke fysieke en sociale omgeving. Die omgeving kan de vorm aannemen van een rustige plaats zonder muziek of gsm-meldingen maar ook een vertrouwde met de schrijfstijl van een rolmodel of leerkracht. Gedragsstrategieën of *behavioral*

strategies duiden vervolgens de motorische activiteiten aan die een schrijver zichzelf aanleert. Voorbeelden zijn het tellen van het aantal geproduceerde pagina's per dag, zichzelf belonen met iets ontspannend na het voltooien van een schrijfo opdracht of het luidop voorlezen van een tekst om te horen hoe die klinkt. De persoonlijke processen of (*covert*) *personal processes* nemen ten slotte de vorm aan van overtuigingen of gevoelens. Hieronder vallen bijvoorbeeld het verlagen van het stressgehalte door een vast moment per dag te reserveren om te schrijven, het opstellen van schrijfdoelen en criteria om zichzelf te beoordelen en het vormen van mentale beelden bij karakterbeschrijvingen.

1.1.2.2 Hayes (1996)

Steunend op inzichten van andere onderzoekers, bracht Hayes in 1996 een nieuw model (zie *Figuur 3*) uit dat hij als een individueel-omgevingsmodel bestempelt (Hayes 1996: 5). Aangezien dit model echter aansluit bij het gevoel van *self-efficacy* zoals beschreven door Zimmerman & Risemberg, past dit model ook binnen de sociaal-cognitieve modellen. Zelf benadrukt Hayes vier grote veranderingen ten opzichte van zijn eerder model uit 1980: toevoeging van het werkgeheugen, focus op visueel-spatiale elementen, belang van motivatie en grotere aandacht voor tekstinterpretatie.

Figuur 3: Herziene schrijfmodel van Hayes (1996: 4)



Dit nieuwe model bestaat niet langer uit drie maar slechts twee dimensies: de taakomgeving en het individu. De taakomgeving omvat nu ook samenwerkingspartners en het schrijfmedium dat verwijst

naar verschillen tussen schrijven met pen en papier of schrijven met een tekstverwerker. Binnen de individuele dimensie zijn naast de cognitieve processen en het langetermijngeheugen nog twee componenten toegevoegd: motivatie en het werkgeheugen. Terwijl de eerstgenoemde component een verklaring biedt voor de link tussen zelfvertrouwen en gepercipieerd succes, fungeert de laatstgenoemde component als een soort doorgeefluik tussen het langetermijngeheugen en de cognitieve processen.

De cognitieve processen, die in dit model te situeren zijn binnen de individuele dimensie, hebben ook ingrijpende veranderingen doorgemaakt. Het model uit 1996 neemt afstand van de drie deelprocessen plannen, vertalen en beoordelen en maakt een nieuwe opdeling bestaande uit tekstinterpretatie, reflectie en tekstproductie. Waar tekstinterpretatie vooral wijst naar het mentaal beeld dat al lezend/luisterend/kijkend gevormd wordt, bestaat reflectie uit algemene vaardigheden zoals het oplossen van problemen, maken van beslissingen en conclusies trekken uit informatie. Tekstproductie komt overeen met het vroegere vertaalproces.

Een verklaring voor het verdwijnen van plannen en beoordelen ligt in de natuur van beide deelprocessen die gereduceerd kunnen worden tot schrijfprocessen op zich, als ze op schrift gesteld worden, of tot deel uitmakend van schrijfdoelen als ze zich in het hoofd van een schrijver afspelen. Zo formuleert Hayes het alleszins zelf in zijn contemplatief artikel *Modeling and remodeling writing* (2012):

“In this view, revision, like planning, is seen not as a separate writing process parallel to the other writing processes [...] but rather as a special application of the writing model.” (Hayes 2012: 376)

Als een student er bijvoorbeeld voor kiest om de structuur van zijn argumentatief essay schematisch op papier te plannen, is dit plannen een apart schrijfproces met een eigen taakomgeving waarvoor zijn werkgeheugen, langetermijngeheugen en interne motivatie vereist zijn. Het doelpubliek is dan de student zelf die het schema moet begrijpen, samenwerking is niet aanwezig, het medium bestaat uit het kladblaadje waarop hij dit schema produceert en elk neergeschreven kernpunt beïnvloedt de argumenten die nog zullen komen. Als deze student zou overgaan tot het nalezen en corrigeren van zijn essay, is dat ook een specifieke schrijfactiviteit waarbij hij als doel heeft om oplossingen voor gedetecteerde problemen in de geschreven tekst te vinden (Hayes 2012: 376). De student zal voor de keuze staan om de constructies waar hij niet tevreden mee is te herschrijven/parafraseren of om deze te diagnosticeren en reviseren (Flower et al. 1986; Hayes et al. 1987).

De toevoeging van het werkgeheugen en van de motivatiecomponent zijn voor Graham de belangrijkste aanpassingen van dit herziene model (Graham 2006, geciteerd in McArthur en Graham 2016: 27).

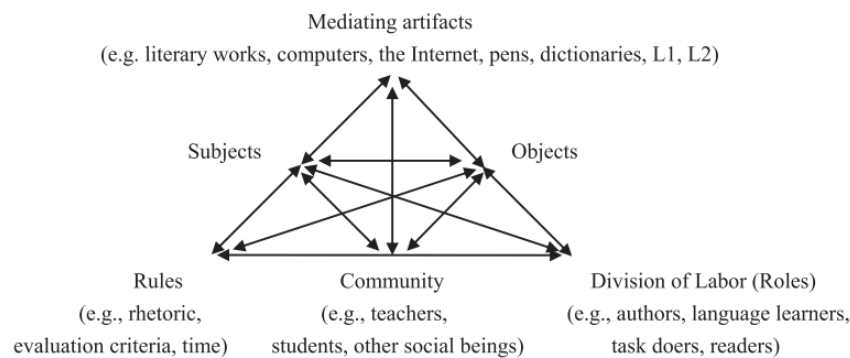
1.1.3 socioculturele modellen

Socioculturele schrijfmodellen verzetten zich tegen het cognitieve paradigma dat zij te beperkt vinden wat betreft de aandacht voor context (Prior 2006: 54). In tegenstelling tot de sociaal-cognitieve modellen, ligt de focus niet langer op cognitieve processen maar op de ruimere context die schrijven omgeeft. Schrijven is voor aanhangers van een socioculturele stroming niet zomaar een materiële activiteit bestaande uit het maken van teksten en ook geen communicatieve activiteit die een boodschap wil overbrengen maar een sociaal gemedieerde en gesitueerde activiteit (Prior 2006: 58; Lei 2008: 219).

Het begrip mediatie ontleen zij aan Vygotski die mediatie beschreef als een indirecte link tussen binnenkomende stimuli en menselijke reacties. Voor Vygotski controleren mensen namelijk hun interne wereld of eigen geest met psychologische tools zoals taal en tekensystemen en de externe buitenwereld aan de hand van technologische tools zoals werkgereedschap (Lei 2008: 219). Mediërende tools verbonden aan het schrijfproces kunnen dus zowel wijzen naar de letters uit ons alfabet en onze kennis over taal en tekststructuren als naar instrumenten zoals een pen, computer, woordenboek, grammatica- of spellingschecker (Englert et al. 2008: 211; Lei 2008: 220). Zelfs een schrijver die geïsoleerd op een zolderkamer schrijft is aan de hand van deze sociaalhistorische hulpmiddelen altijd verbonden met de sociale context van mensen die hem deze middelen aanreikten of aanleerden (Prior 2006: 58).

Mediatie kreeg binnen de *Activity Theory* van Engeström een verdere uitwerking en diende vier verschillende mediators aan te duiden die de interactie tussen een subject en object vormgaven (zie *figuur 4*). Mediërende artefacten of tools zoals hierboven beschreven maar ook regels, de gemeenschap en een arbeids- of rolverdeling. Steunend op het model van Engeström, onderzocht Lei of studenten tijdens het schrijven van opdrachten vooral strategieën gebruikten die aansloten bij artefacten (bijvoorbeeld het navolgen van een literair werk), regels (zoals de evaluatieve criteria van een opdracht), de gemeenschap (gevormd door medestudenten en de leerkracht) of de rolverdeling (met name het verschil tussen een moedertaalspreker en tweede taal verwerver van een taal) (Lei 2008).

Figuur 4: Schrijfmodel volgens Lei gebaseerd op Engeström (2016: 106)



Deze opdeling sluit in grote lijnen aan bij de drie principes die Englert et al. in hun artikel *Tenets of sociocultural theory in writing instruction research* bespreken (2006). Het eerste principe is namelijk de sociale interactie tussen experts/leerkrachten en beginnende schrijvers/studenten. Dit principe is vergelijkbaar met de rolverdelingscomponent uit het model van Lei. Door verschillende strategieën, instructies en *think alouds* aan te reiken, probeert de expert te bereiken dat de onervaren schrijver deze tools internaliseert en na verloop van tijd als coach voor zichzelf kan optreden (zie bijvoorbeeld Lei 2016). Er vindt in essentie een shift plaats van een *intermental* naar een *intramental* niveau waarbij vereende, samengebrachte inspanningen van de coach en zijn leerling (inter-, verwijzend naar meerdere personen) leiden tot individuele denkkaders van de leerling (intra-, verwijzend naar één persoon). De expert moet zich flexibel kunnen opstellen en aan de hand van *step-in* en *step-back* moves aanvoelen wanneer de student voldoende kennis verworven heeft om zelf aan de slag te gaan om de aangereikte strategieën toe te passen. Ook interacties tussen studenten onder elkaar kunnen helpen bij het externaliseren van mentale processen en het verwerven van meer inzichten (Englert et al. 2006: 209-210).

Het tweede principe sluit aan bij de eerder besproken mediërende artefacten en benoemen Englert et al. als *procedural facilitators and tools*. Tools verwijzen naar de mentale, linguïstische en fysieke hulpmiddelen (notatiesystemen, diagrammen, tekststructuren, geheugensteuntjes, spellingcheckers, ...) die de cognitieve last van schrijven verminderen door bepaalde aspecten over te nemen. De benaming procedurale geleider duidt de actie aan van het blootleggen van persoonlijke processen en aanreiken van bovengenoemde tools en cognitieve ondersteuning (Englert et al. 2006: 211-214).

Englert et al. bespreken als derde de gemeenschap, *community of practice*, met aandacht voor conventies, genres en waarden. In het bijzonder binnen de academische wereld is deelname aan de disciplinaire schrijfgemeenschap niet alleen hoe studenten zichzelf bewijzen maar ook hoe ze leren en

hoe ze betekenis toekennen aan schrijven (Lemke 2002, geciteerd in Englert et al. 2006: 215). Ook de dialogische interactie tussen een schrijver en zijn imaginaire lezers die teruggaat op het werk van Bakhtin plaatsen Englert et al. onder het gemeenschapsprincipe.

Een sociocultureel model neemt afstand van de cognitieve modellen en richt zijn blik vooral op de sociale interactie tussen experts en beginners, tussen leden van eenzelfde gemeenschap of tussen gebruikers en bedenkers van tools en denkkaders.

1.2 Academisch schrijven

Academisch taalgebruik verwijst naar het taalgebruik binnen de context van het hoger onderwijs (universiteiten en hogescholen) en wetenschap (Van Kalsbeek & Kuiken 2014: 221). Het is een specifiek register gebruikt door academici (in opleiding) om inhoud over te brengen op een doelpubliek van mede-experten uit het wetenschappelijk vakgebied waar ze zelf toe behoren. Voor studenten in het hoger onderwijs is het vaak een uitdaging om dit specifieke register te leren beheersen en zelf teksten te kunnen produceren die beantwoorden aan de verwachte normen. Verschillende behoefteanalyses en onderzoeksprojecten hebben mogelijke oorzaken en oplossingen voor die gesignaleerde moeilijkheden in kaart gebracht (Berckmoes & Rombout 2009; Bonset 2010; De Bakker et al 2015).

Een sluitende definitie van academisch schrijven geven is moeilijk aangezien gebruiken tussen verschillende taalgebieden en wetenschappelijke disciplines verschillen. Die gebruiken zijn gekaderd binnen ruimere culturele evoluties, zoals de Verlichting in West-Europa (Turner 2011). Verder bevat de academische wereld verschillende subgenres en zullen de talige kenmerken van een onderzoekspaper bijvoorbeeld anders zijn dan die van een academische recensie (De Wachter et al. 2021). Wat alle genres en disciplines binnen een westers kader echter gemeenschappelijk hebben is het streven naar een heldere, beknopte formulering. De boodschap staat centraal en daarom zal er *to the point* gecommuniceerd moeten worden zonder redundantie (Snow 2010).

Binnen de Angelsaksische taalkunde hebben onderzoekers grote corpora van academische teksten samengesteld om tot concrete eigenschappen te komen (zie bijvoorbeeld Hyland 2009). Uit dergelijke onderzoeken bleek dat vooral de structuur, zinsbouw en woordgebruik karakteriserend waren voor academische teksten. Zo zijn er meer complexe naamwoordgroepen, voorzetselconstructies en passief taalgebruik aanwezig door de hoge mate van nominalisering (Van Kalsbeek & Kuiken 2014).

Van den Branden (2010) stelt voor academisch Nederlands vijf typische kenmerken voor die niet alleen op zinsbouw en woordgebruik maar ook op stijl en logische verbanden focussen. Deze vijf kenmerken zijn: een hoog gehalte van niet-frequente woordenschat, complexe grammaticale structuren, impliciete verbanden, vage definities van termen en een onpersoonlijke/abstracte taal. Niet-frequente woordenschat kan enerzijds duiden op vakspecifieke termen uit de wetenschappelijke discipline van de auteur maar kan anderzijds ook algemene academische termen aanduiden. Complexe grammaticale structuren verwijzen naar lange zinnen met neven- en onderschikkingen. Met de impliciete verbanden beschrijft Van den Branden het feit dat causale zinnen elkaar soms opvolgen zonder dat voegwoorden hun verband expliciteren en beroep gedaan wordt op de algemene kennis van de lezer om de juiste verbanden te trekken. De vage definities van termen hangen samen met eerder vermelde kenmerken die ook in verduidelijkende definities aanwezig zijn en hun begrip soms bemoeilijken. De onpersoonlijke stijl ten slotte wordt ook wel als “abstract”, “objectief”, “neutraal”, “formeel” of “zakelijk” benoemd (Goosen & Schoordijk 2017). Een wetenschappelijke stijl is geen verhalende, persoonlijke of essayistische stijl (De Wachter et al. 2021) en laat niet de auteur maar wel de feiten op de voorgrond treden. Gekleurd taalgebruik, eigen meningen en niet-verifieerbare informatie zijn bijgevolg niet gepast in een academische onderzoekspaper.

1.3 Schrijven en technologie

1.3.1 Technologische evoluties

Technologische innovaties veranderen zowel de aard van het schrijfproces als onze opvattingen over schrijven. Deze gedachte komt sterk naar voren in het boek *Digital writing technologies in higher education* onder redactie van Kruse, Rapp, Anson, Benetos, Cotos, Devitt en Shibani (2023). In grote lijnen onderscheiden zij drie fasen binnen de technologische schrijfontwikkelingen. Een eerste stap is de shift van analoog schrijven naar schrijven met *wordprocessors*. De tweede fase bestaat uit de komst van het internet en *networked platforms*. In de derde stap wordt ten slotte aandacht besteed aan *Natural Language Processing* (NLP).

Een gelijkaardige opdeling komt naar voren in het artikel *Digital support for academic writing* van Strobl et al. (2019) dat in sectie 1.3.3 als basis genomen zal worden. Ook Strobl et al. vermelden namelijk eerst de *digitization of writing* met de ingebruikname van *wordprocessors* rond 1980, vervolgens de verhoogde *interactivity* met de komst van het internet en tot slot de *automated text evaluation* ondersteund door onder andere corpusanalyses, *natural language processing* (NLP) en *Latent Semantic Analysis* (LSA) (Strobl et al. 2019: 34).

Het is nuttig om dit groter kader in het achterhoofd te houden en te beseffen dat recente ontwikkelingen zich verhouden tot eerdere ontwikkelingen die ook op een bepaald moment als nieuw en revolutionair geïntroduceerd werden. Vandaag lijkt het misschien evident om schrijfoopdrachten op een laptop uit te typen ondersteund door de automatisch ingebouwde spellingscorrectors, online schrijfhulpen en synoniemendatabanken maar ook deze ontwikkelingen zijn nog recent.

In de volgende secties wordt een kort overzicht geboden van de conceptuele gevolgen van deze technologische evoluties (1.3.2), de bestaande classificaties voor elektronische schrijfondersteuning (1.2.3) en vervolgens ingezoomd op Generatieve AI en ChatGPT (1.3.4, 1.3.5).

1.3.2 Conceptuele evoluties

In het artikel *Writing processes in the digital age* (2023), schetst Cummings hoe theorieën rond schrijfprocessen afhangen van technologische ontwikkelingen en ruimere culturele contexten. Ten tijde van het analoge schrijfproces heerste er een paradigma van sequenties met aandacht voor het eindproduct, de lineaire connecties en het polijsten van een tekst (zie Rohman in 1.1.1). Door de komst van *wordprocessors*, verschoof de paradigmatische aandacht naar de cognitieve processen (zoals beschreven in sectie 1.1.1) die hun blik vooral op revisie, recursiviteit en reflectie richtten. Het wereldwijde web wendde de denkkaders in een sociale richting met een theoretische focus op collaboratie, doelpubliek en participatie tijdens het schrijven (aansluitend bij socio-culturele modellen zoals in sectie 1.1.3). *Software as a Service* (SaaS) en *Content Management Systems* (CMS) passen ten slotte binnen een workflow paradigma en kijken naar schrijven als een *networked ecology* van activiteiten, ruimtes en tools (Cummings 2023: 490). Onder deze laatste groep vallen bijvoorbeeld annotatiesystemen, plagiaatcheckers, en online platformen zoals OneDrive en Google Documenten.

1.3.3 Elektronische schrijfondersteuning

Elektronische schrijfondersteuning in het hoger onderwijs kan de vorm aannemen van computergestuurde tools voor studenten en docenten, toegankelijk via het web of na het downloaden van software. Deze sectie bouwt verder op een classificatie van dergelijke tools opgesteld door Allen et al. (2016) en vervolledigd door Strobl et al. (2019). De classificatie onderscheidt vier soorten tools: *Automated Essay Scoring* (AES), *Automated Writing Evaluation* (AWE), *Intelligent Tutor Systems* (ITS) en *Interactive Writing Platforms* (IWP)

Een *Automated Essay Scoring* (AES) toepassing evalueert de inhoud, structuur of kwaliteit van een geschreven tekst maar biedt geen verdere feedback of instructies (Allen et al. 2016: 317). Er wordt een cijfer of score toegekend aan een bepaalde tekst door de taalkundige kenmerken (zoals gemiddelde

zinslengte, lexicale diversiteit, referentiële cohesie) van deze tekst te vergelijken met verwachte patronen (Allen et al. 2016: 317). Steunend op diverse technieken en voorbeeldcorpora, biedt *AES* een alternatief voor menselijke correctie en verlicht het de verbeterlast van leerkrachten. Voorbeelden zijn *E-rater*, *Writing Pal* en *IntelliMetric*.

Aan de hand van een *Automated Writing Evaluation (AWE)* kan een gebruiker wel feedback ontvangen op een ingediende tekst. Deze systemen maken gebruik van scores berekend met *AES* maar geven ook formatieve feedback mee bij de score. Op die manier kan een student snel en persoonlijk bijgestuurd worden. Binnen deze categorie, leggen bepaalde systemen de nadruk op grammaticale fouten, woorden en zinnen terwijl andere systemen ook de globale structuur van de hele tekst beoordelen. Allen et al. (2016) onderscheiden daarom *low-level feedback* van *high-level feedback* terwijl Strobl et al. (2019) opteren voor de termen *micro-AEW* en *macro-AEW* en deze tweedeling mee in hun classificatie opnemen. Uit een systematisch overzicht van studies rond de bekende Engelstalige spelers zoals *Criterion*, *Grammarly* en *Pigeon* maar ook minder bekende AWE-systemen, blijkt dat deze categorie de laatste jaren vaak onderzocht wordt (Ding et al. 2023; Shadiev et al. 2023).

Intelligent Tutoring Systems (ITS) nemen de feedback nog een stap verder en voorzien geïndividualiseerde instructies aangepast aan de noden van de student in interactieve leeromgevingen. Het aanleren van schrijfstrategieën, aanbieden van extra oefenmogelijkheden en verhogen van motivatie aan de hand van games of filmpjes zijn perspectieven die binnen *ITS* verder uitgewerkt worden volgens Allan et al (2019: 322-324). Een student die op een bepaald onderdeel of bepaalde vaardigheid slecht scoort, zal automatisch meer oefeningen of uitleg hierover voorgeschoteld krijgen en die individualisering onderscheidt deze derde groep van de eerder vermelde categorieën.

De *Interactive Writing Platforms* zijn ten slotte een categorie die Strobl et al (2019) toevoegden aan de drievoudige classificatie van Allen et al (2016) op basis van hun vergelijkend onderzoek van 44 tools. Deze vierde categorie beslaat de schrijfhulpen die focussen op “planning, pre-writing and drafting processes” (Strobl et al. 2019: 44). Het zijn eveneens interactieve en geïndividualiseerde leeromgevingen maar ze verschillen van *ITS* omdat ze geen ingevoerde tekst maar enkel pedagogische ondersteuning bieden (Strobl et al. 2019: 46). In het onderzoek van Strobl et al., behoorden 16 van de 44 bestudeerde tools tot deze categorie terwijl maar één tool als een *ITS* beschouwd werd.

1.3.4 Generatieve Artificiële Intelligentie

Generatieve Artificiële Intelligentie (GenAI) is geen nieuw fenomeen maar heeft sinds november 2022 wereldwijd extra aandacht gekregen met de komst van ChatGPT. Deze chatbot, voluit *Chat Generative Pre-trained Transformer*, werd ontwikkeld door het Amerikaanse startup bedrijf OpenAI en is in staat om op interactieve manier vragen te beantwoorden via chat. Steunend op *Large Language Models* (LLM), functioneert het aan de hand van probabilistische kansen dat talige bouwstenen elkaar zullen opvolgen en creëert het zo een natuurlijk aandoende tekst. Het heeft een *deep-learning neural network* en een *transformer architecture*. Tijdens het trainen van ChatGPT, werd gebruik gemaakt van een grote hoeveelheid tekstuele data, waarvan de oorsprong onbekend is, alsook van supervisie van menselijke beoordelaars die het model bijstuurden (*reinforced learning with human feedback*) (Ray 2023; Kulkarni 2023).

Verschillende versies volgen elkaar snel op maar tijdens het schrijven van deze thesis is gebruik gemaakt van ChatGPT3.5, de versie die in november 2022 uitgebracht werd en gratis toegankelijk was in maart 2024 voor wie een account aanmaakte. Een nieuwere versie, ChatGPT4.0 met uitgebreidere functies, zoals de inclusie van foto- en spraaktechnologieën vereiste op het moment van de uitgevoerde studies van dit werk nog een betalend lidmaatschap. Vergelijkbare generatieve AI-mechanismen zijn onder andere Bing AI (Microsoft), Bard AI of Gemini (Google), Claude (Anthropic) en Copilot (Microsoft).

Academici en studenten merkten snel op dat ze ook voor schrijfoopdrachten gebruik konden maken van generatieve AI. Hoewel er zowel tegenstanders als aanhangers zijn, kunnen algemene tendensen afgeleid worden uit de grote hoeveelheid opiniestukken en artikels over ChatGPT als schrijfhulp.

De grote voordelen zijn vooral de mogelijkheid tot het voortdurend aanvoeren van interactieve en persoonlijke feedback en de afwezigheid van de angst om beoordeeld te worden (Beck & Levine 2023; Su et al 2023). In tegenstelling tot klasgenoten of leerkrachten, biedt ChatGPT advies zonder de schrijver te veroordelen en is het advies zelf 24/7 beschikbaar voor iedereen met toegang tot het internet (Beck & Levine 2023: 67).

Mogelijke nadelen en beperkingen zijn zowel ethisch als technisch van aard. Op een technisch vlak is er een risico op *AI hallucination* of het aanvoeren van informatie die plausibel klinkt maar incorrect of incoherent is (Hwang et al. 2023: 953-954). Die hallucinatie uitte zich oorspronkelijk ook in het geven van foutieve referenties voor onbestaande wetenschappelijke publicaties (zie bijvoorbeeld Day 2023).

Recenter onderzoek uit juni 2023 toont echter dat ChatGPT geen referenties meer levert en de gebruiker aanraadt om zelf wetenschappelijke databases te doorzoeken aan de hand van sleutelwoorden (McGowen et al. 2023; Anghelescu et al. 2023). Verder is de chatbot ook gevoelig voor kleine aanpassingen in de *prompts* en herhaalt hij vaak bepaalde waarschuwingen over zijn begrensde capaciteit die sommige bekritisieren als “excessively verbose” of “unnecessary statements” (OpenAI 2022; Barrot 2023: 4). De ethische beperkingen zijn op hun beurt vierdelig. In de eerste plaats kan de gegenereerde output schadelijke stereotypen en vooroordelen uit de trainingsdata reproduceren, hoewel OpenAI dit probeert te voorkomen aan de hand van een modererende tool (Barrot 2023: 9). Ten tweede, zijn er belangrijke vraagstukken over academische integriteit en mogelijk leerverlies bij studenten (Karakose 2023). Wat als schrijvers gegenereerde output van AI zomaar overnemen in hun tekst en wie is de intellectuele eigenaar van die output dan? Moet ChatGPT wel of niet als coauteur opgenomen worden in wetenschappelijke publicaties? Is er daarvoor al sprake van plagiaat aangezien ChatGPT zelf niet transparant genoeg is over de bronnen waarop zijn antwoorden gebaseerd zijn? Ten derde is er onduidelijkheid over de privacy en gegevensbewaring van ChatGPT wat zeker voor gevoelige informatie zoals patiëntendossiers in een medische context bezwaren oproept (Hwang et al. 2023). Ten slotte kost het trainen van *deep learning* modellen grote hoeveelheden energie en rijzen er ecologische bezwaren (Berthelot et al. 2024)

Voor welke doelen binnen het schrijfproces zou generatieve AI ingezet kunnen worden? De theoretische antwoorden zijn veelvoudig. In hun artikel *Collaborating with ChatGPT in argumentative classrooms* (2023), geven Su et al. toepassingen voor vier verschillende fasen binnen het schrijven van een argumentatief essay. In een voorbereidende fase zou ChatGPT mee ideeën kunnen genereren en feedback geven op de outline en organisatie van argumenten. In een redactiefase zou ChatGPT inhoudelijke feedback op een eerste kladversie kunnen geven of alternatieve perspectieven suggereren. De auteurs vermelden hier het belang van goed geformuleerde *prompts* aangezien de tool slechts functioneert als een deskundige collega die zelf geen initiatief neemt of kritisch nadenkt. In de derde fase, de proeflezingsfase, kan ChatGPT zowel individuele zinnen als een hele tekst nakijken op grammaticale fouten, accurate verwoording, precisie en de academische stijl. In deze fase is kritisch inzicht eveneens vereist om te beoordelen of de suggesties de tekst al dan niet verbeteren maar zou de tool volgens hen ook als query tool kunnen dienen om bepaalde collocaties op te zoeken. De vierde fase of reflectiefase is verwijst de mogelijkheid om een chat op te slaan en aan de hand van dit “archief” als student nieuwe kennis te verwerven door veelvuldige oefening.

Andere studies vermelden de mogelijkheid om automatisch een score te laten toekennen aan een tekst (Barrot et al. 2023: 4), het overkomen van het *blank page syndrom* (Hwang 2023: 953), het

genereren van verschillende kladversies of revisies om die daarna te vergelijken (Beck et al 2023), het samenvatten van een tekst voor een abstract of conclusie (Mahyoob 2023: 283) of het leveren van bepaalde *templates* binnen een bepaald domein (Warschauer et al 2023).

Een exploratieve studie naar de ervaringen van een Chinese doctoraatsstudente die Engels als niet-moedertaal had stelde een duidelijke evolutie vast in haar omgang met ChatGPT als schrijfassistent (Jacob et al. 2023). De doctoraatsstudente gaf oorspronkelijk aan regelmatig gebruik te maken van ChatGPT om ideeën te genereren, verkennend onderzoek naar een bepaald onderwerp te doen en tekst te laten produceren. Ze bleef hierbij kritisch en gebruikte de output eerder als een skelet waarvan ze de informatie en bronnen kritisch checkte en aanvulde. In een follow-up gesprek enkele maanden later bleek haar enthousiasme echter getemperd en gaf ze aan bepaalde patronen te herkennen, de ideeën uit de brainstormfase niet echt creatief te vinden en vaste signaalwoorden en collocaties vaak te zien terugkomen. Daarom verkoos ze de menselijke assistentie van haar collega's en gebruikte ze de chatbot vooral om concrete taalsuggesties te geven en de syntactische variatie te verhogen.

1.3.5 KU Leuven en generatieve Artificiële Intelligentie

De KU Leuven werkte richtlijnen uit rond het verantwoord gebruik van generatieve Artificiële Intelligentie op een universiteitsbreed niveau, over de faculteiten heen. Daarbij staat aangegeven dat “een docent, een opleiding of je faculteit [kan] beslissen om een bepaald gebruik van GenAI niet toe te laten” (KU Leuven, 2024). Het is in het algemeen niet toegelaten om materiaal gegenereerd door GenAI te kopiëren zonder bronvermelding of GenAI zodanig te gebruiken dat de docent de vaardigheden van de student niet meer kan evalueren. Studenten zijn zelf volledig verantwoordelijk voor wat ze indienen en moeten ondubbelzinnig hun verworven competenties tonen. Wat wel toegelaten is, is GenAI inschakelen als zoekrobot of als taalassistent:

“GenAI gebruiken als taalassistent voor het herwerken of verbeteren van zelf geschreven teksten én het model geen nieuwe inhoud toevoegt. Dan is het gebruik gelijk aan de spelling- en grammarcheckers die we vandaag ook al kennen en hoeft je dat gebruik niet expliciet te vermelden. Ook hier blijf je er steeds op letten dat het gebruik van GenAI een evaluatie door de examinerator van de door jou bereikte leerdoelen niet in de weg staat.” (KU Leuven, 2024)

De richtlijnen benadrukken de noodzaak tot overleg met betrokken docenten en promotoren en tonen hoe studenten hun eigenaarschap kunnen behouden en het onderwijs- en examenreglement kunnen respecteren.

Deel 2: Studie (A) vragenlijst

Deze studie steunt zowel op kwalitatief als op kwantitatief onderzoek en bestaat enerzijds uit een enquête (studie A) afgenomen bij studenten aan de KU Leuven en anderzijds uit een schrijfpdracht gecombineerd met *think aloud protocols* (studie B, zie deel 3). Deel 2 en 3 beginnen met een presentatie van de participanten uit de steekproeven (2.1; 3.1), bespreken dan de methode en het materiaal (2.2; 3.2) en sluiten af met de analyse en resultaten (2.3; 3.3).

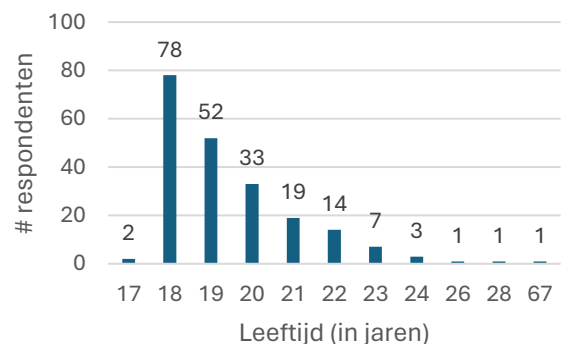
2.1. Respondenten

De eerste studie bestaat uit een vragenlijst afgenomen bij 258 studenten van de associatie KU Leuven. Zoals weergegeven in Tabel 1 geven 128 respondenten aan zichzelf als man te identificeren, 126 respondenten respectievelijk als vrouw, 3 respondenten als 'anders' en kiest 1 respondent om de vraag niet te beantwoorden. Verder hebben 245 van de 258 respondenten een Belgische nationaliteit, 9 respondenten een Nederlandse nationaliteit en zijn er ten slotte 1 Franse, 1 Indonesische en 1 Duits-Spaanse respondent. De moedertaal is Nederlands bij 249 deelnemers, gevolgd door Frans bij 6 deelnemers, Spaans bij 2 deelnemers en Baha Indonesia bij 1 deelnemer. De gemiddelde leeftijd van de deelnemers bedraagt 19,65 jaar met een modus van 18 jaar en een mediaan van 19 jaar. De leeftijdsverdeling omvat een bereik van 17 tot 67 jaar (zie Figuur 5).

Tabel 1: Geslacht, nationaliteit en moedertaal respondenten (n = 258)

Ik identificeer me als...	
Man	128 (49,42%)
Vrouw	126 (48,83%)
Anders	3 (1,16%)
Zeg ik liever niet	1 (0,39%)
Mijn nationaliteit is	
Belg	245 (94,96%)
Nederlander	126 (3,49%)
Frans	1 (0,39%)
Indonesisch	1 (0,39%)
Duits-Spaans	1 (0,39%)
Mijn moedertaal is	
Nederlands	249 (96,51%)
Frans	6 (2,33%)
Spaans	2 (0,78%)
Baha Indonesia	1 (0,39%)

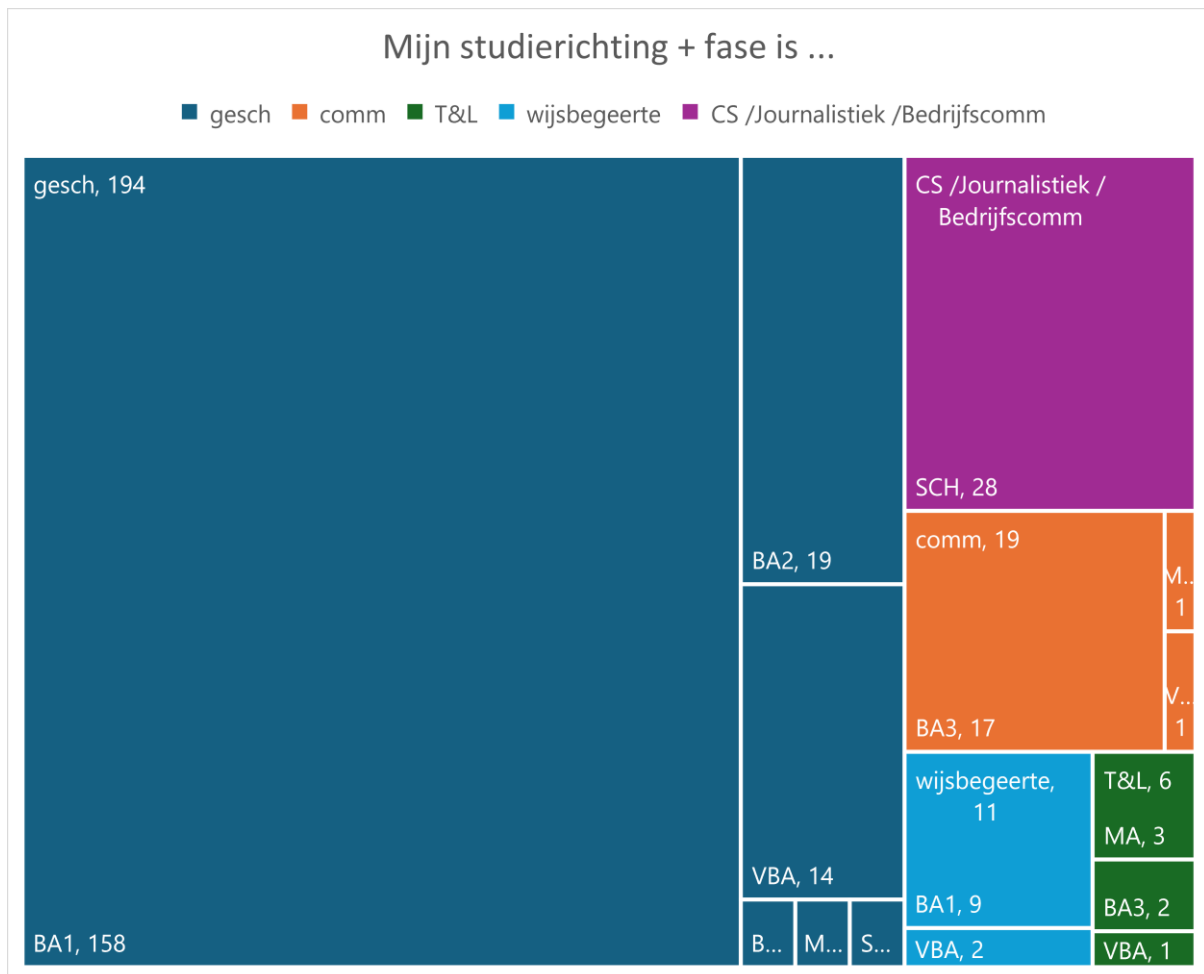
Figuur 5: Leeftijd respondenten (n = 258)



De spreiding over de studierichtingen en fasen is weergegeven in Figuur 6. De meerderheid van de deelnemers volgt de bacheloropleiding geschiedenis (194 studenten of 75,19%), gevolgd door het schakelprogramma voor een master in de journalistiek, bedrijfscommunicatie of culturele studies (28 studenten of 10,85%), de bachelor in de communicatiewetenschappen (19 studenten of 7,63%), de bachelor in de wijsbegeerte met optie

geschiedenis (11 studenten of 4,26%) en ten slotte de bachelor of master in de taal- en letterkunde (6 studenten of 2,32%).

Figuur 6: Studierichting en fase respondenten (n = 258)



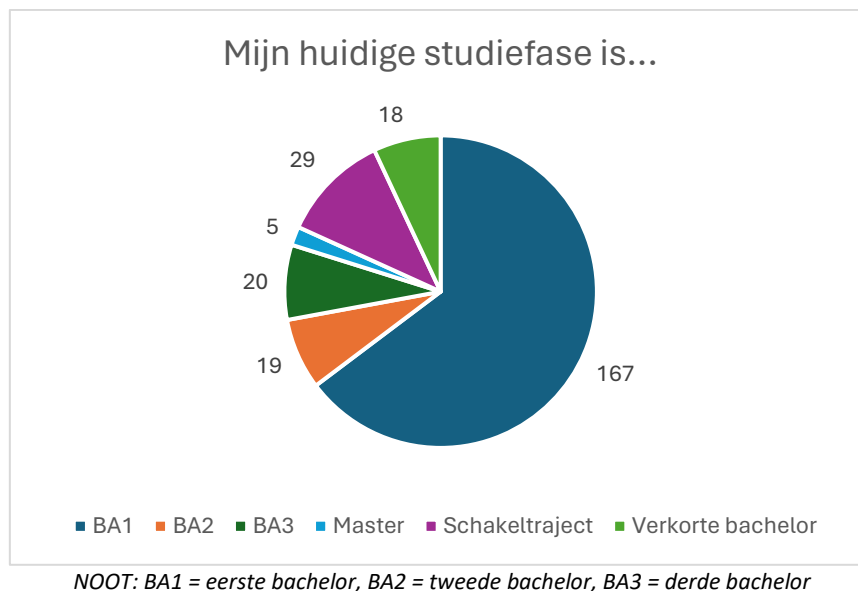
NOOT: BA1 = eerste bachelor, BA2 = tweede bachelor, BA3 = derde bachelor, MA = master, SCH = schakeltraject en VBA = verkorte bachelor // gesch = geschiedenis, comm = communicatiewetenschappen, T&L = taal- en letterkunde, CS = culturele studies, bedrijfscomm = bedrijfscommunicatie, wijsbeg = wijsbegeerte

De groep geschiedenisstudenten bestaat grotendeels uit eerste bachelorstudenten, die 81,44% (n = 158) van die groep vertegenwoordigen, gevolgd door 9,79% (n = 19) tweede bachelorstudenten, 7,2% (n = 14) verkorte bachelorstudenten, en 1,55% (n = 3) derde bachelor- of masterstudenten. De studenten journalistiek, culturele studies of bedrijfscommunicatie hebben allemaal een vooropleiding aan een hogeschool achter de rug en bevinden zich nu in het eerste jaar van hun schakelprogramma voor een universitaire master. De 19 studenten communicatiewetenschappen zijn vooral derde bachelorstudenten (89,47%) aangevuld door een masterstudent en schakelstudent. Bij wijsbegeerte zitten 9 studenten (81,81%) in hun eerste bachelor en de overige 2 studenten in een verkorte bachelor na het behalen van een academische master. De groep taal- en letterkundestudenten is complexer en kleiner omdat die gericht samengesteld is om meer ervaren schrijvers voor studie B te betrekken. Deze 6 studenten zijn gespreid over de derde bachelor (2

studenten), het eerste masterjaar (2 studenten), een tweede educatief masterjaar (1 student) en een verkorte bachelor na het behalen van een masterdiploma binnen de taal- en letterkunde (1 student).

Figuur 7 visualiseert de steekproef die in totaal bestaat uit 167 eerstebachelorstudenten (64,71%), 19 tweedebachelorstudenten (7,36%), 20 derdebachelorstudenten (7,75%), 29 schakelstudenten (11,24%), 5 masterstudenten (1,93%), en 18 verkorte bachelor studenten (6,98%). De analyse maakt een verschil tussen de eerste- en tweedebachelorstudenten enerzijds en de andere groepen anderzijds om op zoek te gaan naar contrasten tussen onervaren (*novice*) en meer ervaren (*advanced*) academische schrijvers. De grens is getrokken tussen studenten die al een bachelor- of masterproef aan het schrijven zijn of geschreven hebben en zij die dit nog niet gedaan hebben. Dit geeft een totaal van 188 *novice* en 60 *advanced writers*. Ter navolging van Bereiter & Scardamalia (zie 1.1.1.2) gebruikt deze studie de Engelstalige terminologie.

Figuur 7: Studiefase respondenten (n = 258)



2.2 Materiaal en methode

De vragenlijst is opgebouwd met de survey tool *Qualtrics* en is volledig anoniem. Ethische goedkeuring is verleend door SMEC onder goedkeuringsnummer G-204-7581. Bijlage 1 bevat een schriftelijke weergave van de enquête.

Tijdens bepaalde colleges rond academische schrijfvaardigheid werd dit masteronderzoek kort toegelicht en werd de link naar de enquête gedeeld via Blackboard. Het betreft de vakken F0LC1A “initiatie in de historische praktijk” voor eerstebachelors geschiedenis en wijsbegeerte met optie geschiedenis, SOC18B “schrijven en presenteren voor mediaprofessionals” voor derdebachelors communicatiewetenschappen en het schakelprogramma voor de master journalistiek en ten slotte

F0BQ9A “academisch Nederlands” voor schakelstudenten binnen de master in de culturele studies, bedrijfscommunicatie of journalistiek. De vragenlijst werd grotendeels ingevuld tussen 28 februari en 6 maart 2024 met uitzondering van sommige deelnemers van studie B die hem later invulden. In de vragenlijst zijn twee controlevragen toegevoegd om na te gaan of de enquête aandachtig ingevuld werd (stelling 1.5 “duid hier de optie ‘soms’ aan” en stelling 3.7 “duid hier de optie ‘altijd’ aan”). Drie respondenten die hierop een fout antwoord gegeven hebben, zijn volledig geëlimineerd uit de verwerking en interpretatie van de enquête.

Deelname aan de vragenlijst is niet verplicht en heeft geen consequenties voor de beoordeling van het vak voor de studenten. De gecontroleerde setting heeft als voordeel een zekere controle over de respondenten en het vermijden van de non-respons die vaak bij online enquêtes optreedt (Baarda 2015). Door de vragenlijst zowel in de eerste bachelor als bij een groep derdebachelor- en schakelstudenten van een master binnen de sociale wetenschappen of letteren af te nemen, zijn er verschillende profielen. Van studenten uit een eerste jaar wordt verwacht dat zij nog weinig ervaring hebben met het academisch discours en er misschien andere strategieën op nahouden dan studenten uit een schakelprogramma die al een opleiding en bachelorproef achter de rug hebben. Deze diversificatie past binnen het exploratief kader van de masterproef.

De vragenlijst zelf bestaat uit drie luiken voorafgegaan door persoonsgegevens (leeftijd, nationaliteit, moedertaal, geslacht, opleiding). Een eerste luik bevraagt algemene ervaringen van de studenten met generatieve artificiële intelligentie: welke vormen ze kennen en hoe vaak ze aangeven het te gebruiken. Een tweede luik onderzoekt in welke fasen van het schrijfproces en voor welke concrete doelen studenten gebruikmaken van generatieve AI. In sterke mate steunend op lopend onderzoek van Lieve De Wachter (KU Leuven), Renske Bouwer (Universiteit Utrecht) en Janneke van der Loo (Universiteit Tilburg) toetsen stellingen met een 5-punts Likertschaal gegroepeerd in matrixtabellen concrete situaties af. De stellingen zijn opgedeeld in drie delen die overeenkomen met een *pre-writing* of planningsfase, een *writing* of vertaalfase en *post-writing* of reflectiefase. Deze opdeling steunt op terminologie uit het sequentieel model van Rohman en op de inzichten van Hayes (1996). Zo is de schrijffase zelf ruim opgevat en komen vaardigheden zoals ‘nalezen’ en ‘corrigeren’ ook binnen die fase aan bod aangezien deze passen binnen de productiecomponent. De *post-writing* fase heeft in dit onderzoek bijgevolg betrekking op eindevaluaties en nabeschouwingen en het eventueel inkorten of samenvatten van een schrijfproduct en niet zozeer op het opsporen van schrijffouten of verbeteren van stijl en samenhang. Het derde luik stelt ten slotte de perceptie van generatieve AI op de proef en toetst ook met stellingen en een 5-punts Likertschaal mogelijke bezwaren uit de literatuurstudie af bij studenten.

2.3 Analyse en resultaten

Deze sectie presenteert de data uit de afgenomen bevraging en is drieledig opgebouwd. Een eerste deel (2.3.1) bespreekt de algemene omgang met generatieve AI of meer bepaald de gebruikte tools en hun frequentie. Een tweede deel (2.3.2) specificeert mogelijke toepassingen opgedeeld in de drie schrijffasen. Het derde luik (2.3.3) toetst de houding tegenover generatieve AI af aan de hand van stellingen rond mogelijke bezorgdheden. Hierbij wordt telkens een onderscheid gemaakt tussen *novice* (n = 188) en *advanced writers* (n = 60) naargelang de studenten al dan niet al een bachelorpaper geschreven hebben of aan het schrijven zijn. Bijlage 3 bevat bovendien alle resultaten uit de bevraging.

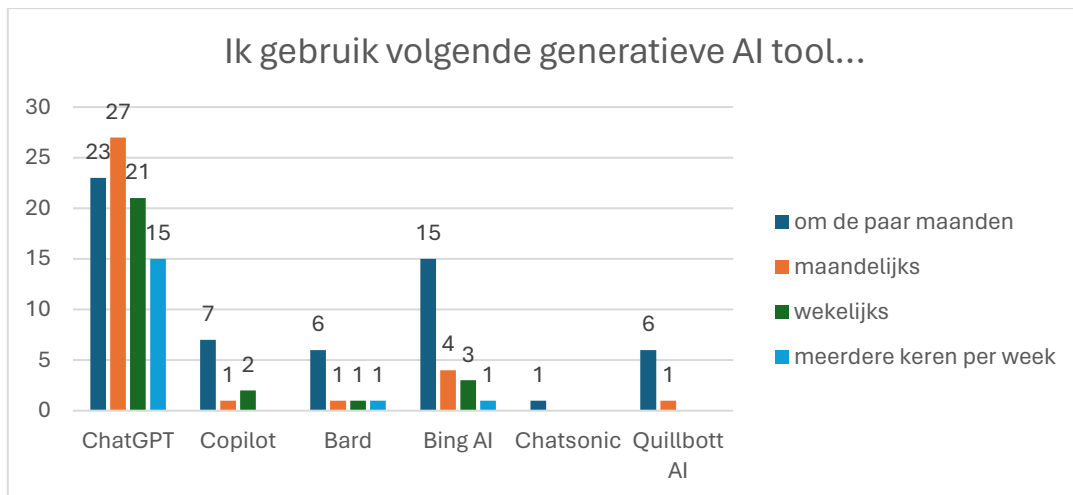
2.3.1 Omgang met generatieve AI: algemeen

De bevraging peilt ten eerste naar de bekendheid van generatieve AI-tools bij de respondenten aan de hand van een lijst waarin ze de gekende opties kunnen aanduiden en zelf aanvullingen kunnen doen. In totaal hebben 255 van de 258 respondenten (98,83%) al van één van de gegeven opties gehoord en zijn slechts drie deelnemers volledig onbekend met GenAI. De studenten zijn voornamelijk vertrouwd met ChatGPT (255 studenten of 98,83%) maar ook Bing AI (50 studenten of 19,37%), Google Bard (27 studenten of 10,46%), Quilbott AI (25 studenten of 9,69%), Microsoft Copilot (17 studenten of 6,59%) en Jasper AI (6 studenten of 2,32%) genieten enige bekendheid. De tools Ecosia AI, Remini en Midjourney worden elks slechts eenmaal aangehaald.

Ten tweede onderzoekt de enquête hoe vaak de deelnemers bepaalde generatieve AI-tools gebruiken in hun dagelijks leven (zie Figuur 8). Ook bij deze vraag blijkt ChatGPT met 214 gebruikers (82,95%) de overige tools veruit te overstijgen. ChatGPT wordt door 64 studenten om de paar maanden gebruikt, door 66 maandelijks, 60 wekelijks en ten slotte door 24 fervente gebruikers meerdere keren per week. De overige tools worden minder gebruikt. Voor Bing AI zijn er 15 occasionele, 4 maandelijks en 3 wekelijkse gebruikers. Bij Microsoft Copilot geven 7 studenten aan het om de paar maanden te gebruiken, 1 student maandelijks en 2 wekelijks. Google Bard/Gemini en Quilbott hebben 6 studenten die er occasioneel gebruik van maken en nog 4 studenten wat vaker.

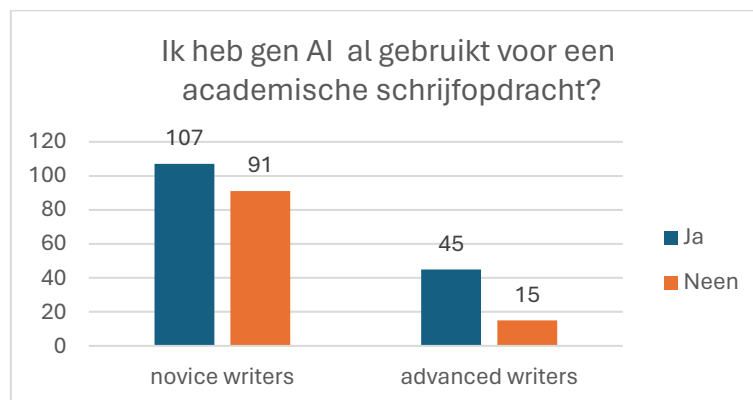
Participanten kunnen ook zelf aangeven welke andere tools ze gebruiken die nog niet in de lijst voorkomen en geven aan Chatpdf, Notion AI, Remini, Midjourney en Pop AI te gebruiken met frequenties wisselend tussen maandelijks (Chatpdf Midjourney, Pop AI, Notion AI) tot wekelijks (Remini). Pop AI en Chatpdf zijn tools om PDF- bestanden mee samen te vatten en Notion AI is actief binnen de workspace Notion terwijl Midjourney en Remini foto's genereren of bewerken.

Figuur 8: Gerapporteerde frequentie gebruik GenAI (n = 258)



Ten derde wordt gevraagd of ze al eens generatieve AI voor een academische schrijfpdracht hebben ingezet. Op deze vraag antwoordt een aanzienlijk deel bestaande uit 106 studenten of 41,08% dat ze dat nog nooit gedaan hebben. Die 106 studenten zijn verdeeld tussen de opleidingen geschiedenis (n = 90), wijsbegeerte (n = 6), journalistiek/ culturele studies/ bedrijfscommunicatie (n = 7), communicatiewetenschappen (n = 2) en taal- en letterkunde (n = 1). Opvallend is dat procentueel gezien bijna de helft (46,46%) van de *novice writers* nooit generatieve AI gebruikt terwijl dit bij de *advanced writers* slechts 25% van de groep beslaat (zie Figuur 9).

Figuur 9: Gerapporteerd gebruik GenAI in academische schrijfpdrachten (n = 258)



De respondenten die GenAI vermijden voor hun schrijfpdrachten mogen hun keuze toelichten. De meest voorkomende redenen om het niet te gebruiken is de angst voor plagiaat of fraude (21 keer vermeld), gevolgd door een hoog vertrouwen in de eigen schrijfvaardigheid en het zelf willen kunnen leren schrijven (13 keer vermeld). Een aanzienlijk aantal vermeldt ook dat GenAI niet voldoet aan hun vereisten (12), ze er geen nood toe voelen (11) of er niet voldoende mee vertrouwd zijn om het efficiënt in te zetten (9). Verder verwijzen studenten ook naar het ontbreken van de bronvermelding (9) of naar de regelgeving en het feit dat het niet toegestaan is (7). Gevoelsmatig geven de deelnemers aan dat ze GenAI beangstigend vinden (6) of er een gevoel van oneerlijkheid mee verbinden (3).

De 152 overige studenten die wel generatieve AI toepassen voor hun papers, geven aan voor hun schrijftaken voornamelijk gebruik te maken van ChatGPT (150 studenten) maar ook in mindere mate van Bing AI (29 studenten), Google Bard (21 studenten) en Quillbot (22 studenten).

2.3.2 Omgang met generatieve AI: per schrijffase

Tabellen 2, 3 en 4 geven aan voor welke handelingen die 152 studenten generatieve AI beweren te gebruiken. De scores zijn afkomstig van een 5-punt Likertschaal en verwijzen naar de frequentie waarop de participanten de voorgestelde handeling uitvoeren waarbij 0 = nooit, 1 = zelden, 2 = soms, 3 = vaak en 4 = altijd. Voor elke stelling is een opdeling gemaakt tussen de *advanced writers* ($n = 45$) en *novice writers* ($n = 107$) en is per groep het gemiddelde (M), de standaarddeviatie (SD), mediaan (Mdn) en modus (Mo) gegeven.

De nulhypothese voor de gegevens is dat er geen verschil is tussen ervaren en onervaren schrijvers wat betreft hun omgang met generatieve AI. Om dit te testen, is de significantie van het verschil of de p -waarde statistisch berekend aan de hand van een Mann-Whitney U test (W). Dat is een niet-parametrische test geschikt om de medianen van twee groepen te vergelijken bij schendingen van de normaliteitsassumptie (Levshina 2015). Voorafgaand aan de keuze voor de Mann-Whitney U-test, is de normaliteit van de data geëvalueerd. Een visuele inspectie van Q-Q plots en een Shapiro-Wilk test leidt tot de bevinding dat de dataset niet normaal verdeeld is en ondersteunt de keuze voor de niet-parametrische test. Een verschil tussen de *novice* en *advanced writers* is significant wanneer de p -waarde kleiner is dan .05 (Van de Mierop 2020).

Bovendien is de effectgrootte bepaald aan de hand van een punt-biseriële correlatiecoëfficiënt (r_{rb}), aangezien het een verband tussen een niet-normaal verdeelde dichotome variabele (schrijversgroep) en ordinale variabele (scores op een vijf puntsschaal) betreft. De correlatiecoëfficiënt is positief en hoog, wanneer hij dicht bij de maximumgrens van 1 ligt. Een correlatie tussen 0.1 en 0.3 wordt als klein beschouwd, een correlatie tussen 0.3 en 0.5 als middelmatig en een correlatie hoger dan 0.5 als hoog (Van De Mierop, 2020).

2.3.2.1 Voor het schrijven

In Tabel 2 zijn zeven stellingen terug te vinden over de *pre-writing* fase. Die stellingen staan in de enquête gegroepeerd in een matrixtabel met als titel “Ik gebruik generatieve AI (zoals ChatGPT) voor het schrijven om ...”. Een overzicht van het gemiddelde (M), de standaardafwijking (SD), mediaan (Mdn), modus (Mo), score van de Mann Whitney U-test (W) met p -waarde en effectgrootte (r_{rb}) bevinden zich in de tabel.

Tabel 2: Gebruik van GenAI voor het schrijven (n = 152)

Stelling	Groep	M	SD	Mdn	Mo	W	p	r _{rb}
1: Tips te vragen over hoe ik de opdracht best aanpak	Adv (n = 45)	1.356	1.151	1	0	2666.5	0.273	0.108
	Nov (n = 107)	1.140	1.177	1	0			
	Tot (n = 152)	1.204	1.170	1	0			
2: Te brainstormen over mijn onderwerp en inspiratie te 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			
3: Bronnen te zoeken of bronnen samen te vatten	Adv (n = 45)	0.889	1.027	1	0	2157.5	0.289	-0.104
	Nov (n = 107)	1.065	1.021	1	0			
	Tot (n = 152)	1.013	1.023	1	0			
4: Moeilijke concepten uit te leggen of te 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			
6: Een tekstplan te genereren	Adv (n = 45)	0.778	0.876	1	0	2811.5	0.065	0.168
	Nov (n = 107)	0.533	0.805	0	0			
	Tot (n = 152)	0.605	0.831	0	0			
7: Feedback te krijgen op een tekstplan dat ik opstelde	Adv (n = 45)	1.244	1.401	0	0	2636.5	0.320	0.095
	Nov (n = 107)	0.981	1.213	1	0			
	Tot (n = 152)	1.059	1.272	0	0			
8: Op gang te komen en niet met een leeg blad te moeten beginnen	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			

NOOT: Adv = *Advanced writer*, Nov = *Novice writers*, Tot = totale groep, M = gemiddelde, SD = standaarddeviatie, Mdn = mediaan, Mo = modus, W = Mann-Whitney U test, p = p-waarde, r_{rb} = biseriële correlatiecoëfficiënt.

Uit een vergelijking van de medianen valt op dat de *advanced writers* meer gebruikmaken van generatieve AI dan de *novice writers* in vier van de zeven situaties (brainstormen, concepten uitleggen, tekstplan genereren, op gang komen) en ze slechts in één situatie een lagere mediaan bekomen (feedback krijgen op een tekstplan). Statistisch onderzoek wijst uit dat het verschil tussen beide groepen slechts voor drie situaties significant is. Ervaren schrijvers doen in vergelijking met de onervaren schrijvers vaker beroep op generatieve AI om de start te vergemakkelijken (W = 3222, p < .001, r_{rb} = 0.338), te brainstormen (W = 3142, p = .002, r_{rb} = .305) en om moeilijke concepten uitgelegd te krijgen (W = 2940, p = .024, r_{rb} = 0.221). De effectgrootte is bij die eerste twee acties middelmatig en bij de laatstgenoemde laag.

Wat de totaalscores betreft, zijn het eveneens de brainstormfase (M = 2.026, md = 2) en de definiëring van moeilijke concepten (M = 2.112, md = 3) die globaal genomen het vaakst gerapporteerd zijn. Het genereren van een tekstplan (M = 0.605, md = 0) of feedback vragen op een tekstplan (M = 1.0596, md = 0) scoren voor de totaalgroep het laagst. De hele *pre-writing* fase samengenomen behaalt een gemiddelde score van 1.317 op 4 en een mediaan van 1 op 4.

In de suggestiebalk (“ik gebruik GenAI voor het schrijven ook nog voor andere doelen, namelijk”) geeft een participant nog aan de hulp van generatieve AI in te schakelen bij het bedenken van tussentitels, wat enigszins thuishoort bij de tweede stelling (brainstormen) waar diezelfde student ook “vaak” had ingevuld. Een andere student specificeert dat hij het gebruikt om “lange teksten in bulletpoints te laten samenvatten”, aansluitend bij stelling 3 die een “soms” kreeg van deze respondent.

2.3.2.2 Tijdens het schrijven

Tabel 3 toont de resultaten voor de 6 geanticipeerde strategieën tijdens het schrijven:

Tabel 3: Gebruik GenAI tijdens het schrijven (n = 152)

Stelling	Groep	M	SD	Mdn	Mo	W	p	r _{rb}
1: Mogelijke voorbeeldzinnen te genereren	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			
2: Een eerste versie van een tekst/alinea te genereren	Adv (n = 45)	1.400	1.053	1	1	3555.5	< .001	0.477
	Nov (n = 107)	0.533	0.839	0	0			
	Tot (n = 152)	0.789	0.988	0	0			
3: Een lijst academische bewoordingen/construc ties te genereren	Adv (n = 45)	1.067	1.116	1	0	2818.5	0.070	0.171
	Nov (n = 107)	0.738	1.003	0	0			
	Tot (n = 152)	0.836	1.45	0	0			
4: Een (aantal) zin(nen) die ik schreef te herformuleren	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			
5: De stijl van mijn tekst te verbeteren en academischer te maken	Adv (n = 45)	2.133	1.272	2	3	3157.5	0.002	0.312
	Nov (n = 107)	1.430	1.190	1	0			
	Tot (n = 152)	1.638	1.253	2	0			
6: De grammatica- en/of spelfouten uit mijn tekst te halen	Adv (n = 45)	1.733	1.437	2	0	2546	0.566	0.058
	Nov (n = 107)	1.589	1.394	2	0			
	Tot (n = 152)	1.632	1.403	2	0			

NOOT: Adv = *Advanced writer*, Nov = *Novice writers*, Tot = totale groep, M = gemiddelde, SD = standaarddeviatie, Mdn = mediaan, Mo = modus, W = Mann-Whitney U test, p = p-waarde, r_{rb} = biseriële correlatiecoëfficiënt.

De zes stellingen uit de *writing* fase gedragen zich enigszins gelijkaardig aan de *pre-writing* fase. Voor vijf van de zes stellingen ligt de mediaan van de ervaren schrijvers namelijk hoger dan die van de onervaren schrijvers. Vier van die stellingen vertonen ook een p-waarde kleiner dan .05 en een middelgrote effectgrootte. *Advanced writers* rapporteren vaker gebruik te maken van generatieve AI om mogelijke voorbeeldzinnen te genereren (W = 329.5, p < .001, r_{rb} = 0.369), een eerste versie te generen (W = 355.5, p < .001, r_{rb} = 0.477), zinnen te herformuleren (W = 332.5, p < .001, r_{rb} = 0.380) en de stijl academischer te maken (W = 3157.5, p = .002, r_{rb} = 0.312).

Ervaren schrijvers rapporteren verder frequenter academische bewoordingen/constructies te laten genereren en vaker grammatica- en spelfouten uit hun tekst te laten halen maar deze acties vertonen

geen significante verschillen met de *novice*-groep. De tweede actie heeft zelfs een gelijke mediaan en maar .144 verschil in gemiddeldes tussen beide groepen.

De hele *writing* fase levert een gemiddelde van 1.326 en een mediaan van 1.5 op. Hoewel er hier geen gemiddelde scores boven de 2 aanwezig zijn, liggen de trends globaal gezien wel hoger en is het laagst behaalde gemiddelde een 0.789 (eerste versie genereren) waar dat in de *pre-writing* fase nog een 0.605 (tekstplan genereren) bedraagt. De verbeter/herformuleerstellingen (stellingen 4-6) zijn opvallend vaker gerapporteerd dan die waarbij nieuwe inhoud gegenereerd wordt (stellingen 1-3). Verder geven 2 participanten in de suggestiebalk van de enquête nog aan dat ze ook synoniemen opzoeken met behulp van generatieve AI.

2.3.2.3 Na het schrijven

Ten derde toont Tabel 4 de resultaten voor de reflectieve *post-writing* fase of de antwoorden op de vraag “ik gebruik generatieve AI na het schrijven om...”:

Tabel 4: Gebruik GenAI na het schrijven (n = 152)

Stelling	Groep	M	SD	Mdn	Mo	W	p	r _{rb}
1: Na te kijken of mijn tekst de juiste inhoud overbrengt	Adv (n = 45)	0.822	1.072	0	0	2768.5	0.092	0.150
	Nov (n = 107)	0.533	0.894	0	0			
	Tot (n = 152)	0.618	0.956	0	0			
2: Een tekst die ik schreef in te korten of langer te maken	Adv (n = 45)	1.600	1.405	2	0	3046.5	0.006	0.625
	Nov (n = 107)	0.944	1.080	1	0			
	Tot (n = 152)	1.138	1.218	1	0			
3: Mijn tekst samen te vatten (bv. voor een abstract of conclusie)	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			
4: Mijn tekst te evalueren door een cijfer toe te kennen	Adv (n = 45)	0.156	0.424	0	0	2452	0.751	0.018
	Nov (n = 107)	0.168	0.558	0	0			
	Tot (n = 152)	0.164	0.520	0	0			
5: Mijn tekst te evalueren door goede/ zwakke punten te geven	Adv (n = 45)	0.356	0.857	0	0	2344	0.717	-0.026
	Nov (n = 107)	0.336	0.764	0	0			
	Tot (n = 152)	0.342	0.790	0	0			
6: Mijn chatlogboek te herbekijken en daar inspiratie uit op te doen	Adv (n = 45)	0.244	0.645	0	0	2240.5	0.339	-0.069
	Nov (n = 107)	0.355	0.743	0	0			
	Tot (n = 152)	0.322	0.715	0	0			

NOOT: Adv = *Advanced writer*, Nov = *Novice writers*, Tot = totale groep, M = gemiddelde, SD = standaarddeviatie, Mdn = mediaan, Mo = modus, W = Mann-Whitney U test, p = p-waarde, r_{rb} = biseriële correlatiecoëfficiënt.

Onmiddellijk valt op dat voor de stellingen uit de fase na het schrijven beduidend lagere score gerapporteerd zijn. Slechts twee van de zes stellingen halen een mediaan groter dan 0 en dit zijn opnieuw de enige stellingen met een significant verschil. De *advanced writers* schakelen vaker generatieve AI in om de lengte van hun schrijfproducten te beïnvloeden (W = 3046.5, p = .006,

$r_{rb} = 0.625$) of er een samenvatting van te genereren ($W = 3189.5$, $p < .001$, $r_{rb} = 0.325$). De effectgrootte is groot en middelgroot en beide verschillen zijn bijgevolg substantieel.

Voor de andere vier stellingen liggen de scores dicht bij elkaar en gedragen beide groepen zich gelijkaardig. De gemiddeldes liggen zeer laag voor het evalueren van een tekst en chatlogboek herbekijken ($M < 0.35$) en het nakijken of een tekst de juiste inhoud bevat ($M = 0.618$). Globaal gezien scoort de derde fase van het schrijfproces aanzienlijk lager dan de voorgaande fasen en ligt het gerapporteerd gebruik van generatieve AI op een totaal gemiddelde van .606 en mediaan van 0. In de open vraag van deze ronde van de enquête zijn geen verdere toepassingen gesuggereerd.

2.3.2.4 Tussentijdse conclusie: novice en advanced writers

In totaal vertonen negen van de achttien stellingen een significant verschil tussen de ervaren en onervaren schrijvers. Het gaat om de handelingen brainstormen, concepten definiëren, op gang komen, voorbeeldzinnen of een eerste versie genereren, zinnen herformuleren of de stijl laten verbeteren en ten slotte een tekst langer/korter maken of samenvatten. Voor alle negen toepassingen rapporteren de ervaren schrijvers een frequenter gebruik dan de onervaren schrijvers, wat ook blijkt uit de positieve effectgroottes. Die trend is grotendeels ook aanwezig bij de overige negen niet-significant verschillende stellingen met uitzondering van stelling zeven uit het *writing* luik (feedback krijgen op een tekstplan) waar de mediaan van de onervaren schrijvers hoger ligt. Voor stelling drie uit de *pre-writing* fase (bronnen zoeken of samenvatten) en stellingen vier en vijf uit de *post-writing* fase (tekst evalueren, chatlogboek bekijken) zijn het eveneens de onervaren schrijvers die een hoger gemiddelde bekomen dan de ervaren schrijvers maar hier lopen de medianen gelijk.

2.3.3 Houding tegenover generatieve AI

Het derde deel van de enquête bestaat uit negen stellingen met opnieuw een 5-punts Likertschaal waarbij respondenten kunnen aangeven in welke mate ze akkoord gaan (0 = niet akkoord, 1 = eerder niet akkoord, 2 = geen mening, 3 = eerder akkoord, 4 = akkoord).

In de enquête zijn sommige stellingen positief geformuleerd ten opzichte van generatieve AI (bijvoorbeeld “ik vind GenAI een handig hulpmiddel) en andere stellingen negatief (bijvoorbeeld “ik vrees dat GenAI op lange termijn leidt tot leerverlies”). Bij de verwerking en analyse zijn alle negatieve stellingen echter geconverteerd naar hun positieve tegenhanger om de resultaten te kunnen generaliseren. Hoe hoger het gemiddelde (M) en de mediaan (Mdn) van een stelling in tabel 5, hoe positiever de ingesteldheid tegenover GenAI.

Tabel 5: Houding ten opzichte van GenAI (n = 152)

Stelling	Groep	M	SD	Mdn	Mo	W	p	r _{rb}
1: Ik maak me bij GenAI geen zorgen over de privacy/gegevensbewaring	Adv (n = 45)	2.267	1.176	3	3	2568.5	0.504	0.067
	Nov (n = 107)	2.131	1.198	2	3			
	Tot (n = 152)	2.171	1.189	2	3			
2: Ik vind GenAI een handig hulpmiddel voor studenten	Adv (n = 45)	3.378	0.576	3	3	2581.5	0.433	0.072
	Nov (n = 107)	3.252	0.715	3	3			
	Tot (n = 152)	3.289	0.677	3	3			
3: Ik vrees niet dat GenAI op lange termijn leidt tot leerverlies van studenten	Adv (n = 45)	1.733	1.053	1	1	2053	0.137	-0.147
	Nov (n = 107)	2.047	1.232	2	3			
	Tot (n = 152)	1.954	1.187	2	3			
4: Het is duidelijk hoe ik GenAI mag gebruiken voor academische opdrachten	Adv (n = 45)	1.978	1.158	2	1	1643	0.001	-0.318
	Nov (n = 107)	2.664	1.205	3	3			
	Tot (n = 152)	2.461	1.228	3	3			
5: Het gebruiken van GenAI voelt niet als het plegen van plagiaat	Adv (n = 45)	2.200	1.254	3	3	2341.5	0.785	-0.027
	Nov (n = 107)	2.262	1.223	2	3			
	Tot (n = 152)	2.243	1.229	2	3			
6: Ik beschouw de antwoorden van GenAI als betrouwbaar	Adv (n = 45)	1.911	0.996	2	1	2456.5	0.837	0.020
	Nov (n = 107)	1.869	1.010	2	3			
	Tot (n = 152)	1.882	1.003	2	3			
7: Door GenAI te gebruiken krijg ik meer vertrouwen in mijn schrijfvaardigheden	Adv (n = 45)	2.067	1.116	2	1	2743.5	0.159	0.140
	Nov (n = 107)	1.766	1.060	2	1			
	Tot (n = 152)	1.855	1.082	2	1			
8: Ik kan mijn schrijftaken beter afronden met GenAI	Adv (n = 45)	2.822	0.886	3	3	3372.5	< .001	0.401
	Nov (n = 107)	2.065	1.075	2	3			
	Tot (n = 152)	2.289	1.077	3	3			
9: Raad vragen aan GenAI is minder eng dan aan een klasgenoot, ouder, vriend	Adv (n = 45)	2.178	1.386	2	3	2715.5	0.203	0.128
	Nov (n = 107)	1.879	1.294	2	3			
	Tot (n = 152)	1.967	1.324	2	3			

NOOT: Adv = *Advanced writer*, Nov = *Novice writers*, Tot = totale groep, M = gemiddelde, SD = standaarddeviatie, Mdn = mediaan, Mo = Modus, W = Mann-Whitney U test, p = p-waarde, r_{rb} = biseriële correlatiecoëfficiënt.

Het gemiddelde van alle stellingen bedraagt 2.235, de mediaan 2 en de modus 3. Over het algemeen nemen de respondenten een neutrale tot eerder positieve houding ten opzichte van GenAI aan. Bij drie van de negen stellingen is de mediaan van beide groepen gelijk aan 3. De studenten geloven bijgevolg in grotere mate dat GenAI een handig hulpmiddel is ($M = 3.289$, $SD = 0.677$, $Mdn = 3$), dat ze hun schrijftaken beter kunnen afronden met GenAI ($M = 2.289$, $SD = 1.077$, $Mdn = 3$) en dat het duidelijk is in welke mate ze GenAI mogen gebruiken in hun opdrachten ($M = 2.461$, $SD = 1.228$, $Mdn = 3$) al zijn de gemiddeldes voor die laatste twee stellingen wel kleiner dan 3. Geen enkele stelling heeft een totaalmediaan kleiner dan 2 en grote bezorgdheden rond bijvoorbeeld de privacy of het risico op plagiaat blijven dus ook uit. Alleen bij de ervaren schrijvers ligt de mediaan van de derde stelling op 1 ($M = 1.733$, $SD = 1.053$, $Mdn = 1$) en gaan de deelnemers dus “eerder niet akkoord” met het feit dat er geen leerverlies optreedt door GenAI te gebruiken. De onervaren schrijvers staan

neutraal tegenover dat leerverlies ($M = 2.047$, $SD = 1.232$, $Mdn = 2$) maar het mentaliteitsverschil tussen de groepen is niet statistisch significant ($p = 0.137$).

Een opvallend verschil tussen de ervaren en onervaren groep treedt op bij de vierde stelling betreffende de duidelijkheid over het gebruik van GenAI in academische opdrachten. De *novice writers* vinden het namelijk duidelijker in welke mate dit mag ($M = 2.664$, $SD = 1.205$, $Mdn = 3$) dan de *advanced writers* ($M = 1.987$, $SD = 1.158$, $Mdn = 2$) en dit verschil is ook statistisch significant en heeft een negatieve effectgrootte van middelmatige grootte ($W = 1643$, $p = .001$, $r_{rb} = -0.318$). De enige andere stelling met een significant verschil tussen de groepen is de achtste stelling over het beter kunnen afronden van schrijftaken met behulp van GenAI. Hier is het verschil omgekeerd en zijn het de *advanced writers* ($M = 2.822$, $SD = 0.866$, $Mdn = 3$) die meer akkoord zijn dan de *novice writers* ($M = 2.065$, $SD = 1.075$, $Mdn = 2$) met een p-waarde die kleiner is dan .001 en een effectgrootte van middelmatige positieve hoogte ($W = 3372.5$, $p < .001$, $r_{rb} = 0.401$).

Over het algemeen vertonen de ervaren en onervaren wel gelijkaardige houdingen tegenover GenAI en zijn die neutraal tot eerder positief. Aangezien ze aangeven dat het een handig hulpmiddel kan zijn ($M = 3.289$, $SD = 0.677$, $Mdn = 3$) maar dat de antwoorden niet altijd betrouwbaar zijn ($M = 1.882$, $SD = 1.003$, $Mdn = 2$), zal studie B onderzoeken op welke manier dat hulpmiddel ingezet wordt en hoe dit kritisch bewustzijn zich in de praktijk uit.

Deel 3: Studie (B) schrijfpdracht met *think aloud protocols*

3.1 Respondenten

Voor het tweede luik van deze masterproef leggen elf studenten van verschillende leeftijden en academische fasen een schrijfpdracht *in real time* af terwijl ze hun gedachten verbaliseren. Na afloop van studie A hadden respondenten de mogelijkheid om hun contactgegevens achter te laten als ze interesse hadden om deel te nemen aan dit vervolgonderzoek. In eerste instantie werden de studenten benaderd die hun gegevens daar hadden achtergelaten en hadden aangegeven GenAI te gebruiken voor academische doeleinden. Wegens een onevenwichtige verdeling en tekort aan ervaren schrijvers, zijn er ook extra respondenten gezocht via persoonlijk netwerken. Dit waren vier derdebachelor- of masterstudenten van taal- en letterkunde. Zij hebben alle vier ook de vragenlijst ingevuld en zijn mee opgenomen in studie A.

De verdeling over studierichtingen is als volgt: wijsbegeerte ($n = 1$), geschiedenis ($n = 3$), taal- en letterkunde ($n = 5$), communicatiewetenschap ($n = 1$), journalistiek ($n = 1$). Er zijn vier studenten die nog een vak uit de eerstebachelor volgen van wie één student (student 3) van richting veranderd is en al voor het derde jaar op rij aan de KU Leuven studeert en een andere student (student 9) een vier jaar lange niet-afgeronde hogeschoolopleiding achter de rug heeft. De overige twee studenten (student 1 en 2) zijn rechtstreekse instromers vanuit het middelbaar onderwijs. Verder zijn er drie derdebachelorstudenten (studenten 4, 5 en 6), één schakelstudent na een hogeschoolopleiding (student 8), één student in de eerste masteropleiding (student 7), één student in het tweede jaar van een integrale educatieve masteropleiding (student 10) en tot slot één student die nog een verkorte bachelor Engels bijdoet na het behalen van een master (student 11).

De spreiding is divers wat academische ervaring betreft en de details over opleiding en fase bevinden zich in Tabel 6, gerangschikt op aantal jaren ervaring aan een hogeschool of universiteit (tussen haakjes). Volgens het ervaringscriterium van deze studie (bezig zijn met of afgerond hebben van een bachelorproef), vallen studenten 1,2,3 en 9 in de *novice*-groep terwijl de overige studenten bij de *advanced writers* gerekend worden. De complexe academische parcours zoals hierboven geschetst tonen echter aan dat studenten heel wat jaren ervaring kunnen hebben zonder aan dat criterium te voldoen. Tabel 6 bevat ook het geslacht, de leeftijd en moedertaal. De nationaliteit is voor alle elf studenten Belgisch.

Tabel 6: Studie, geslacht, leeftijd, moedertaal participanten (n = 11)

Student	Studierichting + fase (# jaren academische ervaring)	Geslacht (M/V/X)	Leeftijd (jaren)	Moedertaal	Groep
Student 1	Geschiedenis: fase 2 (2)	M	20	Nederlands	Nov
Student 2	Wijsbegeerte optie geschiedenis: fase 1 (3)	V	21	Nederlands	Nov
Student 3	Geschiedenis: fase 1-2 (3)	M	21	Frans	Nov
Student 4	Taal- en Letterkunde Frans-Grieks: fase 3 (3)	V	20	Nederlands	Adv
Student 5	Communicatiewetenschappen: fase 2-3 (3)	V	20	Nederlands	Adv
Student 6	T&L Engels-Frans: fase 3 & verkorte bachelor geschiedenis (4)	M	21	Nederlands	Adv
Student 7	T&L Nederlands-Frans: fase 4 (4)	V	21	Nederlands	Adv
Student 8	Journalistiek: schakelprogramma (4)	M	21	Nederlands	Adv
Student 9	Geschiedenis: fase 1 (5)	M	23	Nederlands	Nov
Student 10	T&L Engels-Spaans: fase 5 (5)	V	22	Nederlands	Adv
Student 11	T&L Latijn-Nederlands-Engels: fase 5 (6)	V	23	Nederlands	Adv

Noot: T&L = taal- en letterkunde, M = man, V = vrouw, Nov = novice writer, Adv = Advanced writer

3.2 Materiaal en methode

De tweede fase van dit onderzoek tracht de interactie van studenten met ChatGPT te simuleren binnen de context van een schrijfopdracht. SMEC heeft ook deze studie ethisch goedgekeurd onder dossiernummer G-204-7581. Aangezien 98,83% van de respondenten uit de eerste fase aangeven ChatGPT te kennen, kiest studie B voor een opdracht met ChatGPT.

Daarbij is het belangrijk te vermelden dat de huidige studie tot doel heeft een zo natuurlijk mogelijke interactie met generatieve AI te verkennen en geen instructie- of initiatiesessie voorziet. Er is geen aandacht voor *prompt engineering*, geen extra uitleg over GenAI en studenten moeten ChatGPT openen en tonen wat ze normaal op zichzelf zouden ondernemen. De focus ligt ook niet op het eindproduct van de schrijfopdracht en die zal in geen enkele mate beoordeeld of geëvalueerd worden aangezien het onderzoek niet bestudeert hoe goed teksten geschreven met behulp van GenAI zijn maar wel hoe studenten GenAI al dan niet gebruiken.

Studenten krijgen een uur de tijd om op hun eigen laptop een tekst van maximaal 500 woorden in een academische schrijfstijl te produceren, gericht op een vooraf bepaald onderwerp. Daarbij mogen zij gebruikmaken van ChatGPT naar eigen behoefte en frequentie met als doel een spontane interactie vast te leggen. Het onderwerp van de tekst zijn de factoren die een invloed hebben op de studiekeuze van scholieren na het middelbaar onderwijs. Een bijgevoegde figuur (Figuur 10, zie Bijlage 2) dient ter inspiratie maar het staat vrij eigen inbreng en argumenten toe te voegen. De afkorting (REF) moet in de tekst als fictieve bronverwijzing staan om het academisch karakter te benadrukken en de context van een wetenschappelijk kleinschalig onderzoek na te bootsen. Het is niet vereist om zelf een

literatuurstudie uit te voeren of op echte bronnen te steunen. In Bijlage 2 is de schrijfoopdracht opgenomen zoals getoond aan de participanten. Voor de geluidsopname van start gaat, kunnen de deelnemers nog eventuele vragen stellen over de opdracht;

Gedurende het schrijfproces moeten de studenten simultaan met *think aloud protocols* hun gedachten verbaliseren voor een *verbal protocol analysis*. Die geverbaliseerde uitingen worden opgenomen, getranscribeerd en later geanalyseerd. De geëxporteerde en geanalyseerde conversatie met ChatGPT dient bovendien als supplementaire data om de *think alouds* te verduidelijken.

Verbal protocol analysis (VPA) is een kwalitatieve methode gericht op het blootleggen van cognitieve processen in de gesproken uitingen van een participant tijdens het uitvoeren van een bepaalde taak (Van Someren et al. 1994, Pressley & Afflerbach 1995, Ericsson & Simon 1996, Green 1998). Zoals uit de literatuurstudie (1.1.1) blijkt, is het een vaak toegepaste methode binnen onderzoek naar het schrijfproces. Steunend op een cognitief model waarin sensorische input via een werkgeheugen naar het langetermijngeheugen migreert, probeert VPA de informatie te expliciteren als die nog in het werkgeheugen aanwezig is en nog zo volledig mogelijk is (Van Someren et al. 1994: 19).

De methode is kwalitatief omdat de verzamelde data rechtstreeks leidt tot inferenties zonder numerieke analyses of statistische tussenstappen. Ericsson & Simon (1996) en Green (1998) onderscheiden variaties in soorten verbale protocollen op drie niveaus: de rapportvorm, het tijdstip en de procedurele variatie. Wat betreft de rapportvorm bestaan er zowel *think aloud* als *talk aloud protocols*. Bij een *talk aloud* verbaliseert een participant informatie die al in verbale of woordelijke vorm aanwezig was in zijn hoofd. Een *think aloud* zet daarentegen ook non-verbale informatie zoals visuele, auditieve of tactiele indrukken in woorden om. Als een participant bijvoorbeeld een verwijzing maakt naar de ruimtelijke situering van een alinea op een blad, zal een *think aloud* aan de orde zijn aangezien dit mentaal beeld eerst in woorden omgezet moet worden. Het tijdstip kan zowel gelijktijdig als retrospectief zijn afhankelijk van het feit of de participant tijdens of na het uitvoeren van de taak zijn gedachten verbaliseert. De procedurele variatie wijst ten slotte op de instructies van de onderzoeker die een participant slechts kunnen aansporen om verder te gaan (niet-gemedieerd) of inhoudelijke bijvragen kunnen stellen en kunnen polsen naar extra toelichting (gemedieerd).

Op methodologisch vlak zijn er twee grote beperkingen bij het uitvoeren van Verbal Protocol Analysis (Ericsson & Simon 1994: 226, Yang 2023). Ten eerste is er het reactiviteitsprobleem en vraagt men zich af of het luidop uitspreken een invloed heeft op het gedachteproces zelf en de mentale last verzwakt.

Ten tweede is er ook het veridiciteitsprobleem of de vrees dat de protocollen niet compleet en accuraat genoeg zijn omwille van niet vermelde gedachten, irrelevante acties of net een te grote interpretatie en gebrek aan spontaneïteit. Verder blijkt uit onderzoek dat elk individu op een andere manier cognitieve processen vervult en die anders verbaliseert wat als gevolg heeft dat generalisaties bemoeilijkt worden (Pressley & Aflerbach 1998: 2). Toch hebben vele empirische studies aangetoond dat gelijktijdige verbalisatie een goede weerspiegeling van effectieve strategieën weergeeft en nuttig is.

Dit onderzoek kiest voor het verzamelen van niet-gemedieerde gelijktijdige *think aloud protocols*. Participanten verwoorden hun verbale en niet-verbale indrukken tijdens het uitvoeren van de opdracht en de onderzoeker stelt zich niet-intrusief op en intervenueert alleen als de *think alouds* niet goed worden uitgevoerd. Dit is de variant waarbij het veridiciteitsrisico het laagst is en de rijkste data verzameld wordt voor de exploratieve opzet van deze studie. Om ervoor te zorgen dat de *think alouds* niet vervallen in een beschrijving van de activiteit (bijvoorbeeld “ik lees nu deze zin”) of een interpretatie, is een briefing en oefenronde aangeraden (Green 1998: 42).

Deze studie bevat een abstracte oefenronde die redelijk ver van de schrijfoopdracht staat maar voorgesteld werd in Green (1998) en Van Someren et al. (1994). Studenten krijgen namelijk een redelijk complexe vermenigvuldiging voorgeschoteld die ze luidop dienen op te lossen om vertrouwd te raken met het verbaliseren van tussenstappen en gedachten. Mathematische problemen lenen zich goed tot het blootleggen van denkstrategieën. Na de oefenronde volgt een kort feedbackmoment waarin aangegeven wordt of de *think aloud* al dan niet goed werd uitgevoerd en kan de onderzoeker indien nodig extra instructies geven. Doorheen de opdracht worden de deelnemers ook aangemoedigd om te blijven spreken als er lange stiltes vallen aan de hand van de open vraag “wat denk je op dit moment?”. Het is belangrijk bij een niet-gemedieerde vorm om geen suggestieve vragen te stellen of te polsen naar het “waarom” maar om deelnemers gewoon aan te sporen te blijven verbaliseren (Van Someren et al 1994: 25).

De getranscribeerde uitingen en interacties met ChatGPT krijgen vervolgens een code als ze verwijzen naar stellingen uit de enquête of terugkerende strategieën die niet in de vragenlijst opgenomen waren. Door de uitingen te segmenteren en de codes te overzien, vergelijkt de analyse het schrijfgedrag van de 11 studenten. Bijlage 4 bevat per participant de uiteindelijke tekst, de interactie met ChatGPT en de *think alouds*. De *think alouds* vormen een letterlijke transcriptie van de uitingen en bevatten dus ook fouten of onafgewerkte zinnen die als zodanig in de citaten van deel 3.3 voorkomen.

3.3 Analyse en resultaten

De analyse van studie B is ook driedelig en bestaat opnieuw uit een luik *omgang met ChatGPT* (3.3.1) en een luik *houding tegenover ChatGPT* (3.3.2) aangevuld door een terugkoppeling naar het schrijfmodel (3.3.3). Het eerste luik biedt een algemeen overzicht van de interactie met ChatGPT gevolgd door een presentatie van geobserveerde acties uit de drie schrijffasen. Het tweede luik vergelijkt of het gerapporteerd gedrag uit de enquête overeenkomt met vastgelegd gedrag tijdens de opdracht en bestudeert evaluatieve opmerkingen uit de *think alouds*. Het derde deel bespreekt algemene elementen uit het schrijfmodel die in de schrijfoopdrachten naar voren komen. Interacties met ChatGPT zijn in een kader weergegeven, citaten uit de *think alouds* tussen aanhalingstekens met een tab.

3.3.1 Omgang met ChatGPT

3.3.1.1 Algemeen overzicht

Tabel 7 presenteert per deelnemer het aantal minuten gespendeerd aan de opdracht (richtlijn was maximaal 60 minuten), het aantal woorden van het afgewerkt schrijfproduct en het aantal *prompts* of het aantal vragen gesteld aan ChatGPT.

Tabel 7: Overzicht schrijfoopdrachten: duur, woorden, prompts (n = 11)

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

Een eerste bevinding is dat het aantal keren dat ChatGPT door de participanten ingezet wordt varieert. Eén participant (student 6) kiest ervoor om helemaal geen generatieve AI te gebruiken en bij de overige tien participanten gaat het aantal ingevoerde *prompts* van 2 tot 15. Om te analyseren hoe ChatGPT ingezet werd, wordt student 6 niet opgenomen in de analyse. In een korte nabeschuiving verduidelijkt student 6 waarom hij de tekst volledig op eigen kracht schreef zonder ChatGPT. Het is niet zozeer uit angst voor fraude/plagiaat is maar omdat hij zichzelf als een academisch schrijver ziet en voldoening uit academisch schrijven haalt (zie ook 3.3.2).

De tijd besteed aan de opdracht varieert eveneens sterk met een participant (student 4) die al na 24 minuten tevreden is met haar tekst en een andere participant (student 8) die het geplande uur overschrijdt. De tekst van student 8 is daarom zeker niet de langste tekst en bevindt zich integendeel bij de kortste drie teksten wat het aantal woorden betreft. Er is binnen de analyse niet echt een verband op te merken tussen het aantal *prompts*, de duur van de opdracht en de lengte van het uiteindelijke schrijfproduct. Meer interactie met ChatGPT staat niet gelijk aan langere teksten, zoals het voorbeeld van student 3 getuigt. Participanten met meer academische ervaring doen ook niet beduidend meer of minder beroep op ChatGPT.

De lengte van hun *prompts* zelf is ook verschillend, gaande van instructies die slechts vier woorden tellen tot uitgebreide paragrafen. Ze zijn vaak gebiedend van aard en niet in een vraagvorm geformuleerd. Zo vragen twee studenten bijvoorbeeld hulp bij het schrijven van een inleidende alinea maar voegt student 4 een werkwoord in de imperatief en iets meer contextuele informatie over betreffende studiekeuze (namelijk die na het middelbaar) toe terwijl student 2 het summier houdt:

Student 2: inleiding voor tekst over invloeden op studiekeuze?

Student 4: Schrijf een inleiding over een onderzoek dat gaat over de factoren die de studiekeuzes van jongeren beïnvloeden na het middelbaar

Een gelijkaardig contrast verschijnt in de *prompts* waarin studenten vragen om een stuk tekst te herschrijven en herformuleren. Sommige studenten maken gebruik van de eenvoudige imperatief “herschrijf” (student 5) terwijl andere studenten de aandacht vestigen op de academische taakomgeving (studenten 1, 7, 8, 10) of taalaspecten aanwijzen waar ze het vaak moeilijk mee hebben (studenten 1 en 11). Ook de plaats van de instructie ten opzichte van de geleverde tekst varieert tussen respectievelijk ervoor (studenten 1, 3, 5, 7, 10, 11) of erachter (student 8).

Student 1: Kun deze tekst nalezen op academisch schrijven, dt fouten, juiste gebruik van dit en dat: [..]

Student 3: Verbeter mijn zin indien nodig: [..]

Student 5: Herschrijf: [..]

Student 7: Verbeter deze tekst in een academische stijl: [..]

Student 8: [..] maak deze tekst wetenschappelijk en behoud mijn alinea's

Student 10: verwoord deze alinea op een meer academische en vlottere manier [..]

Student 11: Kun je mij zeggen of er ergens een schrijffout staat in mijn tekst zonder mijn tekst volledig aan te passen?

Wat de verschillende schrijffases betreft, volgt uit de codering dat generatieve AI voornamelijk aangewend is tijdens het schrijven zelf en in mindere mate in de *pre-* en *post-writing* fase. In Tabel 8 zijn de meest frequente handelingen die van het brainstormen, genereren van een eerste versie, het (academisch) herformuleren, laten nalezen op spel- en grammaticafouten en zoeken van synoniemen. Onderstaande secties zullen dieper ingaan op elke fase en de acties illustreren aan de hand van citaten uit de *think alouds* en *prompts*.

Tabel 8: Overzicht schrijfpodracht: codering strategieën (n = 11)

Strategie \ Student	1	2	3	4	5	6	7	8	9	10	11	TOT
VOOR HET SCHRIJVEN	1: Tips over de opdracht vragen	1										1
	2: Brainstormen over onderwerp	1	1	1			4		5	1		13
	3: Bronnen zoeken/samenvatten											
	4: Moeilijke concepten uitleggen	1										1
	5: Tekstplan genereren				1							1
	6: Feedback op tekstplan vragen											
	7: Niet met een leeg blad starten	1			1	1						3
TIJDENS HET SCHRIJVEN	1: Voorbeeldzinnen genereren											
	2: Eerste versie genereren	1	1		2	2						6
	3: Academische bewoordingen genereren	1										1
	4: Zinnen herformuleren	2			6				1			9
	5: Stijl academischer maken		1		1		3	2	1	3		11
	6: Grammatica- en spelfouten verbeteren	2	1	3	1						1	8
	7: Synoniemen opzoeken			4				1	3		1	9
	8: Taalkwesties oplossen			1							1	2
NA HET SCHRIJVEN	1: Nakijken op juiste inhoud											
	2: Inkorten of langer maken				1							1
	3: Samenvatten (abstract/conclusie)				1							1
	4: Cijfer laten geven ter evaluatie											
	5: Sterke en zwakke punten geven								1		1	2
	6: Chatlogboek bekijken								1			1

NOOT: cijfers duiden frequentie aan van gecodeerde strategieën in *think alouds*

3.3.1.2 Voor het schrijven

In een voorbereidende *pre-writing*-fase wordt vooral beroep gedaan op ChatGPT om inhoudelijke argumenten te bedenken en te brainstormen over factoren die een rol spelen bij de studiekeuze. Zo'n brainstormstrategie is aanwezig bij zes van de tien respondenten, een bevestiging van de trend uit de enquête waar die strategie ook de tweede plaats inneemt wat het gemiddelde van alle stellingen betreft ($M = 2.03$ waarbij 0 = nooit en 5 = altijd). De studenten evalueren daarbij de antwoorden kritisch en selecteren relevante informatie om in hun eigen woorden te integreren in hun tekst. Student 3

beaamt al lezend ook dat hij “daar niet aan gedacht [had] maar nu [hij] het lees[t], het ook wel logisch in [zijn] hoofd [klinkt]”.

Student 1: “De redenen hiervoor zijn iets minder duidelijk. Het kan zijn dat. Hmmm. Effe zien of dat ChatGPT iets zegt. ‘Waarom zijn jongeren zo beïnvloedbaar door leeftijdsgenoten?’”

Student 2: “Ik ga eens vragen wat dat ChatGPT allemaal te zeggen heeft over studiekeuze eventueel. Ik typ in ‘welke factoren beïnvloeden de studiekeuze van jongeren?’. Ik lees hier ouders en familie, euhm carrière doelen, interesses en passies. Ja ik denk dat we dat. We kunnen dat zetten bij het feit dat niemand hun beïnvloed heeft omdat dat dan waarschijnlijk komt uit hun interesse en hun passies.”

Student 3: “Hier ga ik aan ChatGPT vragen even voor. ‘Wat zou een mogelijke reden zijn waarom scholieren hun partners niet aanspreken over hun studiekeuze’.”

Student 7: “Dan ga ik vragen in welke mate kan familie een invloed hebben op de studiekeuze van jongeren in het zesde middelbaar. Dan ga ik dat verklaringje overschrijven voor familie.”

Student 9: “Dan ga ik ChatGPT gebruiken om te zien wat die te zeggen heeft over mogelijke statistieken of andere bevindingen rond die factoren en de psychologische invloed daarvan.”

Student 10: “Ik ga aan ChatGPT vragen om mij wat redenen te geven. Ik ga vragen ‘lees deze tekst en geef mij wat argumenten waarom deze personen invloed hebben op de studiekeuze van jongeren’. Oké dan ga ik dat tablad eventjes rechts zetten.”

Opvallend is dat het definiëren van concepten, de *pre-writing* handeling die volgens studie A het meest uitgevoerd zou worden, haast afwezig is in studie B. Slechts één student (student 1) benut ChatGPT om het terminologisch verschil tussen medeleerlingen en vriend te duiden aan de hand van volgende cryptisch geformuleerde *prompt*:

Student 1: verschil tussen medeleerling en vriend
--

De overige handelingen uit de enquête met een gemiddelde frequentie tussen “zelden” en “soms” zijn eveneens in beperkte mate in de *think alouds* terug te vinden. Geen enkele participant vraagt tips om de gehele opdracht aan te kunnen pakken. Student 1 gaat wel te rade bij ChatGPT om te weten uit hoeveel zinnen een conclusie in een academische schrijfofdracht idealiter bestaat. Die *prompt* is bijgevolg onder “tips over opdracht” ingedeeld maar spitst zich eigenlijk toe op een deelopdracht, namelijk het schrijven van een conclusie.

Student 5 daarentegen voert de originele opdracht in zijn geheel in ChatGPT in om vertrekkend van dat antwoord een eigen antwoord op te stellen. Die student baseert een eigen tekstplan op de structuur van de voorbeeldtekst ChatGPT en haalt er vooral inhoudelijke hulp uit:

Student 5: “Oké. Ik ga het nu kopiëren en gewoon plakken in Word en dan ga ik het deftig lezen en dan echt zo zin per zin zien wat dat ik goed vind en wat niet enzo. Want hij heeft wel echt een structuur gemaakt. Er staat zo inleiding, midden, slot en ik vind dat wel goed. Dus ik denk dat ik de structuur wel ongeveer ga houden.”

Over het algemeen plannen de studenten weinig en beginnen ze vaak onmiddellijk een definitieve versie te schrijven zonder lay-out of kladversie. Dit is een algemeen fenomeen van het digitaal schrijven waarbij de tekst makkelijk al schrijvend aangepast kan worden maar toch valt het op dat de deelnemers ook niet vooruitkijken wat de inhoud en argumentatie betreft. Net zoals student 5, is student 2 zich echter wel bewust van de drieledige structuur. Ze begint uit zichzelf aan de opdracht door die drie woorden op te schrijven en wisselt al schrijvend af tussen de delen. Zo schrijft ze al een deel van de conclusie nog voor het middenstuk afgewerkt is:

Student 2: “Voorlopige structuur er al inzetten van een inleiding, midden, slot. [...] De grootste conclusie is dan dat als er iemand invloed heeft, is het de familie. Misschien kan ik dat al in het besluit zetten.”

Een laatste student die wel bewust met het tekstplan bezig is, is student 8. Die begint namelijk informatie uit de grafiek te groeperen aan de hand van kleurtjes en gaat dan over tot een kladversie die hij zelf “draft” noemt waarin hij in volzinnen aangeeft welke informatie hij wil behandelen. Daarna komt hij tot een meer uitgeschreven “probeersel” en ten slotte gaat hij over op een definitieve versie. Bij het schrijven van zijn “probeersel” kijkt hij regelmatig terug welke ingevingen hij had in de planfase en wat nog ontbreekt. Een fragment uit de “draft” zelf en de *think aloud* verduidelijken:

Student 8 (draft): “Starten met inleiding waarin ik zeg dat afstuderen normaal gebeurt op een leeftijd van 18 jaar. Dat is een leeftijd waarin standpunten die gevormd worden nog wel beïnvloed kunnen worden, maar niet voor een volledige omkering zorgen. Daarom dat de optie, niemand beïnvloeden het meest is gekozen in de tekst.”

Student 8: Dan ga ik kijken in mijn draft wat ik heb geschreven waarom ik familie eerst ter sprake breng. [...] Oké effe kijken in mijn draft wat ik nog heb. Ik start met de inleiding en ik zeg dat afstuderen normaal gebeurt op een leeftijd van 18 jaar. Dat heb ik nog niet.

Het *blank page syndrom* overkomen en effectief de eerste zin op papier zetten is een strategie die in beperkte mate in studie B naar voren komt, aansluitend bij de enquête waar die een gemiddelde van 1.22 op 5 behaalt (met 0 = nooit en 5 = altijd). Zo vragen studenten 2 en 4 bijvoorbeeld hulp bij een inleidende zin of alinea (zie 3.3.1.1). De overige studenten schrijven hun eerste woorden of zinnen op eigen kracht maar beamen dat het beginnen een uitdaging is en laten aarzelende pauzes vallen (studenten 3, 4, 7, 11).

Student 3: Ja. Hm. Wacht he. Ik ben aan het denken of euhm hoe hoe ik daaraan zal moeten beginnen. Een eigen weg. Hm

Student 4: Ik ga beginnen met een introductie dat ga ik doen. Euhm. Deze. Euhm pupupum. Euhm. Meestal gebruik ik voor mijn ik ga ChatGPT gebruiken voor een inleiding te schrijven

Student 7: Oké dan zal ik maar beginnen zeker. Oké. Ik ga eerst al een inleidende zin schrijven met? Ik ga effe de vraag hier zetten vanuit de opdracht. Euhm. Ik ga gewoon al een vage eerste zin schrijven

Student 11: Oké. Euhm. Studiekeuze na het middelbaar onderwijs. Euhm. Effe denken. Euhm. Wacht ik zou zelf altijd beginnen met zo'n zin. Er is. Een beetje een algemene zin. Er zijn veel factoren of ja. Er zijn verschillende factoren [...]

3.3.1.3 Tijdens het schrijven

De tweede sectie bestudeert de effectieve *writing stage*, de schrijf- of productiefase. Deze sectie overloopt eerst de stellingen gekoppeld aan studie A maar legt de focus vervolgens op 2 aspecten die niet in de enquête opgenomen waren maar wel gecodeerd zijn.

In de *writing* fase corresponderen de drie vaakst gerapporteerde toepassingen van studie A ($M > 1.6$, $Mdn \geq 2$) grotendeels met de vaakst gecodeerde strategieën in studie B. Dit zijn het herformuleren van zinnen en het laten corrigeren op spel- en grammaticafouten of op het academisch register. In totaal passen de tien studenten allemaal een combinatie van die verbeterstrategieën toe. Daarbij laten drie studenten hun tekst simpelweg herschrijven terwijl de overige zeven specifiek op het academisch register focussen en/of de woorden “fouten”, “spelfouten”, “schrijffouten” of “taalfouten” in hun *prompts* zetten. Sommige studenten focussen in die *prompts* op gekende moeilijkheden, bijvoorbeeld “dt-fouten en dit/dat fouten” (student 1) terwijl anderen het algemeen houden (zie 3.3.1.1).

Om tekst te verbeteren, schakelen de meeste studenten ChatGPT alinea per alinea in. Gefrustreerde uitingen van deelnemers (student 1, student 11) die een langer stuk invoeren, tonen aan dat grote blokken tekst geen goed resultaat opleveren.

Student 1: “Die dingetjes van ChatGPT lijken mij niet echt goed, die verbetering dus ik ga er gewoon zelf nog eens overgaan. Ik denk dat die tekst wat te lang is voor ChatGPT om dat echt deftig te bekijken.”

Student 11: “[Leest van ChatGPT af] ‘Excuses voor de verwarring, ik zie dat mijn eerdere opmerking over herhaling niet helemaal van toepassing is op je tekst’. Jaaa. Als die dat begint te doen, dan denk ik dat het wel goed is.

Niemand laat echt voorbeeldzinnen of academische formuleringen genereren. Er zijn echter wel vier deelnemers (student 1, 2, 4 en 5) die een eerste versie van een tekst of stuk tekst laten genereren. Zo doen studenten beroep op ChatGPT voor een inleiding (student 2 en 4, zie 3.3.1.1), een “inleidingszin”

(student 1) en een academisch slot (student 4). Het genereren van dat academisch slot bij student 4 is ook bij de *post-writing* fase ondergebracht omdat het een samenvatting van de geproduceerde tekst vormt (zie 3.3.1.4). Student 5 kiest er echter resoluut voor om de hele opdracht te kopiëren en in ChatGPT te plakken om zo een voorbeeldversie te bekomen die ze vooral als inhoudelijk skelet gebruikt nadien en zin per zin herschrijft in haar eigen woorden.

Student 5: “Ik ga denk ik eerst gewoon dit, de opdracht letterlijk kopiëren en plakken in ChatGPT en kijken wat die daarvan maakt en dan heb ik zo een idee van wat ik zelf ga doen. En dan ga ik mijn tekst aanpassen [...] ik ga letterlijk niets erbij schrijven. Ik doe gewoon [de opdracht] zo copy paste. Want er staat in maak een argumentatieve tekst dus.”

De enige uiting die gecodeerd is als een gegenereerde “academische bewoording” bevindt zich in de *think aloud* van student 3. Hij zoekt namelijk een concluderend signaalwoord en vraagt specifiek welke optie het meest academisch is:

Student 3: Signaalwoorden die men voor een slot of conclusie kan gebruiken

ChatGPT: Enkele signaalwoorden die je kunt gebruiken om een slot of conclusie aan te geven zijn: **Kortom**, Concluderen, Samengevat, Al met al, Tot slot, Ter afsluiting, Om samen te vatten, Uiteindelijk, Tot besluit

Student 3: wat is meer academisch?

ChatGPT: in academische contexten wordt vaak gezocht [...]

Twee aspecten die niet in de enquête opgenomen waren maar ook frequent toegepast worden, zijn de zoektocht naar synoniemen en hulp bij taalkwesties (juiste lidwoord, juiste meervoudsvorm, ...). Sommige studenten (6, 10, 11) maken ook gebruik van andere websites op het internet voor beide kwesties aangezien dat niet expliciet verboden is en niet tegengehouden wordt. Het is echter ook niet aangemoedigd of voorgesteld waardoor er enigszins een onevenwicht te bespeuren is tussen studenten die de opdracht nauwgezetten volgden en zij die het hele internet erbij betrekken.

In totaal schakelen zes van de elf participanten hulp in bij de zoektocht naar synoniemen: drie studenten exclusief met ChatGPT (studenten 3, 8 en 9), twee studenten gebruiken alleen het internet (studenten 6 en 10) en de laatste student (student 11) maakt eerst gebruik van de website synoniemen.net maar gaat over op ChatGPT wanneer ze geen “nuttig” antwoord vindt.

Student 11: Weet synoniemen [de website synoniemen.net] een ander woord voor instantie. Neen. [...] Ja adviesorgaan zit nu in mijn kop maar we gaan eens aan ChatGPT vragen. Ken jij een ander woord voor adviesorgaan. Dat heeft ook niets nuttigs want dan krijg ik gewoon adviesinstantie en dat is wat ik net wil vermijden.

Onder de categorie taalkwesties zijn situaties gecodeerd die betrekking hebben op de keuze van een lidwoord voor het woord “grafiek” aan de ene kant (zie 3.3.2, student 3) en het akkoord tussen onderwerp “het merendeel” en het vervoegd werkwoord aan de andere kant (student 11). Voor die

laatstgenoemde kwestie is eerst de website *taaladvies* geconsulteerd en is ChatGPT als laatste toevlucht gebruikt hoewel ze ook dan niet volledig overtuigd lijkt te zijn.

Student 11: “Ik moet effe opzoeken *het aandeel houdt of het aandeel houden* want dat weet ik niet. Het aandeel studenten houdt of houden. Taaladvies help mij [...] Hier ben ik nog steeds niets mee dus nu ga ik misschien effe aan ChatGPT het antwoord vragen. [...] Het juiste zinsgebruik hangt af van de context ja euhm [...] Oké gewoon aandeel houdt pakt.”

Naast het rechtstreeks opzoeken van een synoniem, is er ook onrechtstreekse beïnvloeding aanwezig. Zo nemen bepaalde studenten al dan niet bewust woorden over van ChatGPT. Student 1 gebruikt bijvoorbeeld meermaals het woord “scholier” in plaats van “student” nadat ChatGPT dit één keer gebruikt had en neemt ook de formule “verder onderzoek” in de conclusie op nadat die in een inhoudelijk antwoord van ChatGPT over de ideale lengte van een conclusie voorkwam. Student 4 twijfelt zelf eerst over het woord “educatief”, vervangt het door “onderwijsgerelateerd” maar past het uiteindelijk terug aan beïnvloed door ChatGPT:

Student 4: “om sociale motivatie in plaats van educa euhm van educatieve neen dat is een raar woord. Om sociale motivatie in plaats van onderwijsgerelateerde gerelateerde motivatie [...] ChatGPT zegt educatieve motieven te gebruiken. Daar was ik ook al over aan het twijfelen in het begin dus ik ga dat dan toch maar doen.”

Zoals eerder vermeld, is ChatGPT niet het enige betrokken (online) hulpmiddel. Student 8 vormt een interessant voorbeeld van iemand die verschillende tools combineert en volgens een vast patroon te werk gaat. Naast plannen met kleurtjes, opstellen van een draft en die omvormen tot een “probeersel” en ten slotte tot een definitieve versie, combineert deze student namelijk ChatGPT en de SAN Schrijfhulp Academisch Nederlands, een tool binnen de associatie KU Leuven (Vandooren 2022, De Wachter, L., & d’Hertefelt, M. 2018). Zelf licht de student zijn werkwijze als volgt toe:

Student 8: “Dan open ik ondertussen de schrijfhulp van de KUL. Ik doe altijd eerst ChatGPT en dan pas de schrijfassistenthulp omdat ChatGPT ook nog wel schrijffouten durft te maken en ook niet altijd de zinnen die ze maken even goed zijn. En ook ChatGPT gaat als actief schrijver vaak dezelfde woordconstructies gebruiken en wanneer ge dat in de schrijfassistent ziet of plaatst, dan gaat dat opvallen hoe vaak die [voorkomen] en kan ik op die manier op zoek gaan naar synoniemen.”

Chronologisch verzoekt deze student ChatGPT eerst om zijn tekst “wetenschappelijk” te maken en daarbij de alinea’s te behouden, vervolgens om de “passieve constructies eruit te halen zonder inhoudelijk te veranderen” en schakelt hij dan over op de SAN. Met behulp van de SAN, controleert student 8 de spelfouten en gaat hij op zoek naar vaak voorkomende patronen. Zo blijkt dat de uitdrukking *geeft aan dat* zeer frequent is, waarna hij aan ChatGPT vraagt: “geef een alternatief voor de constructie *geeft aan dat*”. In zijn tekst vervangt hij het vaak voorkomende patroon vervolgens door de suggestie van ChatGPT (*vermeldt*) en laat hij ten slotte de SAN tool het omslachtig taalgebruik

nakijken “want dat is soms met ChatGPT ook wel een probleem”. Als ook dat in orde blijkt te zijn, is hij gelukkig met zijn tekst. Wegens tijdstekort heeft student 5 haar schrijfroutine niet volledig kunnen demonstreren maar ook zij vermeldt: “dan zou ik op dit moment het copy pasten en zou ik het zo naar die ene site doen zo SAN schrijfhulp.”

3.3.1.4 Na het schrijven

De derde *post-writing* fase heeft aanzienlijk minder resultaten dan de voorgaande fasen met een totaalscore van slechts vijf gedetecteerde strategieën en nog enkele acties niet verbonden aan de enquête. In de schrijfoopdrachten is er één deelnemer (student 4) die een samenvattend conclusie laat genereren door ChatGPT op basis van haar eigen tekst. Ze vermeldt hypothetisch ook: “als ik zie dat ik 11 woorden te veel heb, dan zou ik vragen [aan ChatGPT] om het bondiger te maken” en gebruikt GenAI dus om de lengte aan te passen en de inhoud samen te vatten.

Student 4: “alles nog eens kopiëren en vragen om een academisch slot te schrijven. Neen ‘schrijf een slot in academische taal van een drietal zinnen’. Voilà en dan neem ik dat over.”

Een andere handeling die onder de derde schrijffase gebracht is, is de vraag van student 11 over de duidelijkheid van zijn ingevoerde tekst. Het antwoord van ChatGPT geeft namelijk goede en sterke punten en evalueert de ingevoerde tekst:

Student 11: Is mijn tekst duidelijk?

ChatGPT: Ja, over het algemeen is je tekst duidelijk en goed gestructureerd. Je hebt helder beschreven hoe verschillende factoren van invloed kunnen zijn op de studiekeuze na het middelbaar onderwijs, zoals eigen interesses en vaardigheden, advies van familie en vrienden, invloed van leraren, en de rol van begeleidingscentra zoals het CLB.

Je hebt ook goed gebruik gemaakt van bronverwijzingen om je punten te ondersteunen. De lezer kan begrijpen welke informatie gebaseerd is op onderzoek en welke op algemene observaties.

Er zijn echter enkele kleine verbeteringen mogelijk, zoals het vermijden van herhalingen en het verduidelijken van sommige zinnen. Over het algemeen is het echter een goed geschreven en informatief stuk.

Een andere student, student 9, polst eveneens naar de sterke en zwakte punten van zijn schrijfproduct door ChatGPT om “suggesties” te vragen:

Student 9: Heb je enige suggesties. Om deze tekst te verbeteren? [..]

ChatGPT: Hier zijn enkele suggesties om de tekst te verbeteren:

1. Introduceer de bronnen van de geciteerde gegevens met volledige referenties, zoals auteurs, titels van onderzoeken of enquêtes, en publicatiedata. Dit versterkt de geloofwaardigheid van de informatie.
2. Vermijd het gebruik van afkortingen zoals "REF". In plaats daarvan, vervang ze door de daadwerkelijke referenties of voeg voetnoten toe aan het einde van de tekst.
3. Verduidelijk de formulering en structuur van sommige zinnen om ze vloeiender te maken en de boodschap duidelijker over te brengen.
4. Gebruik consistente terminologie en stijl doorheen de tekst.

Dezelfde student is ook de enige die bewust terugzoekt in zijn conversatie met ChatGPT en teruggrijpt naar een eerder lijstje met synoniemen voor het woord “factor”. Hij benut het “chatlogboek” en scrollt twee keer terug naar boven om het oorspronkelijk antwoord op zijn eerste *prompt* te herbekijken:

Student 9: dan ga ik misschien een synoniem voor factoren vinden. Ik heb het toch openstaan dus ik ga vragen aan ChatGPT “wat zijn enkele synoniemen voor fact voor het woord factoren”. [...] We gaan die factoren veranderen want ik heb dat al veel te veel gebruikt dus we gaan terug naar het lijstje. Externe variabelen [...]. Niet factoren dat woord gebruik ik al veel te veel. Om deze invloeden. Invloeden, een woord dat ik ook van de lijst van ChatGPT gehaald heb.

Daarbovenop voeren drie studenten extra wijzigingen door wat de macrostructuur betreft na het schrijven van de tekst. Zo voegt student 2 een titel toe (“Wie heeft invloed op de studiekeuzen na het middelbaar onderwijs?”), voegt student 1 tussentitels toe aan elke alinea (“Inleiding”, “Invloed van leeftijdsgenoten op de studiekeuze”, ...) en besluit student 10 om toch een echte bibliografie en bronverwijzing te incorporeren omdat er nog resterende tijd was:

Student 10: “Ik vind het eigenlijk wel al goed. Ik ga inderdaad dat was ik aan het bedenken toch nog wel eventjes kijken naar andere bronnen.”

Student 10 zoekt op Google extra artikels en stoot op een masterscriptie over de studiekeuze na het middelbaar die ze bestempelt als “heel interessant” en iets waar ze “ook meer informatie uit [wil] halen”. Steunend op die scriptie, voegt de student informatie met bijhorende APA-bronvermelding in de tekst zelf toe. Helemaal aan het einde van het schrijfproces na revisierondes maakt ze gebruik van Limo en een online APA-bron generator om een afgewerkte bibliografie toe te kunnen voegen van alle vernoemde bronnen. De bibliografie is duidelijk een opdracht die uitgesteld wordt tot na het schrijven:

Student 10: “Dus ik heb nu al die bronnen maar die heb ik in het rood gezet want die heb ik natuurlijk niet goed geciteerd dus die ga ik nu even zoeken allemaal in het artikel en dan ga ik die op de correcte manier citeren. Dus ja. Euhm. Het is wel vervelend dat ik dat niet kan. Ik ga eerst nog eens alles lezen en daarna eindig ik gewoon met al mijn bronnen te doen.”

Hierbij is het opvallend op te merken dat student 10 als enige gebruik maakt van een zichtbare rode markering om de aandacht te vestigen op aspecten die ze op het einde van het schrijfproces nog moet verfijnen. Niet alleen de bronvermelding maar ook het woord “behoorlijk” staat bijvoorbeeld in het rood omdat hier ontevredenheid over is. In de definitieve versie, wordt “behoorlijk” uiteindelijk ook vervangen door “fundamenteel”. Andere deelnemers vermelden soms wel dat ze later op iets terug willen komen of iets willen aanvullen maar gebruiken hier geen visuele aanduidingen of kleuren voor.

3.3.2 Houding tegenover ChatGPT

Deze sectie is tweeledig en bevat zowel een vergelijking tussen het geobserveerd schrijfgedrag van de elf participanten en hun antwoorden uit de enquête als een samenvatting van evaluatieve opmerkingen geuit tijdens de *think alouds*.

Tabel 9 biedt een overzicht van de scores die de 11 participanten in hun bevraging invulden waarbij het cijfer 0 correspondeert met “ik gebruik generatieve AI nooit om ...” en het cijfer 4 met “ik gebruik generatieve AI altijd om ...”. De groene vakjes duiden de strategieën aan die op basis van hun *think alouds* in de opdracht gedetecteerd zijn.

Tabel 9: Gebruik GenAI volgens enquête en opdrachten (n = 11)

	Stelling \ Student	1	2	3	4	5	6	7	8	9	10	11
VOOR HET SCHRIJVEN	1: Tips over de opdracht	3	0	3	2	3	0	0	3	0	3	1
	2: Brainstormen over onderwerp	2	3	4	2	3	0	2	1	0	3	2
	3: Bronnen zoeken/samenvatten	3	0	2	1	0	0	3	1	0	0	1
	4: Moeilijke concepten uitleggen	3	0	3	3	3	0	3	3	0	2	2
	5: Tekstplan genereren	0	0	2	2	2	0	0	1	0	0	1
	6: Feedback op tekstplan	3	0	2	2	2	0	1	4	0	0	3
	7: Niet met een leeg blad starten	0	3	1	3	4	0	0	2	0	2	0
TIJDENS HET SCHRIJVEN	1: Voorbeeldzinnen genereren	0	1	3	2	4	0	1	3	0	1	1
	2: Eerste versie genereren	0	0	3	1	3	0	1	2	0	1	1
	3: Academische bewoordingen	1	2	3	3	3	0	1	2	0	1	0
	4: Zinnen herformuleren	2	2	3	3	3	0	3	4	0	3	3
	5: Stijl academischer maken	0	1	3	4	3	0	2	3	0	3	3
	6: Grammatica- en spelfouten	3	2	2	4	1	0	3	4	0	0	3
NA HET SCHRIJVEN	1: Nakijken op juiste inhoud	0	0	3	1	1	0	1	2	0	0	3
	2: Inkorten of langer maken	1	0	3	2	1	0	2	4	0	3	1
	3: Samenvatten (voor abstract/conclusie)	2	0	2	2	2	0	2	3	0	3	0
	4: Cijfer geven	0	0	0	1	0	0	0	0	0	0	0
	5: Sterke/zwakke punten geven	1	0	1	3	1	0	2	0	0	0	3
	6: Chatlogboek bekijken	3	0	3	0	1	0	3	0	0	0	0

NOOT: cijfers geven een score weer (0 = nooit, 1 = zelden, 2 = soms, 3 = vaak, 4 = altijd)

Duiding is daarbij vereist voor student 9 die op het moment van het invullen van de enquête nog nooit gebruik had gemaakt van generatieve AI maar tussen de afname van de enquête en zijn deelname aan studie B wel ervaring opdeed. De overige studenten voorspelden hun frequente acties meestal vrij accuraat. Student 1 gaf echter in de bevraging aan dat hij het nooit inzette om een eerste versie te genereren of op gang te komen en vroeg in de opdracht toch een inleidingszin aan ChatGPT. Student 4 vertoont een gedrag in de omgekeerde richting want zij schatte zichzelf in als iemand die vaak academische bewoordingen en voorbeeldzinnen genereert maar liet in de opdrachten een hele versie van inleiding en conclusie genereren in plaats van losse bewoordingen of voorbeeldzinnen. Studenten 3, 5 en 8 lieten in hun antwoorden ook een hogere frequentie van allerlei handelingen uitschijnen dan in hun *think alouds* naar voren kwam.

Wat hun houding tegenover ChatGPT betreft, valt op dat de meeste deelnemers de antwoorden van ChatGPT aan een kritische evaluatie onderwerpen. Als de output niet voldoet aan de eisen die ze zichzelf opleggen of aan de inhoudelijke kennis die ze zelf bezitten over bijvoorbeeld de academische

stijl, stellen ze zich sceptisch over en vertrouwen ze eerder op hun eigen vaardigheden (student 5, 7, 8, 9). Als de interactie met ChatGPT echter niet zo vlot loopt, haken studenten snel af en vertrouwen ze op hun eigen revisievaardigheden (bijvoorbeeld studenten 1 en 7).

Student 1: “Die dinges van ChatGPT lijken mij niet echt goed, die verbetering dus ik ga er gewoon zelf nog eens overgaan. Ik denk dat die tekst wat te lang is voor ChatGPT om dat echt deftig te bekijken.”

Student 5: “Dat zou ik nooit schrijven in een paper zelf dus dat ga ik ook niet overnemen.”

Student 8: “ChatGPT heeft niet alle passieven eruit gehaald dus dan verander ik die zelf.”

Student 9: “Ik ga ChatGPT stijl daar niet zomaar in overpakken”

Student 7: “Ik ben aan het twijfelen of ik de hele tekst die chatgpt nu heeft gemaakt op basis van de tekst die ik al had ga overnemen of niet. Euhm. Ik ga mijn eigen tekst nog eens lezen en zien welke ik beter vind.”

Ze opteren vaak voor een combinatie van hun eigen versie en die van ChatGPT aan de hand van knip- en plakwerk. Dit getuigt van aarzeling tegenover de kwaliteit van de antwoorden van ChatGPT of van een groter kritisch overzicht. Zo voorkomt student 3 bijvoorbeeld herhaling van een signaalwoord dat hij in een eerdere alinea al gebruikt had maar waartoe ChatGPT geen toegang had.

Student 3: “Ah ja het gebruikt ‘aan de andere kant’ maar ik heb het al eerder gebruikt dus ik ga het gewoon zo laten.”

Student 3: “Niet willen belasten met hun studiekeuze, dit ga ik niet nemen. Maar het tweede deel van die zin denk ik wel.”

Die aarzeling is echter niet bij iedereen aanwezig en studenten 2, 5 en 10 nemen grote stukken rechtstreeks over zonder die zelf nog te willen herschrijven. Student 5 merkt wel op dat zij in een “echte” academische setting niet zomaar een antwoord van ChatGPT zou durven kopiëren uit angst voor fraude en plagiaat. Student 7 vraagt ook of ze “eigenlijk rekening [moet] houden met plagiaat?”.

Student 2: “Het is verrassend dat. Oké ja, “het is verrassend” klinkt inderdaad beter. Vervolgopleiding, goed een ander woord voor gebruikt. Daar was ik niet opgekomen.”

Student 5: “deze zin eigenlijk goed dus ik ga die laten staan. ChatGPT heeft dat goed gedaan. Alleen in een echte. Als ik dit zou indienen als paper of ik had meer tijd, zou ik het waarschijnlijk niet laten staan want dit gaat sowieso. Ik heb exact de zin overgenomen en dan zou ik misschien buizen ofzo.”

Student 10: “dan kopieer ik dat en dan lees ik dat gewoon helemaal opnieuw om te kijken of ik dat goed vind. Hm. Ja oké.”

Student 6 opteert ervoor om de tekst volledig zelf te schrijven. Het gebruiken van hulpmiddelen past namelijk niet bij de identiteit die hij zichzelf aangemeten heeft en bij zijn dromen om later een plaats

in de academische wereld te verwerven. Hij geeft wel aan dat hij beseft welke rol technologie in de samenleving speelt en is van plan om zich er vertrouwd mee te maken, vooral om ondervragingen te doen van examenleerstof.

Student 6: “Ik denk dat het heel nuttig zou zijn maar iedere keer als ik het zou proberen beginnen gebruiken zou ik het gevoel hebben van ik heb er niet het werk ingestoken. Ik wil alles lezen en herlezen en schrijven en herschrijven en dan is het van mij. Ik wil er grip over hebben en het hebben behandeld.”

Ook student 11 geeft na afloop van de opdracht aan dat ze voor Engelse schrijfoopdrachten meer gebruik maakt van GenAI dan voor Nederlandse opdrachten.

Student 11: “Voor Engelse opdrachten gebruik ik het veel meer. In de lessen Engels taal & tekst hebben ze ons dat ook echt aangeleerd. Ik gebruik het voor Nederlands wel als thesaurus maar voor taalkwesties gebruik ik liever een betrouwbare website zoals taaladvies of laat ik het door mijn mama of vriend nalezen en kan ik zelf ook beter inschatten of iets de juiste nuance heeft.”

Door het cryptische en summiere karakter van *prompts*, verkrijgen studenten soms ongewenste resultaten en ontstaan er af en toe frustraties. Student 3 twijfelt bijvoorbeeld over het lidwoord van het zelfstandig naamwoord “grafiek” maar formuleert de vraag op een onduidelijke manier en krijgt een algemeen ontwijkend antwoord. De student herformuleert de vraag en krijgt bij de tweede poging wel het beoogde resultaat.

Student 3: het de grafiek

ChatGPT: Het spijt me, maar ik heb geen toegang tot afbeeldingen of grafieken. Als je een specifieke grafiek hebt waar je naar wilt verwijzen of waar je vragen over hebt, kun je proberen het te beschrijven of je vragen te stellen, en ik zal mijn best doen om je te helpen op basis van de beschrijving die je geeft.

Student 3: is het de of het grafiek?

ChatGPT: In het Nederlands wordt "de grafiek" gebruikt. Dus het juiste lidwoord is "de". Dus je zou zeggen: "Ik bekijk de grafiek."

Bij het laten corrigeren van teksten treden soms ook communicatieproblemen op als ChatGPT de tekst niet terugvindt (student 9) of een antwoord begint te genereren alvorens de vraag volledig gesteld is, zoals student 11 aangeeft. Dergelijke communicatieproblemen leiden tot lichte frustraties en soms tot het einde van de interactie met ChatGPT.

3.3.3 Terugkoppeling naar het schrijfmodel

Volgens het model van Flower & Hayes (1980) beïnvloeden zowel het langetermijngeheugen als de taakomgeving het schrijfproces (zie 1.1.1). De taakomgeving bestaat uit het retorisch probleem (voor wie, waarover, welke voorwaarden, ...) en de geproduceerde tekst. De schrijfoopdracht van dit onderzoek behelst een retorisch probleem in een argumentatieve, academische stijl van maximum 500 woorden.

Bijna alle deelnemers verwijzen op een zeker moment tijdens het schrijven naar het academisch register en houden ook het aantal geproduceerde woorden in het oog, wat de invloed van het retorisch probleem bevestigt. Tabel 11 toont per student hoe vaak het aantal woorden gecontroleerd wordt en wat daarbij in hun hoofd omgaat. Ze doen dit niet altijd luidop (student 2), soms pas op het einde (studenten 1,3 en 11) maar vaak regelmatig (studenten 5, 6, 7, 8, 9, 10). Ze berekenen ook verhoudingen ($1/10^e$, $2/5^e$, de helft van het maximum) om de abstracte cijfers tastbaarder te maken.

Tabel 10: Retorisch probleem: nakijken aantal woorden ($n = 10$)

Student?	#	Citaten <i>Think aloud</i> met verwijzing naar aantal woorden
Student 1	1	- Nu zit ik ongeveer aan 500 dus ik ga
Student 3	1	- Ah 203 woorden. Oké, ik moet niets verkorten.
Student 4	1	- Dan zie ik dat ik 11 woorden te veel heb.
Student 5	6	- Ah wow 58 woorden dus ik zit al in $1/10^e$. - Nu heb ik 116 woorden jippie. - Oké ik heb nu 130 woorden en dat is niet veel en ik heb niet veel tijd niet meer oei ik ben een beetje te traag. - Oké ik heb nu 250 woorden. - Ik heb nu 300 woorden dus nog een kleine alinea - Ik heb nu 350 woorden ha oeps. Op zich het was max 500.
Student 6	5	- Ik check graag om de zoveel minuten hoeveel woorden ik heb. Ik zit nu aan 90 - We zitten ongeveer aan 180 aan effe kijken 174 woorden, ongeveer $2/5^e$. - Nu gaan we nog eens kijken naar het aantal woorden. Ik denk dat we nu aan 300 gaan zitten. Wacht he. 307. Oké. Ik denk dat ik tot de 400 misschien 450 ga geraken als ik zo verder werk. - Ik denk dat ik nu misschien 380 woorden heb. 376. Ik denk dat ik een nieuwe concluderende alinea ga beginnen en het wat korter aanpakken. - Wacht we gaan nog eens kijken. 407 woorden.
Student 7	3	- Oei ik heb echt nog heel weinig woorden. - Oké dan heb ik nu één alineaatje van 134 woorden die ik zo ga laten - Wacht eens kijken hoeveel woorden dat ik heb.
Student 8	4	- Nu heb ik 175 woorden - Hoeveel woorden zijn dit? 203 - Dat kan ik nog wel uitbreiden want ik heb nog wel wat woorden over. - Dan check ik hoeveel woorden ik heb. 343
Student 9	2	- Woordcount nakijken en ik zit op 220 dus bijna de helft van het maximum. - En dan gaan we langzaamaan naar een vierde alinea en conclusie werken. Ik ga nakijken hoeveel woorden ik al heb en ik heb er 355.
Student 10	2	- Ik ga nu al eens kijken hoeveel woorden ik heb. Nu 182. - Oké. Ik ga eens kijken hoeveel woorden ik heb.
Student 11	1	- Ik zit aan 480 woorden

Een andere opvallende gelijkenis is het grondig doornemen van de opdracht. Voor de geluidsopname van start gaat, wordt die samen met de onderzoeker overlopen en is er ruimte voor eventuele vragen. Zodra de opname start, leest een groot deel van de deelnemers de opdracht echter al dan niet luidop opnieuw. Sommige studenten onderstrepen belangrijke elementen (student 6) of kopiëren en plakken essentiële informatie naar onder in het Word-document (student 7).

Student 6: “Meestal onderstreep ik de belangrijkste punten van de opdracht. Gebruik de informatie, je eigen ideeën, ChatGPT.”

Student 7: “Ik ga effe de vraag hier zetten vanuit de opdracht”

Student 11: “Ik ga de opdracht opgave nog eens lezen want ik heb dat altijd nodig om dat vijf keer te lezen voordat dat fatsoenlijk bij mij komt.”

De invloed van de reeds geproduceerde tekst komt ook duidelijk naar voren in de *think aloud protocols* wanneer participanten al schrijvend hun eigen zinnen herlezen en soms de beperkingen daarvan aanvoelen en opnieuw beginnen. Onderstaande citaten illustreren dit.

Student 3: “En kan een externe visie voornamelijk. Neen. En kan een externe visie. Neen wacht. Kan de visie van familieleden helpen.”

Student 4: “Een mogelijke verklaring hiervoor is dat familie vaak een niet-intimiderende intimiderende bron is van informatie en makkelijker te raadplegen zijn. Plegen zijn. Hm. Neen ik ga dat veranderen. Mogelijke verklaring is dat familie. Dat familie vaak een hechte band hebben met de student met de toekomstige student en deze daarom makkelijk terecht kan bij.”

Ook het langetermijngeheugen drukt duidelijk een stempel door in de *think alouds*. Verschillende deelnemers denken terug aan lessen wetenschappelijk schrijven of sporen doelbewust bepaalde uitdrukkingen op (bijvoorbeeld lege woorden zoals “heel” of persoonlijk taalgebruik zoals “ik” en “wij”) waarvan hen aangeleerd is die te vermijden in een academische tekst.

Student 3: “Ja ik ga niet ‘denk ik’ schrijven want het moet academisch blijven.”

Student 7: “Maar dat is niet academisch dus dat ga ik straks nog veranderen.”

Student 8: “Uit onderzoek blijkt dat jongeren heel vaak. Niet heel vaak. Heel mag niet.”

Student 8: “In het onderzoek van de standaard zien we. Zien we dat mag niet. Oh ik heb dat geleerd. Oké. Dus wetenschappelijk geschreven.”

Student 9: “Nu ben ik aan het terugdenken naar de sessies wetenschappelijk schrijven en dit is eigenlijk mijn algemeen brede context dus nu moet ik naar mijn specifiek brede context.”

Deel 4: Discussie

Deel 4 bespreekt en interpreteert opvallende resultaten en probeert aan de hand van verklaringen de brug te maken met het theoretisch kader en de onderzoeksvragen rond het gebruik van en de houding tegenover GenAI bij (on)ervaren academische schrijvers te beantwoorden. De discussie focust zich op de beperkingen van het onderzoek (4.1), de triomf van ChatGPT (4.2), het verschil tussen (on)ervaren

schrijvers (4.3), de al dan niet kritische houding van studenten (4.4) en komt tot een conclusie over een schrijfproces waarin GenAI gebruikt wordt (4.5).

4.1 Beperkingen

De opzet van deze masterproef is verkennend van aard en heeft ook niet de omvang om diep in te gaan op alle aspecten van schrijfvaardigheid en AI. De uitgevoerde studies bevatten echter nog een reeks beperkingen.

Zo focust de analyse van de enquête zich op het verschil tussen ervaren en onervaren schrijvers maar besteedt die geen aandacht aan andere groepen (bijvoorbeeld verschil tussen mannelijke en vrouwelijke schrijvers of het gedrag van de negen niet-moedertaalsprekers). Ook zijn de statistische analyses beperkt gebleven en is er geen statistische power berekend of analyse met meerdere factoren uitgevoerd.

De opzet van de schrijfoopdracht is artificieel omdat het eindproduct niet beoordeeld wordt en de tijd tot 60 minuten gelimiteerd is. Voor de meeste academische schrijfoopdrachten is er meer tijd beschikbaar en kan het schrijfproces mogelijks meer gefaseerd verlopen. Studenten nemen andere houdingen aan aangezien het niet op punten staat en laten bijvoorbeeld meer zinnen van ChatGPT staan of ronden de opdracht sneller af. De elf respondenten van studie B vormen ook een kleine en niet-representatieve steekproef en schetsen een interessant maar beperkt beeld van mogelijke interacties. Er is geen controlegroep aanwezig die de opdracht uitvoert zonder ChatGPT.

Ten slotte zijn de methodologisch twijfels over de *think alouds* aanwezig. Gezien de moeilijkheden die student 5 en student 11 aangeven, is er wel duidelijk een cognitieve last als studenten luidop moeten redeneren. De grote individuele verschillen in aantal uitgesproken zinnen en aard daarvan versterken de opmerkingen van Ericsson & Simon (1994) en Yang (2023). Het coderen van de *think alouds* is maar door één onderzoeker gebeurd waardoor er geen controle over de beoordeling en codering aanwezig is. Door het exploratief karakter heeft niet elke uiting een code gekregen en zijn sommige uitingen onder twee overlappende categorieën gebracht.

4.2 Gebruik van GenAI (ChatGPT) in schrijfoopdrachten

Studie A toont in de eerste plaats het succes van ChatGPT, gekend bij 255 van de 258 studenten (98,83%) en gerapporteerd gebruikt door 214 van hen (82,95%). Dit hangt samen met de grote media-aandacht en naamsbekendheid die sinds november 2022 aan ChatGPT gegeven is als een soort

“pionier”. Interessanter om te bestuderen is hoeveel studenten aan de KU Leuven GenAI durven inzetten voor hun academische opdrachten en taken. Binnen deze studie rapporteren 152 studenten (58,92%) dat ze het inzetten, bestaande uit 75% van de groep *advanced writers* (45 van de 60) en 53,53% van de groep *novice writers* (107 van de 198).

Hieruit blijkt dat de eerste- en tweedebachelorstudenten minder geneigd zijn GenAI in te zetten dan hun oudere en meer ervaren tegenhangers. Dit valt misschien te verklaren door de aard van hun opleiding die in de beginjaren nog niet zoveel academische schrijftaken bevat en vooral gericht is op het reproduceren van kennis. Een alternatieve hypothese is het feit dat een groot deel van hen nog in het middelbaar onderwijs zat toen ChatGPT uitkwam en dat er daar met strenge regels geëvalueerd werd en gebruik van GenAI vaak verboden was. Ten derde zijn hun antwoorden mogelijks ook beïnvloed door sociaal wenselijk gedrag (Baarda 2015).

De meest voorkomende reden om geen GenAI te gebruiken voor schrijfp opdrachten blijft ook verbonden met die strenge regels en het oneerlijk gevoel van plagiaat of fraude te plegen of iets te doen dat “niet mag”. Uit de *think alouds* komt dit ook naar voren hoewel daarbij op te merken valt dat de deelnemers wisten dat het schrijfproduct niet beoordeeld zou worden en zich daarom misschien meer vrijheden veroorloofden (zie bijvoorbeeld student 5, 3.3.2). Hoewel de setting zoveel mogelijk het spontaan gebruik probeert te activeren, blijft het een kunstmatige situatie waarin participanten sowieso proberen om zich van hun beste kant te laten zien en mogelijk ander gedrag vertonen dan bij een echte academische paper. Een tweede opvallende motivatie voor het vermijden van GenAI is het vertrouwen dat de studenten in hun eigen schrijfkwaliteiten hebben en het plezier dat ze halen uit het produceren van hun eigen teksten. In de *think alouds* verklaart student 6 hoe dergelijke hulpmiddelen botsen met zijn zelfbeeld en ambities en stappen andere deelnemers ook soms snel over op hun eigen revisievaardigheden of ideeën omdat ze die beter of mooier vinden.

4.3 Omgang met ChatGPT bij (on)ervaren schrijvers

Binnen de groep van de 152 studenten die wel gebruikmaken van GenAI in hun schrijfp opdrachten, vertonen de ervaren en onervaren schrijvers een vrij gelijkaardig gedrag waarbij de medianen voor de strategieën uit de enquête schommelen tussen 0 (= nooit) en 2 (= soms).

De strategie die het vaakst gerapporteerd wordt in de enquête is de verduidelijking van concepten of termen ($M = 2.112$, $Mdn = 3$), gevolgd door inhoudelijk brainstormen ($M = 2.026$, $Mdn = 2$) en herschrijven/verbeteren op het vlak van stijl of spellings- en grammaticafouten (stellingen 4-6 van

tijdens schrijven bij elkaar opgeteld, $M\ 1.175$, $Mdn = 2$). De *think alouds* verrassen hier door hun bijna volledige afwezigheid van de verduidelijking van concepten (slechts 1x gecodeerd). De aard en het thema van de opdracht zijn echter herkenbaar en begrijpelijk voor studenten uit het hoger onderwijs wat ervoor zou kunnen zorgen dat ze alle gebruikte termen volledig begrijpen en geen extra uitleg nodig hebben. Brainstormen en zinnen herformuleren/verbeteren zijn wel in hoge mate aanwezig in de *think alouds* net als de zoektocht naar synoniemen en in mindere mate het genereren van een eerste versie van een zin/alinea om te helpen bij de start van het schrijfproces. Omdat fictieve referenties volstonden is de strategie van het opzoeken of samenvatten van bronnen ook logischerwijs bijna volledig afwezig in de *think alouds* en kan het gerapporteerde contrast tussen *novice* en *advanced writers* daarover niet afgetoetst worden.

De zoektocht naar synoniemen, hulp bij het bedenken van (tussen)titels en bij grammaticale kwesties waren alle drie niet opgenomen in de enquête maar zijn wel gerapporteerd of gecodeerd in beide studies en zouden dus zeker in vervolgonderzoek opgenomen en verder onderzocht moeten worden.

Voor negen van de achttien stellingen uit luik 2 van de bevraging is het verschil tussen de *novice* en *advanced writers* significant en zijn het telkens de *advanced writers* die een frequenter gebruik rapporteren. De overige negen stellingen tonen dezelfde trend zonder statistisch significantie met uitzondering van de derde stelling uit de *pre-writing* fase (zoeken en samenvatten van bronnen) en stellingen 5 en 6 uit de *post-writing* fase (een evaluatie van sterke/zwakke punten vragen en chatlogboek gebruiken) wat vaker gerapporteerd wordt door *novice writers*. Gezien de lage scores voor de *post-writing* fase worden die laatste even buiten beschouwing gelaten. Een hypothese voor de divergerende aanpak wat bronnen betreft, is de academische ervaring van de *advanced writers* die al voor de komst van AI schrijfp opdrachten hebben moeten vervullen en gewoonweg meer vertrouwd zijn met academische databanken en bibliotheken. Zij staan bijgevolg kritischer tegenover het inschakelen van GenAI voor zo'n essentieel onderdeel van wetenschappelijk onderzoek, al ligt het gemiddelde en de mediaan van de *novice writers* voor dit onderdeel ook aan de lage kant. In de bevraging geven studenten ook aan dat ze omwille van een ontoereikend referentiesysteem en onvoldoende transparantie over bronnen, ervoor kiezen om geen GenAI in te zetten in schrijfp opdrachten.

De *think alouds* bewijzen sterk dat schrijven als cognitief proces bij elk individu anders verloopt. Een korte vergelijking tussen twee ex-masterstudenten taal- en letterkunde, student 10 en student 11, toont bijvoorbeeld aan dat evenveel academische ervaring niet automatisch tot gelijkaardig gedrag

leidt. Student 10 vraagt inhoudelijke argumenten aan ChatGPT en laat elke alinea herschrijven op een “academische en vlotte manier” en voegt eveneens een bibliografie toe terwijl student 11 focust op het detecteren van schrijffouten en de hele tekst tegelijk invoert. Het schrijfproces van student 11 is ook langer, chaotischer en stressvoller van aard. Verder vermelden bijvoorbeeld ook twee participanten, student 5 en 9, dat ze ChatGPT gebruiken in complementariteit met de SAN maar maken de andere ervaren schrijvers hier geen vermelding van en blijkt ook dit vooral individueel en niet ervaring gebonden te zijn. Alle besproken verschillen tussen *novice* en *advanced writers* moeten tegen die nuance van individualiteit bekeken worden.

4.4 Houding tegenover ChatGPT bij (on)ervaren schrijvers

De resultaten uit het derde luik van studie A tonen dat de 258 respondenten een neutrale tot positieve houding tegenover GenAI aannemen. De ervaren schrijvers geven hier bij zeven van de negen stellingen opnieuw iets hogere scores en zijn dus iets positiever dan hun onervaren collega's. Wat betreft het leerverlies stellen zij zich echter negatiever op dan de *novice writers* en vrezen ze dat GenAI wel tot een leerverlies leidt, één van de enige twee stellingen waarbij het verschil ook significant is. Ook vinden zij het minder duidelijk dan *novice writers* in welke mate GenAI gebruikt mag worden maar dit verschil is niet significant.

Het verschil in mentaliteit tussen beide groepen heeft waarschijnlijk opnieuw te maken met het feit dat de ervaren schrijvers twee à drie jaar langer in academische milieus hebben doorgebracht en zelf een aantal academische jaren zonder GenAI afgerond hebben. Het feit dat net die oudere studenten met een gemiddelde van slechts 1.978/5 begrijpen hoe GenAI ingezet mag worden leunt aan bij de grote angst rond plagiaat en het overtreden van regels (zie 4.2) en toont aan dat de situatie nog niet helder is voor studenten aan de KU Leuven. De richtlijnen (zie 1.3.5) zijn ofwel nog niet voldoende gekend ofwel te vaag geformuleerd. Korte nabeschoouwingen met respondenten uit studie B en suggesties uit studie A vermelden bovendien dat *good practices* en regels sterk afhankelijk zijn van het vak en de professor en die verschillende regels verwarring veroorzaken.

Ondanks inzichten uit de literatuur (1.3.4) gaan de deelnemers niet echt akkoord met het feit dat raad vragen aan GenAI minder eng is dan raad vragen aan ouders, vrienden of medeleerlingen en krijgen ze ook niet opvallend meer vertrouwen in hun eigen schrijfvaardigheden. Ze geven aan de antwoorden niet als altijd betrouwbaar te beschouwen en bewaren dus een zekere mate van kritisch bewustzijn. In hun eigen teksten, bewaren ze ook meestal een sterke vorm van *agency* en vergelijken ze de suggesties van ChatGPT met hun eigen kennis en inzichten. Slechte ervaringen met ChatGPT kunnen ook leiden

tot het afschrijven daarvan, zoals de motivaties uit de enquête aangeven (zie 3.3.1.1) en ook het *self-efficacy* model van Zimmerman en Risemberg (1.1.2.1) voorspelt. Voor een teksten in een vreemde taal zoals het Engels kan het zijn dat hun houdingen anders liggen (zie student 11 in 3.3.2). Uit hun omgang blijkt duidelijk hoe groot de rol van mediërende tools is voor schrijven binnen de academische *community*.

4.5 ChatGPT en het schrijfproces

In dit onderzoek wordt GenAI voornamelijk ingeschakeld in de stellingen uit de *writing* en *pre-writing* fasen en is er een opvallend verschil met de *post-writing* fase waar amper beroep gedaan wordt op ChatGPT. Dit valt enerzijds ook te verklaren door de aard van de opdracht die zich met maximaal 500 woorden niet echt leent voor een samenvattend abstract/conclusie of voor een grondige evaluatie van positieve en negatieve punten aangezien zo'n korte tekst nog makkelijk zelf te overzien is. Anderzijds dient ook genuanceerd te worden dat de opdeling van de stellingen een artificiële selectie is en het herschrijven/verbeteren in andere studies wel gezien wordt als een *post-writing* actie. In navolging van Hayes (1996) kiest deze studie er echter voor om verbeteringen mee in de schrijffase in te delen omdat die vaak simultaan optreden met het schrijven zelf (zie bijvoorbeeld student 4 in 3.3.3). Ook als het niet simultaan gebeurt, blikken deelnemers soms vooruit en nemen ze zich voor later terug te komen op bepaalde aspecten (zie student 7, 8, 10), wat de ruime opvatting van de schrijffase versterkt en de cyclische aard van de componenten uit het schrijfproces bevestigt.

Student 7: “Daar moet nog een verklaring tussen. Die ga ik straks wel vinden.”

Student 8: “In het onderzoek van de standaard zien we. Zien we mag niet maar dat pas ik seffens aan of ChatGPT gaat dat aanpassen.”

Student 10: “Een behoorlijke ga ik eventjes schrijven ik kan daar nog op terugkomen dus ik markeer dat in het rood.”

Daarbovenop is een belangrijk inzicht van dit onderzoek dat het inzetten van GenAI ook de aard van het schrijfproces ingrijpend verandert. Schrijvers moeten niet langer eerst in bronnen of hun langetermijngeheugen op zoek gaan naar inhoud om die te vertalen naar een tekst en die tekst vervolgens te reviseren. Ze kunnen nu ook aan ChatGPT vragen om verklaringen te geven of zelfs een hele eerste versie van een tekst te genereren en hun eigen schrijfproces begint dan als het ware in de revisiefase wanneer ze de tekst van ChatGPT proberen te herschrijven in hun eigen woorden (zie bijvoorbeeld student 5 in 3.3.1.3 en bijlage 3).

De *think alouds* tonen bovendien dat studenten in beperkte mate plannen en vaak zonder lay-out of kladversie te werk gaan. Hier kan ChatGPT net tegengewicht bieden omdat antwoorden daarvan vaak beginnen met een lay-out of algemene tips.

Het is ten slotte belangrijk op te merken dat ChatGPT niet het enige hulpmiddel is dat ingeschakeld wordt bij het schrijven van een academische tekst. De deelnemers van studie B rekenen daarbuiten op de spellingchecker van Microsoft Word, zoeken synoniemen op in de databank van de website synoniemen.net, maken gebruik van de SAN – Schrijfhulp Academisch Nederlands en zoeken grammaticale of inhoudelijke kwesties nog steeds op het internet op. De komst van generatieve AI vervangt de voorgaande hulpmiddelen dus niet en vormt alleen een aanvulling of een stap in een uitgewerkte schrijfroutine. De gerapporteerde scores en gecodeerde strategieën tonen echter ook aan dat veel studenten nog steeds voor een groot deel op hun eigen vaardigheden rekenen wanneer ze schrijven of de output van ChatGPT grondig herschrijven en redigeren. De bezorgdheid dat academisch schrijven volledig door AI overgenomen zal worden, lijkt voorlopig dus enigszins weggenomen te kunnen worden.

Deel 5: Conclusie

Deze masterthesis voert een verkennend onderzoek uit naar de rol van generatieve artificiële intelligentie binnen het academisch schrijfproces van studenten aan de KU Leuven. Het theoretisch kader (deel 1) schetst een achtergrond rond schrijfvaardigheid, academisch schrijven en technologie. In de eerste plaats presenteert het enkele invloedrijke schrijfmodellen binnen cognitieve, sociaal-cognitieve en sociaal-culturele paradigma die een terminologisch houvast bieden. Ten tweede bespreekt het theoretische luik hoe technologische innovaties de aard van het schrijfproces beïnvloeden en welke categorieën van elektronische schrijfondersteuning bestaan. Het theoretische luik sluit af met een introductie van GenAI en de richtlijnen van de KU Leuven rond het verantwoord gebruik daarvan.

Dat verantwoord gebruik wordt vervolgens onderzocht in een tweeledige studie (deel 2 en 3). Enerzijds worden algemene bevindingen en gebruik van GenAI afgetoetst met een enquête afgenomen bij 258 studenten sociale wetenschappen en letteren (studie A). Anderzijds bieden elf uitgevoerde schrijfoopdrachten in interactie met ChatGPT aan de hand van *think aloud protocols* een verdieping en vergelijking van de gerapporteerde strategieën (studie B). De discussie (deel 4) bespreekt en verklaart de resultaten.

Uit Studie A blijkt dat een aanzienlijke meerderheid van de studenten bekend is met GenAI en het regelmatig inzet in hun schrijfwerk. De enquête toont aan dat studenten GenAI voornamelijk gebruiken in de *pre-writing* fase en *writing*-fase, bijvoorbeeld om concepten uit te leggen, te brainstormen en eerste versies grammaticaal of academisch te verrijken. Na het schrijven hun teksten laten evalueren of samenvatten worden beduidend minder gerapporteerd. Ondanks deze toepassingen, blijft er een zekere mate van voorzichtigheid bestaan onder de studenten wat de betrouwbaarheid en academische integriteit van de gegenereerde inhoud aangaat.

In Studie B verschaffen de *think aloud*-protocollen aanvullende inzichten en tonen ze dat studenten bijvoorbeeld ook synoniemen opzoeken met ChatGPT en soms eerste versies van de hele opdracht of een inleiding/slot laten genereren. Het gedrag verschilt hier zeer sterk van individu tot individu en sluit ook niet volledig aan bij het gerapporteerd gedrag uit de enquête. Door de aard van de opdracht worden bijvoorbeeld geen verklaringen voor concepten gezocht of samenvattingen gevraagd. Een eerste belangrijk inzicht uit deze studie is dat de volgorde van de componenten uit het schrijfproces zoals gepresenteerd door Flower & Hayes (1980) verandert nu studenten niet zelf alle inhoud moeten bedenken en uit hun langetermijngeheugen moeten halen maar ook beroep kunnen doen op een eerste versie geschreven door GenAI die zij dan als het ware reviseren en corrigeren. Een tweede belangrijk inzicht is dat de noodzaak voor sterke schrijfvaardigheden niet vermindert.

Om het retorisch probleem in een prompt om te zetten, moeten studenten de juiste informatie uit een opdracht kunnen filteren en vervolgens goed geformuleerde instructies aan ChatGPT geven. Frustraties, gefaalde interacties en vage *prompts* (bijvoorbeeld “herschrijf”) uit de *think alouds* tonen aan dat die vaardigheden nog niet sterk aanwezig zijn bij de elf respondenten en dat het hoger onderwijs hier sterk op moet inzetten als ze vaardige digitale schrijvers af wil leveren. Ook de angst rond plagiaat en academische integriteit die heerst, moet weggenomen worden met nog duidelijkere of wijder verspreide richtlijnen.

Verder onderzoek zou zich kunnen buigen over *prompt engineering* en aan de hand van initiatiesessies kunnen vergelijken of het schrijfproces en de teksten kwalitatiever worden als studenten hier meer op getraind worden. Een uitgebreidere schrijfopdracht zou een bronnenzoektocht of langere tekst kunnen betrekken en een langere duur kunnen voorzien. Ook zou een verdere verkenning van het gedrag en de houding van studenten van andere faculteiten of instellingen boeiend zijn. Een laatste suggestie is een studie van GenAI voor tweedetaalverwerving en de zoektocht naar hoe

Nederlandstalige studenten bijvoorbeeld voor Engelse teksten voordeel kunnen halen uit ChatGPT. De technologie staat nooit stil en dit onderzoek uit maart 2024 zal dus spoedig aangevuld moeten worden met nieuwe inzichten en gegevens. Hier is ten slotte een krachtige slotzin van deze masterproef gegenereerd door ChatGPT:

“Door een zorgvuldige en kritische integratie van generatieve artificiële intelligentie in het academisch schrijfonderwijs, kunnen we een nieuwe generatie studenten voorbereiden die niet alleen technisch vaardig is, maar ook beschikt over de intellectuele diepgang en kritische denkvaardigheden die essentieel zijn voor de academische en professionele uitdagingen van de toekomst.” (ChatGPT)

Bibliografie

Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). Computer-Based Writing Instruction. In C. A. McArthur, S. Graham, & J. Fitzgerakd (Red.), *Handbook of writing research* (second edition, pp. 316-329). The Guilford Press.

Anghelescu, A., Ciobanu, I., Munteanu, C., Anghelescu, L. A. M., & Onose, G. (2023). ChatGPT: ‘To be or not to be’ ... in academic research. The human mind’s analytical rigor and capacity to discriminate between AI bots’ truths and hallucinations. *Balneo and PRM Research Journal*, 14(Vol.14, 4), 614. <https://doi.org/10.12680/balneo.2023.614>

Baarda, B., de Goede, M. P. M., & Kalmijn, M. (2015). *Basisboek enquêteren: handleiding voor het maken van een vragenlijst en het voorbereiden en afnemen van enquêtes*. (Vierde druk). Noordhoff Uitgevers.

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>

Beck, S. W., & Levine, S. R. (2023). Backtalk: ChatGPT: A powerful technology tool for writing instruction. *Phi Delta Kappan*, 105(1), 66-67. <https://doi.org/10.1177/00317217231197487>

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Erlbaum.

Berckmoes, D., & Rombouts, H. (2009). Intern rapport verkennend onderzoek naar knelpunten taalvaardigheid in het hoger onderwijs. Linguapolis/Universiteit Antwerpen.

Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). Estimating the environmental impact of Generative-AI services using an LCA-based methodology. *Procedia CIRP*, 122, 707–712. <https://doi.org/10.1016/j.procir.2024.01.098>

Bonset, H. (2010). Deel 1. Nederlands in voortgezet en hoger onderwijs: hoe sluit dat aan? *Levende Talen Magazine*, 3, 16-21.

Cummings, L. (2023). Writing Processes in the Digital Age: A Networked Interpretation. In O. Kruse, C. Rapp, C. M. Anson, K. Benetos, E. Cotos, A. Devitt, & A. Shibani (Red.), *Digital Writing Technologies in Higher Education* (pp. 485-497). Springer International Publishing. https://doi.org/10.1007/978-3-031-36033-6_30

Day, T. (2023). A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT. *The Professional Geographer*, 75(6), 1024–1027.
<https://doi.org/10.1080/00330124.2023.2190373>

De Bakker, I., Meuffels, B., Muris, M., & Rietstap, E. (2015). Diagnose en remediëring van schrijfvaardigheidsproblemen bij eerstejaarsstudenten Geesteswetenschappen. *Tijdschrift voor Taalbeheersing*, 37(2), 127-157. <https://doi.org/10.5117/TVT2015.2.BAKK>

De Wachter, L., & d’Hertefelt, M. (2018). Academic Writing Assistant – Effective and process-oriented writing support. *JoSch – Journal der Schreibberatung*, 23-32.

De Wachter, L., Fivez, K., & Van Soom, C. (2021). *Academisch schrijven: een praktische gids* (zevende druk). Acco.

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12402-3>

Englert, C.S., Mariage, T.V., Dunsmore, K. (2006). Tenets of sociocultural theory in writing instruction research. In C. A. McArthur, S. Graham, & J. Fitzgerald (Red.), *Handbook of writing research* (first edition, pp. 54-66). The Guilford Press.

Ericsson, K. A., & Simon, H. A. (1996). *Protocol analysis: verbal reports as data* (Rev. ed., 2nd pr.). MIT press.

Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365-387. <https://doi.org/10.2307/356600>

Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, Diagnosis, and the Strategies of Revision. *College Composition and Communication*, 37(1), 16.
<https://doi.org/10.2307/357381>

Goosen, M., & Schoordijk, F. (2017). *Basisvaardigheden academisch schrijven* (eerste druk, derde oplage). Coutinho.

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Red.), *Handbook of educational psychology* (pp. 457-477). Erlbaum.

Green, A. (1998). *Verbal Protocol Analysis in Language Testing Research: a Handbook*. Cambridge University Press.

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J. F., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenerg (Red.), *Advances in Applied Psycholinguistics II: reading, writing and language processes* (pp. 176-240). Cambridge University Press.

Hayes, J.R. (1996) A new framework for understanding cognition and affect in writing. In C.M Levy, Ransdell S. (Red.) *The science of writing* (pp. 1-27). Erlbaum.

Hayes, J. R. (2012). Modeling and Remodeling Writing. *Written Communication*, 29 (3), 369-388.

<https://doi.org/10.1177/0741088312451260>

Hwang, S. I., Lim, J. S., Lee, R. W., Matsui, Y., Iguchi, T., Hiraki, T., & Ahn, H. (2023). Is ChatGPT a “Fire of Prometheus” for Non-Native English-Speaking Researchers in Academic Writing? *Korean Journal of Radiology*, 24(10), 952. <https://doi.org/10.3348/kjr.2023.0773>

Hyland, K. (2009). *Academic Discourse: English in a Global Context*. Continuum International Publishing Group.

Jacob, S. R., Tata, T., & Warschauer, M. (2023). Emergent AI-assisted discourse: Case study of a second language writer authoring with ChatGPT. *ArXiv.Org*.

<https://doi.org/10.48550/arxiv.2310.10903>

Karakose, T. (2023). The Utility of ChatGPT in Educational Research—Potential Opportunities and Pitfalls. *Educational Process International Journal*, 12(2). <https://doi.org/10.22521/edupij.2023.122.1>

KU Leuven. (2024, 15 februari). Verantwoord gebruik van Generatieve Artificiële Intelligentie.

Student aan KU Leuven. Geraadpleegd op 28 mei 2024, van

<https://www.kuleuven.be/onderwijs/student/onderwijstools/artificiele-intelligentie>

Kulkarni, A., Shivananda, A., Kulkarni, A., & Gudivada, D. (2023). *Applied Generative AI for Beginners: Practical Knowledge on Diffusion Models, ChatGPT, and Other LLMs*. Apress.

<https://doi.org/10.1007/978-1-4842-9994-4>

Lei, X. (2008). Exploring a sociocultural approach to writing strategy research: Mediated actions in writing activities. *Journal of Second Language Writing*, 17(4), 217-236.

<https://doi.org/10.1016/j.jslw.2008.04.001>

Lei, X. (2016). Understanding writing strategy use from a sociocultural perspective: The case of skilled and less skilled writers. *System*, 60, 105-116. <https://doi.org/10.1016/j.system.2016.06.006>

Levshina, N. (2015). *How to do linguistics with R*. John Benjamins Publishing Company.

McArthur, C., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. McArthur, S. Graham, & J. Fitzgerald (Red.), *Handbook of writing research* (second edition, pp. 24-41). The Guilford Press.

McGowan, A., Gui, Y., Dobbs, M., Shuster, S., Cotter, M., Selloni, A., Goodman, M., Srivastava, A., Cecchi, G. A., & Corcoran, C. M. (2023). ChatGPT and Bard exhibit spontaneous citation fabrication during psychiatry literature search. *Psychiatry Research*, 326, 115334.

<https://doi.org/10.1016/j.psychres.2023.115334>

OpenAI (2022, november). Introducing ChatGPT. *OpenAI Blog*. Geraadpleegd 17 januari 2023, van

<https://openai.com/blog/chatgpt/>

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: the nature of constructively responsive reading*. Erlbaum.

Prior, P. (2006). A sociocultural theory of writing. In C. A. McArthur, S. Graham, & J. Fitzgerald (Red.), *Handbook of writing research* (first edition, pp. 54-66). The Guilford Press.

Ray, 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>

Shadiev, R., & Feng, Y. (2023). Using automated corrective feedback tools in language learning: A review study. *Interactive Learning Environments*, 1-29. <https://doi.org/10.1080/10494820.2022.2153145>

Snow, C. E. (2010). Academic Language and the Challenge of Reading for Learning About Science. *Science*, 328(5977), 450-452. <https://doi.org/10.1126/science.1182597>

Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33-48. <https://doi.org/10.1016/j.compedu.2018.12.005>

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. <https://doi.org/10.1016/j.asw.2023.100752>

Turner, J. (2011). *Language in the academy: Cultural reflexivity and intercultural dynamics*. Multilingual Matters.

Van De Mierop, D. (2020). Nederlandse taalkunde III: Taalbeheersingsonderzoek. Acco.

Van den Branden, K. (2010). Taalbeleid in het hoger onderwijs. In Peters, E. & Van Houtven, T. (Eds.), *Taalbeleid in het hoger onderwijs: de hype voorbij?* (pp. 213-223). Acco.

Vandooren, L. (2022). Studenten en de Schrijfhulp: de juiste klik? [Ongepubliceerde masterproef]. KU Leuven

Van Kalsbeek, A., & Kuiken, F. (2014). Academisch taalgebruik in het Nederlandse wetenschappelijk onderwijs: Stand van zaken en blik op de toekomst. *Internationale Neerlandistiek*, 52(3), 221-236. <https://doi.org/10.5117/IN2014.3.KALS>

Van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: a practical guide to Modelling cognitive processes*. Academic press.

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, 101071. <https://doi.org/10.1016/j.jslw.2023.101071>

Yang, C., & Zhang, L. J. (2023). *Think-Aloud Protocols in Second Language Writing: A Mixed-Methods Study of Their Reactivity and Veridicality* (1st ed. 2023.). Springer. <https://doi.org/10.1007/978-3-031-39574-1>

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a Self-Regulated Writer: A Social Cognitive Perspective. *Contemporary Educational Psychology*, 22(1), 73-101. <https://doi.org/10.1006/ceps.1997.0919>

Bijlagen

Bijlage 1: Enquête uit studie (A)

Sectie 0: Metadata

- Mijn leeftijd is: ... (vul een getal in cijfers in)
- Ik identificeer me als:
 - Man
 - Vrouw
 - Anders
 - Zeg ik liever niet
- Mijn nationaliteit is:
 - Belg
 - Anders, namelijk: ...
- Mijn moedertaal is:
 - Nederlands
 - Anders, namelijk: ...
- Mijn hoogst behaalde diploma is:
 - Middelbare school
 - Professionele bachelor
 - Academische bachelor
 - Academische master
- Mijn huidige studierichting is: ...
- Mijn huidige studiefase is: (als je verschillende fases combineert, duid dan de fase aan waarvan je de meeste studiepunten opneemt)
 - 1^e bachelor
 - 2^e bachelor
 - 3^e bachelor
 - Master
 - Schakelprogramma
 - Andere, namelijk ...
- Dit is mijn ... jaar aan een instelling van het hoger onderwijs (een hogeschool of universiteit)
 - Eerste
 - Tweede
 - Derde
 - Vierde
 - Vijfde of meer

Sectie 1: Gebruik en frequentie?

- Welke generatieve AI-tools die ingezet kunnen worden tijdens het schrijfproces ken je bij naam? Selecteer:
 - ChatGPT
 - Microsoft Copilot
 - Google bard/ Gemini
 - Bing AI
 - Chatsonic
 - Jasper AI
 - Perplexity AI
 - Quillbott AI

- Geen van bovenstaande
- Andere, namelijk:
- Hoe vaak gebruik jij volgende generatieve AI-tools gemiddeld gezien? Zet de schuiver op het corresponderend getal waarbij 0 = nooit/ 1 = om de paar maanden/ 2 = maandelijks/ 3 = wekelijks/ 4 = meerdere keren per week.
 - ChatGPT
 - Microsoft Copilot
 - Google bard/ Gemini
 - Bing AI
 - Chatsonic
 - Claude AI
 - Jasper AI
 - Perplexity AI
 - Quillbott AI
 - Nog andere, namelijk:
- Heb je generatieve AI-tools als gebruikt voor een academische schrijfo opdracht (een paper, essay, ...)
 - Ja/nee
 - Indien neen: Je gaf in de vorige vraag aan dat je geen gebruik maakt van generatieve AI voor academische schrijfo opdrachten. Kun je die keuze kort toelichten en uitleggen waarom niet? ... → Sectie 4
 - Indien ja → sectie 3

Sectie 2: schrijfstrategieën en gebruik van GenAI in academische opdrachten

- Je gaf in de vorige vraag aan dat je gebruikmaakt/hebt gemaakt van generatieve AI voor academische schrijfo opdrachten. Welke tools heb je hiervoor gebruikt/gebruik je hiervoor?
 - ChatGPT
 - Microsoft Copilot
 - Google bard/ Gemini
 - Bing AI
 - Chatsonic
 - Claude AI
 - Jasper AI
 - Perplexity AI
 - Quillbott AI
 - Nog andere, namelijk:
- Ik gebruik generatieve AI (zoals ChatGPT) in schrijfo opdrachten voor het schrijven om... [matrixtabel met Nooit – Zelden – Soms – Vaak – Altijd]
 - Tips te vragen over hoe ik een schrijfo opdracht best aanpak
 - Te brainstormen over mijn onderwerp en inspiratie te krijgen
 - Bronnen te zoeken of bronnen samen te vatten
 - Moeilijke concepten uit te leggen of te definiëren
 - Dit is een controlevraag: duid hier “soms” aan
 - Een tekstplan te genereren
 - Feedback te krijgen op een tekstplan dat ik zelf opstelde
 - Op gang te komen en niet met een leeg blad te moeten beginnen
 - Andere, namelijk: ...
- Ik gebruik generatieve AI (zoals ChatGPT) in schrijfo opdrachten voor het schrijven om... [matrixtabel met Nooit – Zelden – Soms – Vaak – Altijd]
 - Mogelijke voorbeeldzinnen te genereren

- Een eerste versie van een tekst/paragraaf te genereren
- Een lijst academische bewoordingen en constructies te genereren
- Een (aantal) zin(nen) die ik schreef te herformuleren
- De stijl van mijn tekst te verbeteren en academischer te maken
- De grammatica- en/of spelfouten uit mijn tekst te halen
- Andere, namelijk: ...
- Ik gebruik generatieve AI (zoals ChatGPT) in schrijfopdrachten voor het schrijven om...
[matrixtabel met Nooit – Zelden – Soms – Vaak – Altijd]
 - Na te kijken of mijn tekst de juiste inhoud overbrengt
 - Een tekst die ik schreef in te korten of langer te maken
 - Mijn tekst samen te vatten (bijvoorbeeld voor een abstract/conclusie)
 - Mijn tekst te evalueren door een cijfer toe te kennen
 - Mijn tekst te evalueren door goede en sterke punten te geven
 - Mijn chatlogboek te herbekijken en te zien welke strategieën vorige keer gegeven werden
 - Dit is een controlevraag. Duid hier "Altijd" aan
 - Andere, namelijk:

Sectie 3: Houding ten opzichte van Generatieve AI

- Duid aan in welke mate je akkoord gaat met volgende stellingen
 - Ik maak mij zorgen over de manier waarop generatieve AI omgaat met privacy en gegevensbewaring
 - Ik vind generatieve AI een handig hulpmiddel voor studenten
 - Ik vrees dat generatieve AI op lange termijn leidt tot leerverlies van studenten
 - Het is duidelijk in welke mate ik gebruik mag maken van generatieve AI voor mijn academische opdrachten
 - Het gebruiken van generatieve AI voor schrijfopdrachten voelt als het plegen van plagiaat
 - Ik beschouw de antwoorden van generatieve AI als betrouwbaar
 - Ik kan mijn schrijftaken beter afronden als ik gebruik kan maken van generatieve AI
 - Raad vragen aan ChatGPT vind ik minder eng dan raad vragen aan een ouder/vriend/klasgenoot/...
- Indien je graag extra toelichting geeft bij een antwoord of je eigen gebruik van generatieve Artificial Intelligence, dan is dat hier mogelijk: ...

Bijlage 2: Schrijfopdracht uit studie (B)

Schrijf een korte argumentatieve tekst (max 500 woorden) waarin je de factoren verklaart die van invloed zijn op de studiekeuze van jongeren na het middelbaar onderwijs. Gebruik de informatie uit onderstaande grafiek (*grafiek 1*), je eigen ideeën en ChatGPT wanneer gewenst. Produceer de tekst in een academische stijl en zet (REF) op de plaatsen waar je een bronvermeldingen zou toevoegen (het is niet nodig om echt bronnen te zoeken!). Je mag zeker zelf redenen en argumenten bedenken. Verwoord je gedachten luidop en spreek uit wat je denkt tijdens het schrijven. Om vertrouwd te raken met deze *think aloud* techniek, zullen we eerst een korte oefenronde doen.

Figuur 10: Beïnvloedende factoren studiekeuze (in absolute aantallen)



Bijlage 3: Alle resultaten enquête uit studie (A)

Apart bijgevoegd

Bijlage 4: teksten, chatgeschiedenis en think alouds studie (B)

Apart bijgevoegd