De rol van AI-geletterdheid en prestatieverwachting bij de acceptatie en het gebruik van AI-systemen in de publieke sector

Anne-Fleur Kremer - 556748ak

Erasmus School of Social and Behavioural Sciences

Erasmus Universiteit Rotterdam

Begeleider: Prof. dr. Wolfgang Ebbers

Tweede beoordelaar: Dr. ir. Lizet Kuitert

25 juni 2025

Aantal woorden: 10080

Abstract

Het gebruik van AI-systemen binnen de publieke sector biedt kansen, maar brengt ook risico's met zich mee. Om de kansen van AI-systemen te kunnen benutten en de daarmee verbonden risico's te mitigeren, is het belangrijk dat ambtenaren het nut zien van de technologie bij wat ze moeten doen (prestatieverwachting) en tegelijkertijd ook kennis hebben over het gebruik van AI-systemen op een verantwoorde manier (AI-geletterdheid). Onderzoek naar AI-geletterdheid is nog in een beginstadium en is zodoende voornamelijk exploratief en kwalitatief van aard. Daarnaast is er in algemene zin nog weinig onderzoek gedaan naar de acceptatie van AI-systemen binnen de publieke sector. Het hoofddoel van dit kwantitatieve onderzoek was het toetsen van hypothesen over de invloed van AI-geletterdheid op het construct prestatieverwachting binnen het UTAUT-model om inzicht te krijgen in de determinanten van de acceptatie en het gebruik van AI-systemen in de publieke sector. De data voor dit onderzoek werd verzameld door middel van een online survey onder 278 ambtenaren en werd hoofdzakelijk geanalyseerd met behulp van PROCESS. De analyse laat een aantal belangrijke bevindingen zien. Ten eerste, blijkt dat AI-geletterdheid positief gerelateerd is aan gebruikersintentie en dat deze relatie gedeeltelijk wordt verklaard door prestatieverwachting. Ten tweede, blijkt gebruikersintentie positief gerelateerd te zijn aan gebruikersgedrag. Ten derde, laat dit onderzoek zien dat AI-geletterdheid in ieder geval tweeledig is met enerzijds een technisch/praktisch aspect en anderzijds een ethisch aspect. Daarmee weerspreken de resultaten van dit onderzoek het theoretische idee dat AI-geletterdheid bestaat uit vier dimensies (bewustzijn, gebruik, evaluatie en ethiek). De bevindingen van dit onderzoek wijzen op het belang van AI-geletterdheid en prestatieverwachting bij de acceptatie en het gebruik van AI-systemen in de publieke sector. Overheidsorganisaties worden aanbevolen om, in lijn met de verplichting vanuit de AI-verordening, de AI-geletterdheid onder haar werknemers te bevorderen. Hierbij moet in ieder geval aandacht worden besteed aan zowel de technische/praktische kant van AI-geletterdheid, alsook de ethische kant.

Keywords: artificiële intelligentie (AI), AI-geletterdheid, prestatieverwachting, technologische acceptatie, UTAUT, AI-acceptatie, publieke sector

Inhoudsopgave

1 Inleiding

In hoofdstuk 1 wordt ingegaan op de aanleiding van dit onderzoek, de probleemstelling, de vraagstelling en de wetenschappelijke en maatschappelijke relevantie.

1.1 Aanleiding

Het gebruik van AI-systemen binnen organisaties neemt toe. Waar in 2023 door 13% van de Nederlandse ondernemingen gebruik werd gemaakt van AI-systemen, was dit percentage in 2024 al gestegen naar 23% (Eurostat, 2024). Ook binnen de publieke sector neemt het gebruik van AI-systemen toe. In 2024 werden er binnen de publieke sector maar liefst 266 AI-toepassingen gevonden, terwijl dit er in 2019 nog 75 waren (TNO, 2024). Er bestaan verschillende opvattingen over wat AI-systemen zijn. De Europese Unie hanteert een heel algemene definitie: “een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen” (*Verordening (EU) 2024/1689*, 2024). Ook generatieve AI valt onder deze definitie. Generatieve AI is een vorm van AI waarmee nieuwe data wordt gecreëerd op basis van geleerde patronen uit bestaande data, waarbij het er sterk op lijkt dat deze data door de mens zelf is gegenereerd (Banh & Strobel, 2023). Hoewel AI zich niet beperkt tot generatieve AI, is er door de opkomst van generatieve AI momenteel massale belangstelling voor AI en is het directe gebruik van AI-systemen veel wijdverspreider geworden.

Om deze toename in het gebruik van AI-systemen in goede banen te leiden, is op 1 augustus 2024 de AI-verordening in werking getreden (Europees Parlement, 2021; Europese Commissie, 2024). De AI-verordening is een Europese wet over de ontwikkeling en het gebruik van Kunstmatige Intelligentie (AI) binnen de Europese Unie. De doelstelling van de AI-verordening is het stimuleren van het gebruik van betrouwbare AI. Een van de bepalingen uit de AI-verordening is dat organisaties die AI-systemen ontwikkelen of gebruiken, vanaf 2 februari 2025 moeten inzetten op de AI-geletterdheid van haar werknemers (*Verordening (EU) 2024/1689*, 2024). In de AI-verordening wordt AI-geletterdheid gedefinieerd als: “vaardigheden, kennis en begrip die aanbieders, gebruiksverantwoordelijken en betrokken personen ... in staat stellen geïnfor-

meerd AI-systemen in te zetten en zich bewuster te worden van de kansen en risico's van AI en de mogelijke schade die zij kan veroorzaken". In overweging 20 van de AI-verordening staat verder over het benodigde niveau aan AI-geletterdheid dat: "kennis kan verschillen naargelang de context en kan inzicht omvatten in de correcte toepassing van technische elementen tijdens de ontwikkelingsfase van het AI-systeem, de maatregelen die moeten worden toegepast tijdens het gebruik ervan, hoe de output van het AI-systeem moet worden geïnterpreteerd, en, in het geval van betrokken personen, de kennis die nodig is om te begrijpen hoe beslissingen die met behulp van AI worden genomen, op hen van invloed zullen zijn".

1.2 Probleemstelling

Het gebruik van AI-systemen binnen de publieke sector kan voordelen bieden, zoals een verhoogde efficiëntie en kostenbesparingen (Al Naqbi et al., 2024). Daarnaast kan het bijdragen aan slimme oplossingen voor maatschappelijke opgaven. Bij de Provincie Zuid-Holland wordt bijvoorbeeld gebruikgemaakt van AI beeldherkenning voor het analyseren van satellietbeelden om natuurgebieden die gevoelig zijn voor stikstof in kaart te brengen, om vervolgens maatregelen te kunnen nemen wanneer zich veranderingen in het natuurgebied voordoen (Het algoritmeregister, 2024). Het gebruik van AI-systemen is echter ook niet zonder risico. Eén van de kenmerken die het gebruik van (onder andere) generatieve AI-systemen risicovol maakt is algoritmische vooringenomenheid (Akter et al., 2021). Algoritmische vooringenomenheid houdt in dat een algoritme vooroordelen bevat waardoor het tot onjuiste output kan komen. Deze vooringenomenheid ontstaat onder andere doordat algoritmes ontworpen en getraind worden door mensen die zelf ook vooringenomen zijn. Daarnaast wordt er steeds vaker gebruikgemaakt van AI-systemen die dusdanig ingewikkeld zijn dat niet meer te begrijpen valt hoe deze AI-systemen tot hun output komen, de zogeheten 'black box van AI', (Arrieta et al., 2020). Wanneer de output van dit soort AI-systemen wordt gebruikt voor het maken van beslissingen, leidt dit tot beslissingen die niet uit te leggen zijn.

Hoewel het gebruik van AI-systemen dus niet zonder risico is, wordt het gebruik van AI-systemen (met uitzondering van hoog-risico AI-systemen) niet verboden (*Verordening (EU) 2024/1689*, 2024). In plaats daarvan wordt door middel van de AI-verordening ingezet op het stimuleren van verantwoord gebruik van AI-systemen. Dit heeft alles te maken met de eerder genoemde kansen die AI-systemen bieden. Voorwaardelijk voor het benutten van de kansen die AI-systemen bieden is dan wel dat werknemers kennis hebben over het gebruik van

AI-systemen op een verantwoorde manier (AI-geletterdheid) en dat zij denken dat deze technologie nuttig is bij wat ze moeten doen (prestatieverwachting). In dat kader is het interessant om aan de hand van een specifiek onderdeel van het *Unified Theory of Acceptance and Use of Technology-model* (UTAUT) van Venkatesh et al. (2003), namelijk het construct prestatieverwachting, te onderzoeken in hoeverre het begrip dat werknemers van AI-systemen hebben (AI-geletterdheid) en de mate waarin zij denken dat deze technologie nuttig is bij wat ze moeten doen (prestatieverwachting) in relatie staat tot de acceptatie en het gebruik van AI-systemen.

1.2.1 Doelstelling

De doelstelling van dit onderzoek is het toetsen van hypothesen over de invloed van AI-geletterdheid op het construct prestatieverwachting binnen het UTAUT-model, door middel van een kwantitatieve survey-analyse onder werknemers in de publieke sector, om inzicht te krijgen in de determinanten van de acceptatie en het gebruik van AI-systemen.

1.2.2 Onderzoeksvraag en deelvragen

Op basis van bovenstaande doelstelling is de centrale vraag in dit onderzoek: In hoeverre hebben AI-geletterdheid en prestatieverwachting invloed op de acceptatie en het gebruik van AI-systemen? Om antwoord te geven op deze hoofdvraag zijn een aantal deelvragen opgesteld.

Theoretische deelvragen:

1. Wat zegt de bestaande theorie over UTAUT (in relatie tot AI-systemen)?
2. Wat zegt de bestaande theorie over AI-geletterdheid?

Empirische deelvraag:

1. In hoeverre is prestatieverwachting van invloed op de acceptatie en het gebruik van AI-systemen?
2. In hoeverre is AI-geletterdheid van invloed op de acceptatie en het gebruik van AI-systemen?

1.3 Wetenschappelijke relevantie

Sinds de publicatie van Long & Magerko (2020) waarin zij het concept AI-geletterdheid hebben gedefinieerd, is er een enorme toename aan publicaties over AI-geletterdheid. Een zoek-

opdracht naar (“ai-literacy” AND “artificial intelligence literacy”) in SCOPUS laat zien dat waar er in 2020 slechts 11 artikelen werden gepubliceerd over AI-geletterdheid, dit in 2024 al 542 artikelen waren. Doordat AI-geletterdheid een relatief nieuw concept is, is een groot deel van deze artikelen exploratief van aard waarbij gebruik wordt gemaakt van een kwalitatieve benadering. Dit onderzoek gebruikt daarentegen een kwantitatieve benadering. Daarnaast is het bestaande onderzoek naar AI-geletterdheid hoofdzakelijk uitgevoerd binnen het onderwijs en is er nog niks bekend over AI-geletterdheid onder werknemers in de publieke sector. Tot slot, draagt dit onderzoek in algemene zin bij aan bestaand onderzoek over de acceptatie en het gebruik van AI-systemen.

1.4 Maatschappelijke relevantie

Naast dat dit onderzoek inspringt op een gat in de wetenschappelijke literatuur, is dit onderzoek maatschappelijk ook heel relevant. Hoewel de verplichting vanuit de AI-verordening om AI-geletterdheid onder werknemers te stimuleren voor iedere organisatie geldt (*Verordening (EU) 2024/1689*, 2024), is AI-geletterdheid misschien wel extra belangrijk voor overheidsorganisaties zoals de Provincie Zuid-Holland, vanwege haar verantwoordelijkheid in het beschermen van de grondrechten van burgers. Door AI-geletterdheid te onderzoeken in relatie tot de acceptatie en het gebruik van AI-systemen, kan inzicht worden verkregen in hoeverre er sprake is van geïnformeerd AI-gebruik waardoor verantwoorde inzet van AI-systemen mogelijk is.

1.5 Leeswijzer

De structuur van deze scriptie is als volgt. In hoofdstuk 2 staat het theoretisch kader centraal met een uiteenzetting van de belangrijkste theoretische concepten. In hoofdstuk 3 wordt de methodologie besproken, inclusief beschrijving van het onderzoeksontwerp, wijze van steekproefselectie, beschrijving van de steekproef, procedure, meetinstrumenten en analyses. In hoofdstuk 4 worden de belangrijkste bevindingen gerapporteerd en wordt er ook stilgestaan bij de validiteit en betrouwbaarheid van het onderzoek. In hoofdstuk 5 wordt allereerst antwoord gegeven op de onderzoeksvraag. Tot slot, wordt in dit hoofdstuk het onderzoek ook bediscussieerd.

2 Theoretisch kader

In dit theoretisch kader worden de belangrijkste theorieën en concepten die in dit onderzoek centraal staan uiteengezet. De verantwoording voor het literatuuronderzoek dat hieraan voorafging is te vinden in bijlage A. Het hoofdstuk begint met het concept technologische acceptatie. Daarna volgt het concept AI-geletterdheid. Het hoofdstuk sluit af met het conceptueel model en de daarop gebaseerde hypothesen.

2.1 Technologische acceptatie

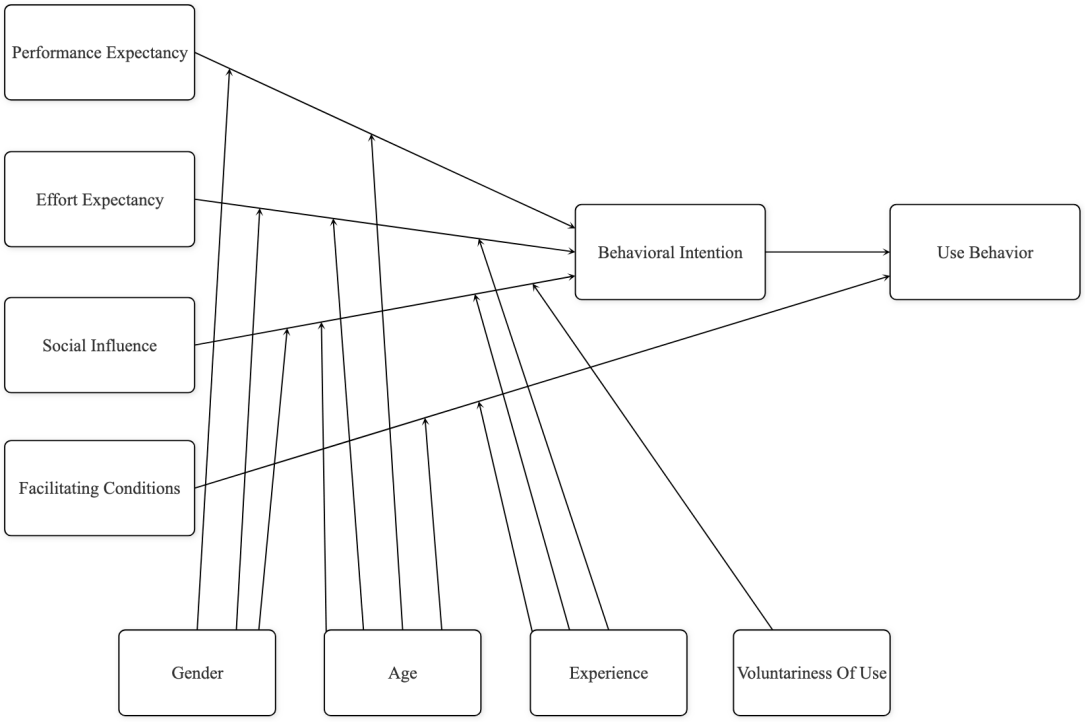
Er zijn verschillende theoretische modellen over de acceptatie en het gebruik van nieuwe technologie. Het *Technology Acceptance Model* (TAM) van Davis (1989) is het eerste theoretische model dat probeert te verklaren waarom gebruikers een nieuwe technologie accepteren en gebruiken. Het model is ontwikkeld omdat nieuwe technologie volgens Davis (1989) de potentie heeft om werkprocessen substantieel te verbeteren, maar dat daarvoor acceptatie en gebruik door beoogde gebruikers cruciaal is. Twee centrale concepten in het model zijn *perceived usefulness* en *perceived-ease-of-use*. Perceived usefulness (waargenomen nut) gaat over of gebruikers denken dat de nieuwe technologie nuttig is voor wat ze moeten doen. Perceived-ease-of-use (waargenomen gebruiksgemak) gaat over de mate waarin gebruikers denken dat de nieuwe technologie makkelijk te gebruiken is. Volgens het TAM heeft het waargenomen nut directe invloed op gebruikersintentie en is het waargenomen gebruiksgemak, via het waargenomen nut, van indirecte invloed op gebruikersintentie. Gebruikersintentie is volgens het model op zijn beurt van directe invloed op daadwerkelijk gebruikersgedrag.

Het TAM staat niet op zichzelf, maar bouwt voort op de *Theory of reasoned action* (TRA). De TRA gaat niet specifiek over de acceptatie en het gebruik van technologie, maar over hoe attituden en subjectieve normen invloed hebben op gebruikersintentie en daaropvolgend gedrag (Fishbein & Ajzen, 1975). Net als dat het TAM voortbouwt op de TRA, zijn er in navolging van het TAM een veelvoud aan vergelijkbare modellen ontwikkeld die nog beter proberen te verklaren hoe individuen tot (technologische) acceptatie komen, zoals het *Motivational Model* (MM), de *Theory of Planned Behaviour* (TPB), het *Model of PC Utilisation* (MPCU), *Innovation Diffusion Theory* (IDT) en de *Social Cognitive Theory* (SCT) (Venkatesh et al., 2003). Op basis van deze grote hoeveelheid aan acceptatiemodellen, heeft Venkatesh et al. (2003) de *Unified Theory of Acceptance and Use of Technology* (UTAUT) ontwikkeld. UTAUT combineert

elementen van maar liefst acht eerdere acceptatiemodellen en kan daarmee 70% aan variantie in gebruikersintentie verklaren. Twee belangrijke concepten uit UTAUT zijn *performance expectancy* (prestatieverwachting) en *effort expectancy* (inspanningsverwachting). Deze concepten komen overeen met de eerdergenoemde concepten waargenomen nut en waargenomen gebruiksgemak uit TAM. In dit onderzoek wordt het UTAUT-model als uitgangspunt genomen omdat het zich specifiek richt op technologische acceptatie door werknemers en vanwege zijn sterke theoretische en empirisch onderbouwing. UTAUT2 (een latere variant op de originele UTAUT) wordt niet gebruikt, omdat dit model de acceptatie en het gebruik van technologie door consumenten probeert te verklaren (Venkatesh et al., 2012).

2.1.1 UTAUT

UTAUT (weergegeven in Figuur 1) onderscheidt drie variabelen die van directe invloed zijn op gebruikersintentie, namelijk: prestatieverwachting, inspanningsverwachting en sociale invloed (Venkatesh et al., 2003). Daarnaast onderscheid het model twee variabelen die van directe invloed zijn op gebruikersgedrag, te weten: gebruikersintentie en faciliterende condities. Het UTAUT-model bevat ook vier moderatoren die bepalend zijn voor de uitkomst van het model: gender, leeftijd, ervaring en vrijwilligheid van gebruik. Gender heeft invloed op zowel prestatieverwachting, inspanningsverwachting als sociale invloed. Leeftijd heeft invloed op alle vier de voorspellende factoren. Ervaring heeft invloed op inspanningsverwachting, sociale invloed en faciliterende condities. Tot slot heeft vrijwilligheid van gebruik enkel invloed op sociale invloed. Van de vier voorspellende factoren, is prestatieverwachting volgens Venkatesh et al. (2003) de sterkste voorspellende factor voor gebruikersintentie binnen UTAUT. Dit wordt ook onderbouwd door het literatuuronderzoek van Williams et al. (2015). Zoals eerdergenoemd hangt de sterkte van deze relatie af van twee moderatoren, namelijk leeftijd en gender. Op die manier dat de positieve relatie tussen prestatieverwachting en gebruikersintentie het sterkst is voor jonge mannen (Venkatesh et al., 2003).



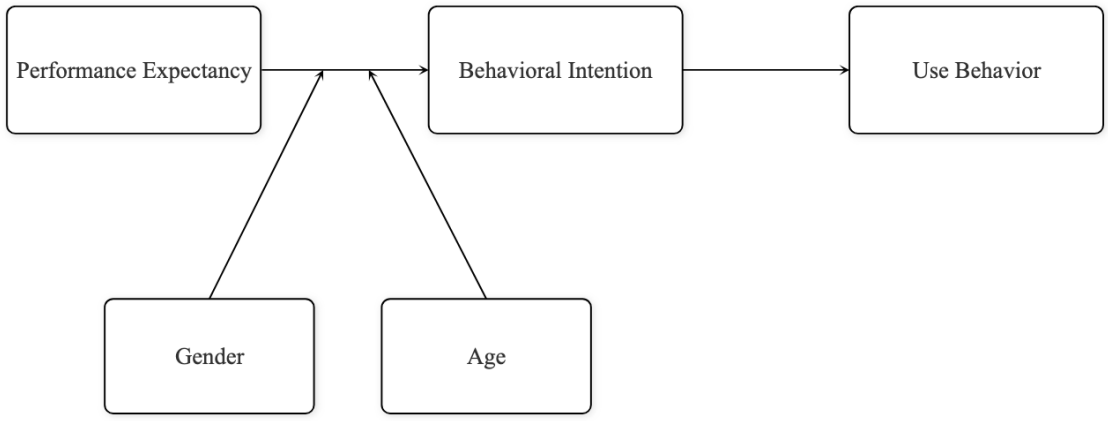
Figuur 1: Schematische weergave van UTAUT

2.1.2 UTAUT in dit onderzoek

Dit onderzoek richt zich op de acceptatie en het gebruik van een specifieke technologie, namelijk AI-systemen. Hoewel AI nog een relatief nieuwe technologie is, wordt er vanwege de snelle opkomst van AI en de potentie die AI zou hebben, steeds meer onderzoek verricht naar de acceptatie en het gebruik van AI-systemen door werknemers. Net als in dit onderzoek worden hiervoor technologische acceptatiemodellen zoals TAM en UTAUT als uitgangspunt gebruikt. Vanwege de beperkt mogelijke scope van dit onderzoek en in lijn met het hoofddoel van dit onderzoek, wordt in dit onderzoek niet de hele UTAUT getoetst, maar wordt er enkel gefocust op het concept prestatieverwachting. Hierdoor zal UTAUT zoals gebruikt in deze studie een deel van zijn eerdergenoemde verklarende variantie verliezen. Toch wordt in dit onderzoek voor UTAUT gekozen, omdat UTAUT in tegenstelling tot TAM, moderatoren bevat zoals gender en leeftijd, die van invloed zijn op prestatieverwachting. In Figuur 2 is de aangepaste UTAUT schematisch weergegeven.

In lijn met de verwachtingen van het UTAUT-model met betrekking tot technologie in het algemeen, brengt bestaand onderzoek het concept prestatieverwachting vrij consequent in verband met gebruikersintentie en/of daadwerkelijk gebruikersgedrag in de context van AI-

systemen. Onderzoek onder ondernemers laat zien dat prestatieverwachting van positieve invloed is op gebruikersintentie (Upadhyay et al., 2022). Dezelfde resultaten gelden voor de acceptatie van AI-systemen onder studenten (Gado et al., 2022). Ook hier is prestatieverwachting van positieve invloed op gebruikersintentie. In de conservatieve industrie (i.e. bouw-, olie- en gasindustrie) wordt prestatieverwachting ook positief in verband gebracht met gebruikersintentie en wordt gebruikersintentie op zijn beurt gerelateerd aan daadwerkelijk gebruikersgedrag (Khan et al., 2023). Vergelijkbare resultaten zijn gevonden in een onderzoek naar adoptie van AI-systemen onder zakenlui (Baabdullah, 2024). In dit onderzoek wordt onderscheid gemaakt tussen gebruikers en niet-gebruikers. Onder zowel gebruikers als niet-gebruikers is prestatieverwachting van positieve invloed op gebruikersintentie. Er is slechts één onderzoek uitgevoerd naar de acceptatie van AI-systemen in de publieke sector (Assaf et al., 2024). In dit onderzoek blijkt prestatieverwachting ook positief gerelateerd te zijn aan gebruikersintentie.



Figuur 2: Aangepaste schematische weergave van UTAUT

2.2 AI-geletterdheid

Gilster (1997) introduceerde in 1997 de term digitale geletterdheid, wat hij definieerde als “the ability to understand and use information in multiple formats from a wide range of sources when it is presented via computers”. Volgens Gilster (1997) is digitale geletterdheid een onmisbare vaardigheid in het digitale tijdperk. Op dezelfde wijze dat digitale geletterdheid als essentieel wordt gezien in het digitale tijdperk, zou AI-geletterdheid als essentieel kunnen worden beschouwd in het AI-tijdperk. Met de huidige toename in het gebruik van AI-systemen, is

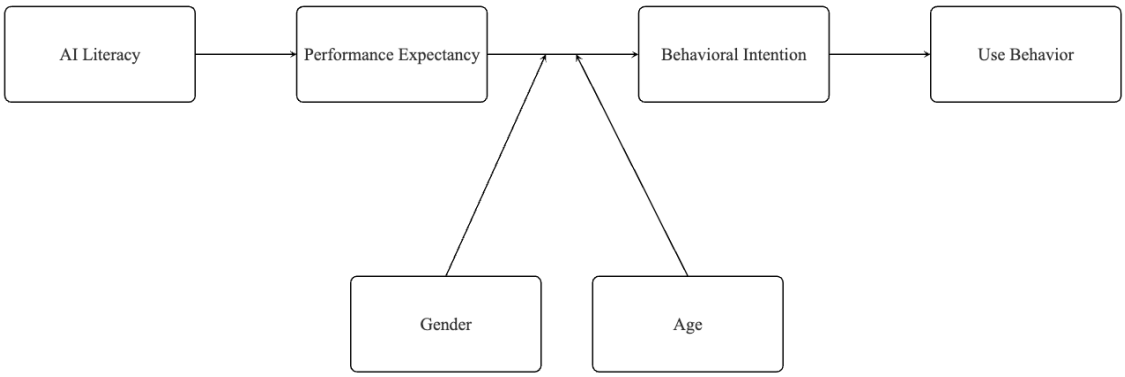
er steeds meer aandacht voor AI-geletterdheid. Want hoewel AI-systemen steeds vaker worden gebruikt, wil dit niet per definitie zeggen dat er ook bij iedereen sprake is van een voldoende kennisniveau om AI-systemen op een effectieve en verantwoorde manier te gebruiken.

Er zijn verschillende definities van AI-geletterdheid ontwikkeld, waarvan de meeste definities gebaseerd zijn op verschillende soorten geletterdheden (Ng et al., 2021). Een van de meest gebruikte definities van AI-geletterdheid is de definitie van Long & Magerko (2020), die AI-geletterdheid definiëren als "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace." Vergelijkbaar is de conceptualisatie van Ng et al. (2021), die stellen dat AI-geletterdheid bestaat uit vier aspecten: het kennen en begrijpen van AI, het toepassen van AI, het evalueren en creëren van AI en het hebben van ethische overwegingen met betrekking tot AI. Hoewel het concept AI-geletterdheid relatief nieuw is, zijn er in navolging van de definities van Long & Magerko (2020) en Ng et al. (2021) al een aantal schalen ontwikkeld om AI-geletterdheid te meten (Wang et al., 2023; Laupichler et al., 2023; Ng et al., 2024). Wat deze definities en schalen met elkaar gemeen hebben, is dat AI-geletterdheid niet enkel gaat over het hebben van technische kennis van AI-systemen, maar ook over het hebben van ethische overwegingen met betrekking tot het gebruik van AI-systemen.

Hoewel onderzoek naar de relatie tussen AI-geletterdheid en de acceptatie en het gebruik van AI-systemen op dit moment nog schaars is, wordt AI-geletterdheid in enkele studies al wel in verband gebracht met de acceptatie van AI-systemen. Uit onderzoek van Ma & Lei (2024) onder studenten blijkt AI-geletterdheid van positieve invloed te zijn op prestatieverwachting en prestatieverwachting op zijn beurt van positieve invloed te zijn op gebruikersintentie. Vergelijkbaar onderzoek van Schiavo et al. (2024) onder de algehele populatie laat dezelfde resultaten zien, maar laat daarnaast ook een direct effect van AI-geletterdheid op gebruikersintentie zien. Het onderzoek van Al-Abdullatif (2024) onder docenten laat een vergelijkbaar direct effect zien namelijk, tussen AI-geletterdheid en AI-acceptatie. Zij vinden daarentegen geen bewijs voor de invloed van AI-geletterdheid op prestatieverwachting. Tot op heden is er nog geen onderzoek gedaan naar de invloed van AI-geletterdheid op het concept prestatieverwachting uit het UTAUT-model binnen de publieke sector.

2.3 Conceptueel model

Op basis van bovenstaand theoretisch kader, ziet het conceptueel model er als volgt uit (zie Figuur 3).



Figuur 3: Conceptueel model

2.4 Hypothesen

Op basis van het conceptueel model kunnen onderstaande hypothesen worden opgesteld:

- Hypothese 1: AI-geletterdheid is van positieve invloed op prestatieverwachting
- Hypothese 2: Prestatieverwachting is van positieve invloed op gebruikersintentie
- Hypothese 3: AI-geletterdheid is van positieve invloed op gebruikersintentie
- Hypothese 4: Prestatieverwachting medieert de relatie tussen AI-geletterdheid en gebruikersintentie
- Hypothese 5: Bij mannen is de relatie tussen prestatieverwachting en gebruikersintentie sterker
- Hypothese 6: Bij jongere individuen is de relatie tussen prestatieverwachting en gebruikersintentie sterker
- Hypothese 7: Gebruikersintentie is van positieve invloed op gebruikersgedrag

3 Methode

In dit hoofdstuk wordt ingegaan op de methodologische opzet van dit onderzoek. Het hoofdstuk begint met een algemene toelichting op het gekozen onderzoeksontwerp. Daarna volgt een berekening van de minimale steekproefgrootte, een beschrijving van de wijze van steekproefselectie, de participanten, de procedure, de metingen en de analyses.

3.1 Onderzoeksontwerp

In dit onderzoek is gebruikgemaakt van een kwantitatieve onderzoeksbenadering. Een kwantitatieve onderzoeksbenadering draait om het toetsen van hypothesen op basis van bestaande theorieën (Creswell & Creswell, 2017). Dit komt overeen met de primaire doelstelling van dit onderzoek, namelijk het toetsen van hypothesen over de invloed van AI-geletterdheid op het construct prestatieverwachting binnen het UTAUT-model om inzicht te krijgen in de determinanten van acceptatie en gebruik van AI-systemen. Binnen de kwantitatieve onderzoeksbenadering kunnen verschillende onderzoeksontwerpen worden onderscheiden. In dit onderzoek is gekozen voor survey-onderzoek als onderzoeksontwerp. Het doen van survey-onderzoek biedt meerdere voordelen. Allereerst, is survey-onderzoek een kosten-effectieve manier om data te verzamelen (Creswell & Creswell, 2017). Ten tweede, is survey-onderzoek een manier om data te verzamelen over een relatief grote steekproef, waardoor de kans groter is dat de onderzoeksresultaten ook extern valide zijn. Ten derde, biedt survey-onderzoek anonimiteit waardoor individuen eerder geneigd zijn eerlijk te antwoorden (Ong & Weiss, 2000). Dit is extra belangrijk bij onderzoek naar gevoelige onderwerpen. Mogelijk is het gebruik van AI-systemen voor ambtenaren een gevoelig onderwerp vanwege de terughoudende overheidsbrede visie op het gebruik van generatieve AI-systemen (Rijksoverheid, 2024). Tot slot, is er gekozen voor een cross-sectionele methode waarbij alle data op één moment in de tijd wordt verzameld (Creswell & Creswell, 2017). Dit sluit goed aan bij dit onderzoek, omdat dit onderzoek zich richt op het doen van een momentopname. Ook zou het gegeven de beperkte scope van dit onderzoek praktisch niet haalbaar zijn om gebruik te maken van een longitudinale methode.

3.1.1 Steekproefselectie

De vereiste steekproefgrootte voor dit onderzoek is berekend door middel van een a priori poweranalyse. Deze poweranalyse is gedaan met behulp van versie 3.1 van G*Power (Faul et

al., 2009). De in- en output van de poweranalyse is te vinden in bijlage B. Aan de input valt af te lezen dat de berekening is gebaseerd op een gemiddelde effectgrootte ($f^2 = .5$), een alfa van .05, een gewenste power van .80 en zes predictoren. Er is gekozen om uit te gaan van een gemiddelde effectgrootte, omdat dit gebruikelijk is bij een nieuw onderzoek met relatief weinig bestaande literatuur (Cohen, 2013). Uit de poweranalyse blijkt dat een steekproefgrootte van 92 participanten vereist is voor het detecteren van significante effecten.

Voor het selecteren van individuen voor de steekproef is gebruikgemaakt van *convenience sampling*. Convenience sampling is een vorm van *non-probability sampling* waarbij individuen worden geselecteerd op basis van gemak en beschikbaarheid (Creswell & Creswell, 2017). Het nadeel van convenience sampling is dat niet ieder individu uit de onderzoekspopulatie een even grote kans heeft om geselecteerd te worden, wat ten koste gaat van de representativiteit van de steekproef. Convenience sampling heeft daarentegen als voordeel dat er in korte tijd data verzameld kan worden over een relatief grote groep individuen.

3.1.2 Participanten

Er zijn op verschillende manieren participanten gerekruteerd voor dit onderzoek. Allereerst hebben alle werknemers van de Provincie Zuid-Holland een mail ontvangen met de vraag om deel te nemen aan het onderzoek. Daarnaast is er een oproep tot deelname op het intranet van de Provincie Zuid-Holland geplaatst. Ook zijn alle deelnemers van het Interprovinciaal Overleg (IPO) gevraagd om de survey rond te sturen binnen de eigen provinciale organisatie. Er werd een cadeaukaart van 25 euro verloot om deelname aan het onderzoek te stimuleren. Aan de survey hebben uiteindelijk 315 participanten deelgenomen. Hiervan hebben 33 participanten de survey niet volledig ingevuld. Deze participanten zijn uit de dataset verwijderd. De mediaan van de duur van het invullen van de survey was 229 seconden. 2 participanten hebben de survey in minder dan 100 seconden ingevuld en zijn om die reden ook uit de dataset verwijderd. Slechts 2 participanten hebben als geslacht ‘anders’ ingevuld. Er is besloten om deze participanten uit de dataset te verwijderen, omdat het de resultaten van de analyses kan vertekenen en de groep zelf te klein is om betrouwbare uitspraken over te doen. Uiteindelijk bleef er een dataset met 278 participanten over. De meerderheid van de participanten was werkzaam bij de Provincie Zuid-Holland (66.2%). De andere participanten waren werkzaam bij de Provincie Groningen (28.8%), Provincie Zeeland (3.2%), Provincie Limburg (.7%) en Provincie Friesland (.4%). Nog eens 2 participanten (.7%) waren werkzaam bij een niet nader gespecificeerde

organisatie binnen de publieke sector. Aan het onderzoek hebben 146 mannen (52.5%) en 132 vrouwen (47.5%) meegedaan. De leeftijd varieerde van 21 tot en met 67 jaar (*M*= 43.47, *SD*= 12.20).

Een groot deel van de participanten gaf aan helemaal geen gebruik te maken van AI-systemen in hun werk (28.4%). De andere participanten minder dan één keer per week (25.4%), ongeveer één keer per week (10.4%), meerdere keren per week (18.7%), ongeveer één keer per dag (5.8%) en meerdere keren per dag (11.5%). Het merendeel (71.6%) maakt dus in meer of mindere mate regelmatig gebruik van AI-systemen in hun werk. 11 participanten die eerder aangaven helemaal geen gebruik te maken van AI-systemen in hun werk, gaven wel aan gebruik te maken van de voorgestelde antwoorden in Outlook (een vorm van AI). Van de participanten die aangaven gebruik te maken van AI-systemen in hun werk, zegt 59.7% gebruik te maken van een openbare chatbot, 22.3% van een interne chatbot, 14.7% van een hulpmiddel om spraak automatisch om te zetten in tekst, 14.0% van een schrijfhulpmiddel, 11.2% van een afbeeldingsgenerator en 9.7% van een programmeerassistent. Een belangrijke kanttekening hierbij is dat op het moment van dit onderzoek, alleen medewerkers van de Provincie Zuid-Holland toegang hadden tot een interne chatbot.

3.1.3 Procedure

Voor dit onderzoek is een online survey in Qualtrics gemaakt. Voorafgaand aan het versturen van de definitieve survey is er eerst een pilottest gedaan om te evalueren of de survey geschikt was voor de beoogde doelgroep. Er is op basis van de pilottest één aanpassing aan de survey gedaan. De controlevraag over het gebruik van AI-systemen (zie de laatste vraag van bijlage E) is aangepast zodat het duidelijker is dat hier gedoeld wordt op de antwoordsuggesties die Outlook doet bij het beantwoorden van een e-mail en niet om het verzenden van een automatische antwoord (out of office) in Outlook.

Voorafgaand aan het invullen van de survey kregen de participanten eerst algemene informatie over het doel van het onderzoek, de duur van het onderzoek en wat voor soort vragen er gesteld zouden worden. Daarnaast zijn de participanten geïnformeerd over waarvoor de data gebruikt zal gaan worden, dat er geen risico's aan het onderzoek verbonden zijn, dat deelname aan het onderzoek vrijwillig is en op elk moment gestopt kan worden, dat er geen vertrouwelijke informatie of persoonlijke gegevens worden gebruikt in de uitkomsten van het onderzoek en dat de data voor een periode van één jaar bewaard zal blijven. Er waren twee vereisten

voor deelname aan het onderzoek. Participanten moesten 18 jaar of ouder zijn en werkzaam zijn binnen het openbaar bestuur of de overheidsdiensten. De volledige inleidende tekst van de survey is te vinden in bijlage C. Na het geven van schriftelijke toestemming tot deelname aan het onderzoek, werden de participanten gevraagd enkele demografische informatie over zichzelf te verstrekken (zie bijlage D). Vervolgens kregen de participanten vragen over hun gebruik van AI-systemen in hun werk. Alleen wanneer participanten aangaven wel gebruik te maken van AI-systemen in hun werk, kregen zij vervolgvragen over het soort AI-systemen waar zij gebruik van maken (zie bijlage E). De participanten die zeiden helemaal geen gebruik te maken van AI-systemen, kregen enkel de laatste vraag hiervan te zien. Deze vraag was bedoeld als controlevraag om te achterhalen of individuen mogelijk toch onbewust gebruikmaken van AI-systemen in hun werk. Tot slot, werden aan alle participanten stellingen voorgelegd met betrekking tot AI-geletterdheid, prestatieverwachting en gebruikersintentie.

3.1.4 Metingen

Er is in dit onderzoek gefocust op de relatie tussen AI-geletterdheid en gebruikersintentie via prestatieverwachting, en de modererende rol van gender en leeftijd in de relatie tussen prestatieverwachting en gebruikersintentie. Ook stond de relatie tussen gebruikersintentie en gebruikersgedrag centraal.

Allereerst, is het concept AI-geletterdheid gemeten door middel van de *AI literacy scale* (AILS) van (Wang et al., 2023). De AILS meet zelfgerapporteerde AI-geletterdheid aan de hand van vier constructen: bewustzijn, gebruik, evaluatie en ethiek. De definities van deze constructen zoals gedefinieerd door Wang et al. (2023) zijn af te lezen in tabel 1. Er is voor deze schaal gekozen omdat deze schaal is ontwikkeld voor het meten van AI-geletterdheid onder de algehele populatie. Veel andere schalen zijn specifiek ontwikkeld voor het meten van AI-geletterdheid bij leerlingen, (medisch) studenten of docenten. Daarnaast is voor de kwaliteit van de AILS het meest robuuste bewijs gevonden (Lintner, 2024). De schaal is gevalideerd door andere onderzoekers en laat sterk positief bewijs zien voor structurele validiteit, interne consistentie en construct validiteit. Daarnaast is er zwak positief bewijs voor inhoudsvaliditeit en betrouwbaarheid. In tegenstelling tot andere schalen, waarvoor geen enkel bewijs voor inhoudsvaliditeit en betrouwbaarheid is gevonden. De AILS bevat 12 items met drie items per construct (Wang et al., 2023). Een voorbeeld van een item bedoeld om het construct ethiek te meten is: “ik ben altijd alert op misbruik van AI-technologie”. De AILS maakt gebruik van

een zevenpunts Likertschaal van ‘sterk mee oneens’ tot ‘sterk mee eens’. In tabel 2 zijn alle items uit de AILS opgenomen. De items zijn voor dit onderzoek vertaald van het Engels naar het Nederlands. De originele AILS bevat drie items (2, 5 en 11) die ten opzichte van de andere items in omgekeerde volgorde zijn geformuleerd. In dit onderzoek is dit aangepast zodat alle items in dezelfde richting zijn geformuleerd. Omgekeerd geformuleerde items zijn bedoeld om response bias te verminderen, maar leiden juist vaak tot vertekening van resultaten door onoplettendheid of verwarring (Sonderen et al., 2013).

Tabel 1 Definities van de constructen van de AILS

Construct	**Definitie**	**Bron**
Bewustzijn	“Het vermogen om AI-technologie te identificeren en te begrijpen tijdens het gebruik van AI-gerelateerde toepassingen.”	Wang et al. (2023)
Gebruik	“Het vermogen om AI-technologie toe te passen en te benutten om taken vakkundig uit te voeren.”	Wang et al. (2023)
Evaluatie	“Het vermogen om AI-toepassingen en de resultaten daarvan te analyseren, selecteren en kritisch te evalueren.”	Wang et al. (2023)
Ethiek	“Het vermogen om zich bewust te zijn van de verantwoordelijkheden en risico’s die gepaard gaan met het gebruik van AI-technologie.”	Wang et al. (2023)

De concepten prestatieverwachting, gebruikersintentie en gebruikersgedrag komen alle drie uit UTAUT (Venkatesh et al., 2003). Voor de concepten prestatieverwachting en gebruikersintentie geldt dat de gevalideerde schalen van Venkatesh et al. (2003) zijn gebruikt, die zijn opgenomen in tabel 2. De items zijn vanuit het Engels naar het Nederlands vertaald en aangepast naar de context van deze studie, namelijk naar de acceptatie en het gebruik van AI-systemen. Het concept prestatieverwachting is gemeten door middel van vier items. Een voorbeeld van een item is: “het gebruik van AI-systemen stelt mij in staat taken sneller uit te voeren”. In dit onderzoek is het eerste item van de originele schaal van Venkatesh et al. (2003) iets aangepast zodat het item beter aansluit bij zowel individuen die reeds gebruikmaken van AI-systemen in hun werk als individuen die (nog) geen gebruikmaken van AI-systemen in hun werk.

Het concept gebruikersintentie is gemeten door middel van drie items. Een voorbeeld van een item is: “ik ben van plan om AI-systemen te gebruiken in de komende 3 maanden”. De participanten die eerder hebben aangegeven reeds gebruik te maken van AI-systemen in hun

werk, krijgen een subtiel aangepaste versie hiervan te zien. In plaats van aan te geven in hoeverre zij van plan zijn AI-systemen te gebruiken in de komende 3 maanden, geven zij aan in hoeverre zij van plan zijn AI-systemen te *blijven* gebruiken in de komende 3 maanden. Dit geldt ook voor de andere twee items van de variabele gebruikersintentie. Zowel het concept prestatieverwachting als gebruikersintentie zijn in dit onderzoek gemeten door middel van een zevenpunts Likertschaal van 'sterk mee oneens' tot 'sterk mee eens'. Voor het concept prestatieverwachting komt dit overeen met het onderzoek van Venkatesh et al. (2003). Echter, wordt gebruikersintentie door Venkatesh et al. (2003) gemeten door middel van een driepuntsschaal. In dit onderzoek is er ten behoeve van de consistentie voor gekozen om gebruikersintentie net als AI-geletterdheid en prestatieverwachting te meten door middel van een zevenpunts Likertschaal.

Tot slot, wordt het concept gebruikersgedrag door Venkatesh et al. (2003) gemeten door middel van de daadwerkelijke gebruiksduur op basis van systeemlogboeken. Omdat er in dit onderzoek enkel gebruikt wordt gemaakt van een online survey voor het verzamelen van data, is er voor gekozen om gebruikersgedrag te meten op dezelfde manier als in het onderzoek van Davis (1989), waar participanten werden gevraagd op een zespuntsschaal aan te geven in welke mate zij op dit moment gebruikmaken van een bepaalde technologie. De antwoordopties waren: "Helemaal niet", "Minder dan één keer per week", "Ongeveer één keer per week", "Meerdere keren per week", "Ongeveer één keer per dag", "Meerdere keren per dag". Ook dit item is aangepast naar de context van dit onderzoek, namelijk naar het gebruik van AI-systemen (zie tabel 2).

Tabel 2 Items per construct

Construct	Items	Gebaseerd op
AI-geletterdheid	1. Ik kan slimme apparaten onderscheiden van niet slimme apparaten (bewustzijn) 2. Ik weet hoe AI-technologie mij kan helpen (bewustzijn) 3. Ik kan de AI-technologie identificeren die wordt gebruikt in de applicaties en producten die ik gebruik (bewustzijn) 4. Ik kan AI-applicaties of producten vaardig gebruiken om mij te helpen met mijn dagelijkse werk (gebruik) 5. Het is meestal makkelijk voor mij om een nieuwe AI-applicatie of product te leren gebruiken (gebruik) 6. Ik kan AI-applicaties of producten gebruiken om mijn werk efficiëntie te verbeteren (gebruik) 7. Ik kan de mogelijkheden en beperkingen van een AI-applicatie of product evalueren nadat ik het een tijdje heb gebruikt (evaluatie) 8. Ik kan een geschikte oplossing kiezen uit verschillende oplossingen die worden aangeboden door een slimme agent (evaluatie) 9. Ik kan de meest geschikte AI-applicatie of product kiezen uit een verscheidenheid voor een specifieke taak (evaluatie) 10. Ik houd altijd de ethische principes in acht bij het gebruik van AI-applicaties of producten (ethiek) 11. Ik ben altijd alert op privacy- en informatiebeveiligingsproblemen bij het gebruik van AI-applicaties of producten (ethiek) 12. Ik ben altijd alert op misbruik van AI-technologie (ethiek)	Wang et al. (2023)
Prestatieverwachting	1. Het gebruik van AI-systemen is nuttig in mijn werk 2. Het gebruik van AI-systemen stelt mij in staat om taken sneller uit te voeren 3. Het gebruik van AI-systemen verhoogt mijn productiviteit 4. Als ik AI-systemen gebruik, zal ik mijn kans op een salarisverhoging vergroten	Venkatesh et al. (2003)
Gebruikersintentie	1. Ik heb de intentie om AI-systemen te (blijven) gebruiken in de komende 3 maanden 2. Ik voorspel dat ik AI-systemen zal (blijven) gebruiken in de komende 3 maanden 3. Ik ben van plan AI-systemen te (blijven) gebruiken in de komende 3 maanden	Venkatesh et al. (2003)
Gebruikersgedrag	1. Hoe vaak maakt u momenteel gebruik van AI-systemen in uw werk?	Davis (1989)

3.1.5 Analyses

Voorafgaand aan het uitvoeren van de analyses zijn de scores op de items van de variabelen AI-geletterdheid, prestatieverwachting en gebruikersintentie per participant tot een gemiddelde geaggregeerd. Daarnaast is de variabele gender gehercodeerd in een dummy variabele zodat deze variabele geschikt was voor het doen van een regressieanalyse. Hierbij kregen mannen de waarde 0 en vrouwen de waarde 1.

Voor de beschrijving van de steekproef, de analyse van de betrouwbaarheid en validiteit, de beschrijvende statistieken, het controleren van de regressie assumpties en het toetsen van de hypothese over de invloed van gebruikersintentie op gebruikersgedrag is gebruikgemaakt van versie 30 van *Statistical Package for Social Sciences (SPSS)*. De invloed van gebruikersintentie op gebruikersgedrag is onderzocht door middel van een lineaire regressieanalyse.

Voor het toetsen van de hypothesen van de rest van het conceptuele model is gebruikgemaakt van versie 4.3 van PROCESS, een macro voor SPSS waarmee mediatie- en moderatieanalyses kunnen worden uitgevoerd (Hayes, 2012). Er is gekozen om gebruik te maken van model 16 van Hayes (2012), omdat dit model bijna het gehele conceptuele model in een keer kan toetsen. Alleen voor de relatie tussen gebruikersintentie en gebruikersgedrag was een aanvullende lineaire regressieanalyse in SPSS nodig.

4 Resultaten

In dit hoofdstuk worden de resultaten van dit onderzoek gerapporteerd. Het hoofdstuk begint met een betrouwbaarheidsanalyse. Daarna volgt een beoordeling van de validiteit door middel van een correlatie- en associatieanalyse en een factoranalyse. Vervolgens worden de beschrijvende statistieken getoond. Daarna volgt een controle van de data op de assumpties voor het doen van een regressieanalyse. Het hoofdstuk sluit af met het toetsen van de hypothesen, waarvan de resultaten worden samengevat in twee grafische weergaven.

4.1 Betrouwbaarheid

Voor het beoordelen van de betrouwbaarheid van de gebruikte schalen is Cronbach's alpha berekend. In tabel 3 valt af te lezen dat voor al de gebruikte schalen een zeer hoge betrouwbaarheid geldt. Van de schaal die gebruikt is om de variabele prestatieverwachting te meten is het laatste item ("als ik AI-systemen gebruik, zal ik mijn kans op een salarisverhoging vergroten") verwijderd om de betrouwbaarheid van de schaal te vergroten. Door het verwijderen van het item nam Cronbach's alpha toe van .856 naar .940. Voor alle andere schalen geldt dat de betrouwbaarheid afnam zodra een item verwijderd werd. Het is niet mogelijk om een betrouwbaarheidsanalyse uit te voeren voor de variabele gebruikersgedrag, omdat deze variabele door middel van slechts één item is gemeten.

Tabel 3 Cronbach's alpha per schaal

Schaal	Cronbach's alpha	*N* items
AI-geletterdheid (AIG)	.916	12
Prestatieverwachting (PV)	.940	3
Gebruikersintentie (GI)	.986	3

4.2 Validiteit

4.2.1 Correlatie- en associatieanalyse

Er is een correlatie- en associatieanalyse uitgevoerd om de onderlinge samenhang tussen de verschillende verklarende onderzoeksvariabelen inzichtelijk te maken en eventuele multicollineariteit te detecteren. Multicollineariteit treedt op wanneer er binnen een model sterke samenhang is tussen twee of meer verklarende variabelen (Draper & Smith, 1998). In dit onderzoek

worden twee modellen onafhankelijk van elkaar getoetst: 1) het hoofdmodel met alle variabelen behalve gebruikersgedrag en 2) het aanvullende model met alleen de variabelen gebruikersintentie en gebruikersgedrag. Aangezien het aanvullende model slechts één verklarende variabele bevat, kan hier geen sprake zijn van multicollineariteit.

Het hoofdmodel daarentegen bevat wel meerdere verklarende variabelen, namelijk: AI-geletterdheid, prestatieverwachting, leeftijd en gender. Voor het berekenen van de correlatie tussen AI-geletterdheid, prestatieverwachting en leeftijd is gebruikgemaakt van Spearmans rangcorrelatiecoëfficient. Er is gebruikgemaakt van Spearmans rangcorrelatiecoëfficient omdat deze correlatiemaat geschikt is voor variabelen die zijn gemeten op zowel ordinaal als ratio niveau. De resultaten zijn zichtbaar in tabel 4 en laten zien dat alle variabelen significant met elkaar samenhangen. Dit vormt echter geen multicollineariteitsprobleem, maar onderbouwt juist het gekozen model. Het gekozen model gaat er namelijk vanuit dat de verklarende variabelen met elkaar samenhangen door mediatie van prestatieverwachting en moderatie van leeftijd.

Omdat er geen maat bestaat die gelijktijdig de samenhang kan berekenen tussen zowel nominale, ordinale als ratio variabelen, is daarnaast Cramérs' V gebruikt om de samenhang te berekenen tussen gender en respectievelijk AI-geletterdheid en prestatieverwachting. Er is gekozen voor Cramér's V omdat deze associatiemaat geschikt is voor variabelen op zowel nominaal als ordinaal niveau. Uit de analyse blijkt een sterke associatie tussen geslacht en AI-geletterdheid (V= .59, p= <.01), en een matige associatie tussen geslacht en prestatieverwachting (V= .30, p= <.05). Ook hiervoor geldt dat dit geen multicollineariteitsprobleem vormt, omdat dit in lijn ligt met de verwachtingen van het model.

Tabel 4 Spearmans rangcorrelatiecoëfficient

	AI-geletterdheid	Prestatieverwachting	Leeftijd
AI-geletterdheid	1.00		
Prestatieverwachting	.57**	1.00	
Leeftijd	-.37**	-.16**	1.00

**. *Correlatie is significant op 0.01 niveau (2-tailed)*

4.2.2 Factoranalyse

In dit onderzoek wordt gebruikgemaakt van een relatief nieuwe schaal om het concept AI-geletterdheid te meten. Om die reden is er een *Principal Component Analysis* (PCA) uitgevoerd

om te onderzoeken of de door Wang et al. (2023) veronderstelde factorstructuur ook zichtbaar is in de data van dit onderzoek. Voorafgaand aan de PCA is gecontroleerd of de onderzoeksdata geschikt is voor het doen van een factoranalyse. Dit is gedaan door middel van de Kaiser-Meyer-Olkin (KMO) test en *Bartlett's Test of Sphericity*. De KMO-test laat een waarde van .92 zien en Bartlett's Test of Sphericity is significant (χ2= 1885, p= <.001). Hieruit blijkt dat de onderzoeksdata geschikt is voor een factoranalyse.

De resultaten van de factoranalyse zijn te vinden in bijlage F. De factorstructuur kan worden bepaald door te kijken naar twee criteria: het aantal componenten met een eigenwaarde groter dan 1 en het aantal eigenwaarden boven de zogeheten 'elleboog' in de screeplot (Cattell, 1966; Kaiser, 1960). Uit de resultaten blijkt dat er twee componenten zijn met een eigenwaarde groter dan 1 en dat deze componenten samen 64% van de totale variantie verklaren (zie tabel 10). Dit komt overeen met de screeplot waarin twee eigenwaarden boven de elleboog zichtbaar zijn (zie figuur 6). Op basis hiervan valt een tweedimensionale factorstructuur op te maken. Dit komt niet overeen met de vierdimensionale factorstructuur, bestaande uit bewustzijn, gebruik, evaluatie en ethiek, die Wang et al. (2023) veronderstelt. De geroteerde componentenmatrix (zie tabel 11) laat de ladingen van de verschillende items op de twee componenten zien. Hieruit valt af te lezen dat items 1 t/m 9 op component 1 laden en item 10 t/m 12 op component 2 laden. Alleen de items die bedoeld zijn om het construct 'ethiek' te meten lijken daarmee een onderscheidende dimensie te vormen.

4.3 Beschrijvende statistieken

In tabel 5 worden de onderzoeksvariabelen beschreven aan de hand van verschillende soorten spreidingsmaten. Voor de variabelen leeftijd en gender geldt dat alle relevante beschrijvende statistieken zijn opgenomen in sectie 3.1.2.

Tabel 5 Beschrijvende statistieken van de onderzoeksvariabelen

	AIG	PV	GI	GG
N	278	278	278	278
Gemiddelde	4.66	4.93	5.09	2.83
Mediaan	4.75	5.00	5.67	2.00
Modus	5	5	7	1
Std. Deviatie	1.17	1.48	1.84	1.68
Variantie	1.36	2.19	3.37	2.82
Minimum	1	1	1	1
Maximum	7	7	7	6

GG= Gebruikersgedrag

4.4 Regressieanalyse

4.4.1 Assumpties

Voorafgaand aan het uitvoeren van de regressieanalyses, is er gecontroleerd of er wordt voldaan aan de belangrijkste assumpties voor het doen van een regressieanalyse, namelijk: (1) de residuen moeten normaal verdeeld zijn, (2) de residuen moeten homoscedastisch zijn, (3) de onafhankelijke en afhankelijke variabelen moeten een lineair (of geen) verband vertonen en (4) er mag geen sprake zijn van multicollineariteit. In sectie 4.2.1 was al vastgesteld dat zich in beide modellen geen multicollineariteitsprobleem voordoet. De andere assumpties worden in deze sectie besproken.

Allereerst is voor zowel het hoofdmodel als het aanvullende model gecontroleerd of de residuen normaal verdeeld zijn. Uit de grafieken in bijlage G kan worden opgemaakt dat voor beide modellen geldt dat de residuen normaal verdeeld zijn. De histogrammen laten immers een normale verdeling zien en in de P-Plots liggen de punten min of meer op de diagonale lijn. Daarna is voor beide modellen gecontroleerd of de residuen homoscedastisch zijn. In de scatterplots van de residuen in bijlage H is te zien dat de residuen van beide modellen niet uniform verdeeld zijn. Dit betekent dat er sprake is van heteroscedasticiteit en er niet wordt voldaan aan de homoscedasticiteitsassumptie. Voor de niet-uniforme verdeling zal voor beide modellen worden gecorrigeerd door middel van bootstrapping. Tot slot, is er gecontroleerd of alle onafhankelijke variabelen (behalve gender) en afhankelijke variabelen een lineair (of geen) verband met elkaar vertonen. Hierbij zijn ook de interactietermen gecontroleerd. Op basis van de scatterplots in bijlage I valt te concluderen dat voor beide modellen aan de lineariteitsassumptie wordt voldaan. Alle onafhankelijke en afhankelijke variabelen laten een lineair verband met

elkaar zien. Alleen tussen leeftijd en gebruikersintentie is geen enkel verband te zien, maar hiermee wordt de lineariteitsassumptie zoals gezegd niet geschonden.

4.4.2 Hypothesetoetsing

Directe effecten

Voor het toetsen van hypothesen 1 t/m 6 is een multiple lineaire regressieanalyse uitgevoerd. Dit is gedaan door middel van PROCESS-model 16 van (Hayes, 2012). Om te corrigeren voor de aanwezige heteroscedasticiteit, is hierbij gebruikgemaakt van bootstrapping met 5000 resamples. De regressiemodellen zijn zichtbaar in tabel 6 en 7. Uit de resultaten blijkt dat AI-geletterdheid positief gerelateerd is aan prestatieverwachting (B= .632, p= < .001) en dat prestatieverwachting positief gerelateerd is aan gebruikersintentie (B= .824, p= < .001). Dit betekent dat zowel hypothese 1 “AI-geletterdheid is van positieve invloed op prestatieverwachting” als hypothese 2 “prestatieverwachting is van positieve invloed op gebruikersintentie” is bevestigd. Daarnaast blijkt er ook een directe relatie tussen AI-geletterdheid en gebruikersintentie (B= .338, p= < .001). Hiermee is hypothese 3 “AI-geletterdheid is van positieve invloed op gebruikersintentie” ook bevestigd. Deze resultaten worden ook onderbouwd door de bootstrapresultaten in tabel 12 en 13 in bijlage J. Uit deze tabellen valt immers af te lezen dat voor geen van de relaties geldt dat de waarde 0 in het bootstrap-betrouwbaarheidsinterval zit. Op basis hiervan kunnen de nulhypotheses definitief worden verworpen.

Tabel 6 Resultaten multiple regressieanalyse

	B	*SE*	*t*	*p*
Constante	-2.945	.318	-9.264	.000
AI-geletterdheid	.632	.066	9.549	.000

Afhankelijke variabele: Prestatieverwachting

Tabel 7 Resultaten multiple regressieanalyse

	B	*SE*	*t*	*p*
Constante	3.422	.370	9.259	.000
AI-geletterdheid	.338	.074	4.581	.000
Prestatieverwachting	.824	.068	12.066	.000
Gender	.161	.142	1.137	.257
Gender x Prestatieverwachting	-.006	.095	-.064	.949
Leeftijd	.004	.006	.597	.551
Leeftijd x Prestatieverwachting	-.007	.004	-1.982	.049

Afhankelijke variabele: Gebruikersintentie

Voor het toetsen van hypothese 7 is nog een aanvullende lineaire regressieanalyse uitgevoerd. Ook hierbij is vanwege de aanwezige heteroscedasticiteit gebruikgemaakt van bootstrapping met 5000 resamples. Het regressiemodel is zichtbaar in tabel 8. Uit de resultaten blijkt dat gebruikersintentie positief gerelateerd is aan gebruikersgedrag (B= .644, p= < .001). Dit betekent dat hypothese 7 "gebruikersintentie is van positieve invloed op gebruikersgedrag" kan worden bevestigd. Dit wordt ook onderbouwd door de bootstrapresultaten in tabel 14 in bijlage J. Ook hier kan de nulhypothese worden verworpen, omdat de waarde 0 niet in het bootstrap-betrouwbaarheidsinterval zit.

Tabel 8 Resultaten lineaire regressieanalyse

	B	*SE*	*t*	*p*
Constante	-.455	.212	-2.147	.033
Gebruikersintentie	.644	.039	16.466	<.001

Afhankelijke variabele: gebruikersgedrag

Mediatie effect

Hypothese 4 "prestatieverwachting medieert de relatie tussen AI-geletterdheid en gebruikersintentie" is getoetst door te kijken naar de bootstrap-betrouwbaarheidsintervallen van het indirecte effect van AI-geletterdheid op gebruikersintentie via prestatieverwachting. Tabel 9 laat de indirecte effecten zien voor drie verschillende waarden van de moderatoren. Uit de resultaten blijkt dat voor alle waarden een significant indirect effect geldt, omdat de waarde 0 in geen van de bootstrap-betrouwbaarheidsintervallen zit. Dit betekent dat ook hypothese 4 is bevestigd. Er is echter geen sprake van volledige mediatie, omdat eerder al was vastgesteld dat er ook een direct effect is tussen AI-geletterdheid en gebruikersintentie. De resultaten wijzen in plaats daarvan op gedeeltelijke mediatie.

Tabel 9 Indirect effect van AI-geletterdheid op gebruikersintentie via prestatieverwachting

Gender	Leeftijd	Effect	BootSE	BootLLCI	BootULCI
.000	-12.196	.578	.077	.433	.730
.000	.000	.521	.075	.379	.670
.000	12.196	.463	.087	.297	.639
1.000	-12.196	.574	.080	.420	.741
1.000	.000	.517	.071	.383	.666
1.000	12.196	.460	.077	.320	.621

Moderatie effecten

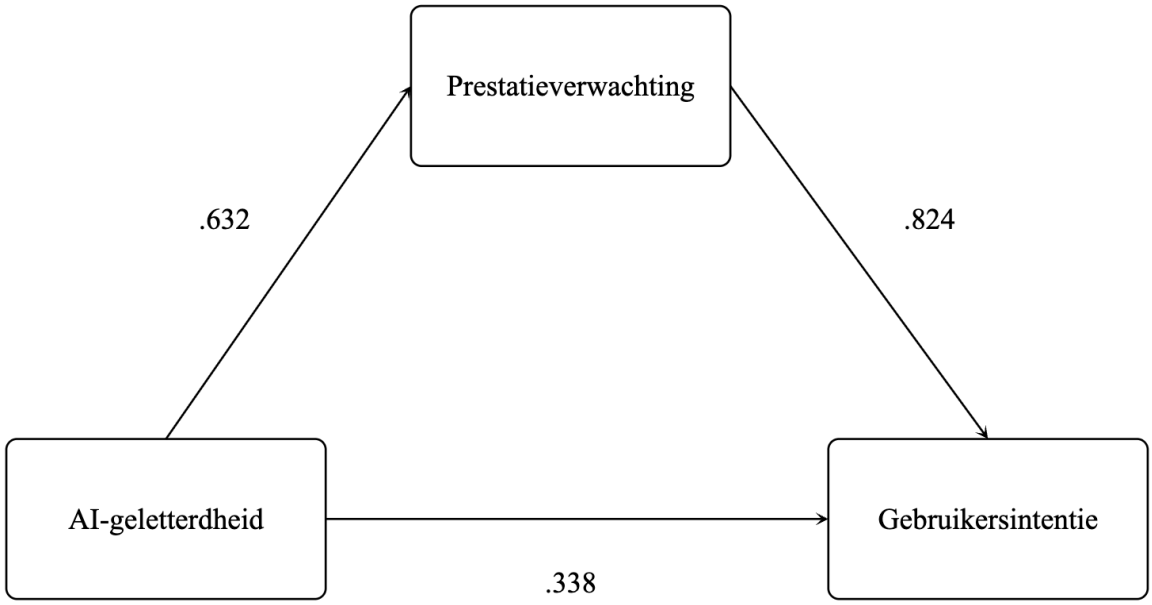
Zoals af te lezen is in het regressiemodel in tabel 7, is de interactie tussen gender en prestatieverwachting niet significant (*B*= -.006, *p*= .949). Dit betekent dat de relatie tussen prestatieverwachting en gebruikersintentie niet afhankelijk is van gender. Hypothese 5 “bij mannen is de relatie tussen prestatieverwachting en gebruikersintentie sterker” is dus niet bevestigd. Dit wordt ook bevestigd door de bootstrapresultaten in tabel 13 in bijlage J. Daarin valt namelijk af te lezen dat de waarde 0 in het bootstrap-betrouwbaarheidsinterval van de interactie zit.

De interactie tussen leeftijd en prestatieverwachting lijkt op basis van het regressiemodel wel nipt significant te zijn (*B*= -.007, *p*= < .05). Dit wordt echter niet bevestigd door de bootstrapresultaten, wederom in tabel 13 in bijlage J. Ook hiervoor geldt dat de waarde 0 in het bootstrap-betrouwbaarheidsinterval van de interactie zit, waardoor de nulhypothese toch moet worden aangenomen. Hypothese 6 “bij jongere individuen is de relatie tussen prestatieverwachting en gebruikersintentie sterker” is dus niet bevestigd. De bootstrapresultaten komen overeen met de index van gedeeltelijke gemodereerde mediatie die voor zowel gender (BootLLCI= -.125, BootULCI= .128) als leeftijd (BootLLCI= -.010, BootULCI= .001) niet significant is. Uit tabel 9 viel ook al op te maken dat het indirecte effect voor alle waarden van de moderatoren min of meer gelijk was. Dit bevestigd eens te meer dat het gevonden mediatie-effect niet wordt beïnvloed door de moderatoren.

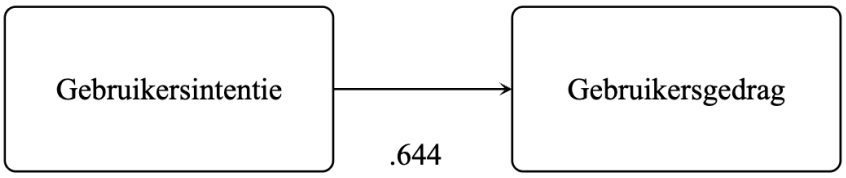
4.5 Samenvatting resultaten

De onderzoeksresultaten kunnen worden samengevat in de twee onderstaande modellen. In figuur 4 worden de resultaten van het hoofdmodel samengevat en figuur 5 geeft het resultaat weer van het aanvullende model. De pijlen geven de richting van elk verband aan en de coëfficiënten

de sterkte van elk verband. De moderatoren gender en leeftijd zijn logischerwijs niet opgenomen in het hoofdmodel, omdat er geen bewijs is gevonden voor gemodereerde mediatie. Om de verklarende variantie van het gevonden mediatie-model te berekenen is PROCESS-model 4 van Hayes (2012) gerund. Hieruit blijkt dat het mediatie-model 61.3% aan variantie in gebruikersintentie kan verklaren.



Figuur 4: Grafische weergave van de resultaten van het hoofdmodel



Figuur 5: Grafische weergaven van de resultaten van het aanvullende model

5 Conclusie en discussie

Dit hoofdstuk begint met het beantwoorden van de empirische deelvragen teneinde antwoord te kunnen geven op de hoofdvraag van dit onderzoek. Vervolgens worden de limitaties van dit onderzoek benoemd. Daarna volgen de theoretische implicaties en praktische aanbevelingen. Het hoofdstuk sluit af met suggesties voor toekomstig onderzoek.

5.1 Conclusie

De doelstelling van dit onderzoek was het toetsen van hypothesen over de invloed van AI-geletterdheid op het construct prestatieverwachting binnen het UTAUT-model, door middel van een kwantitatieve survey-analyse onder werknemers in de publieke sector, om inzicht te krijgen in de determinanten van de acceptatie en het gebruik van AI-systemen. Daarbij was de hoofdvraag: in hoeverre hebben AI-geletterdheid en prestatieverwachting invloed op de acceptatie en het gebruik van AI-systemen?

Uit dit onderzoek blijkt dat AI-geletterdheid van significante en positieve invloed is op prestatieverwachting. Dit betekent dat wanneer individuen zeggen bekwaam te zijn in het gebruik van AI-systemen op een verantwoorde manier, zij eerder denken dat AI-systemen nuttig zijn bij wat ze moeten doen. Dit komt overeen met het nog heel schaarse onderzoek dat tot op heden is gedaan naar de relatie tussen AI-geletterdheid en prestatieverwachting (Al-Abdullatif, 2024; Ma & Lei, 2024; Schiavo et al., 2024).

Daarnaast blijkt dat prestatieverwachting op zijn beurt, van significante en positieve invloed is op gebruikersintentie. Met andere woorden, wanneer individuen denken dat AI-systemen nuttig zijn bij wat ze moeten doen, zeggen zij eerder bereid te zijn om gebruik te maken van AI-systemen. Deze bevinding is in lijn met de verwachtingen van verschillende technologische acceptatiemodellen als het gaat om technologie in het algemeen (Davis, 1989; Venkatesh et al., 2003) en onderzoek naar de acceptatie van AI-systemen in het specifiek (Assaf et al., 2024; Baabdullah, 2024; Gado et al., 2022; Khan et al., 2023; Upadhyay et al., 2022).

Daarnaast blijkt er ook een direct effect te bestaan tussen AI-geletterdheid en gebruikersintentie. Anders gezegd, wanneer individuen zeggen bekwaam te zijn in het gebruik van AI-systemen op een verantwoorde manier, zijn zij eerder bereid gebruik te maken van AI-systemen. Deze bevinding is in overeenstemming met het onderzoek van Schiavo et al. (2024) waaruit ook een direct effect blijkt tussen AI-geletterdheid en gebruikersintentie en is te verge-

lijken met het onderzoek van Al-Abdullatif (2024) waarin een directe relatie wordt gevonden tussen AI-geletterdheid en AI-acceptatie.

Kortom, de bevindingen van dit onderzoek laten zien dat er sprake is van gedeeltelijke mediatie van prestatieverwachting in de relatie tussen AI-geletterdheid en gebruikersintentie. Met andere woorden, AI-geletterdheid is zowel van indirecte invloed op gebruikersintentie via prestatieverwachting als van directe invloed op gebruikersintentie zonder tussenkomst van prestatieverwachting. Dit betekent dat prestatieverwachting voor een deel verklaart waarom individuen die zeggen AI-geletterd te zijn, eerder bereid zijn om gebruik te maken van AI-systemen, maar dat prestatieverwachting niet het enige werkzame mechanisme is in deze relatie. Andere factoren die in dit onderzoek niet zijn onderzocht spelen ook een rol.

Verder blijkt uit dit onderzoek dat gebruikersintentie van significante en positieve invloed is op gebruikersgedrag. Dit betekent dat wanneer individuen zeggen bereid te zijn om AI-systemen te gebruiken, zij ook daadwerkelijk meer gebruikmaken van AI-systemen. Deze bevinding is niet verrassend gegeven het bestaande onderzoek dat gebruikersintentie, naast prestatieverwachting, identificeert als een van de sterkste voorspellers van het UTAUT-model (Williams et al., 2015). Ook specifiek met betrekking tot AI-systemen, biedt bestaand onderzoek onderbouwing voor de bevinding dat gebruikersintentie positief gerelateerd is aan daadwerkelijk gebruikersgedrag (Baabdullah, 2024; Khan et al., 2023).

Gender en leeftijd daarentegen, blijken geen significante moderatoren te zijn in de relatie tussen prestatieverwachting en gebruikersintentie. Deze bevinding is verrassend, aangezien het UTAUT-model voorspelt dat de relatie tussen prestatieverwachting en gebruikersintentie het sterkst is bij jongen mannen (Venkatesh et al., 2003). Volgens Venkatesh et al. (2003) zou dit komen doordat mannen taakgerichter zijn dan vrouwen en daardoor eerder bereid zijn gebruik te maken van een technologie waarvan zij denken dat dit hen kan helpen bij het verrichten van taken. Verder zouden jongere individuen extrinsieke beloning belangrijker vinden dan oudere individuen en daardoor eerder bereid zijn gebruik te maken van een technologie waarvan zij denken dat dit hen kan helpen om beter te presteren.

Mogelijk is er in dit onderzoek geen bewijs gevonden voor de modererende rol van gender en leeftijd vanwege de specifieke technologie die in dit onderzoek centraal stond. Op dit moment is er massale belangstelling voor AI vanwege de potentie die de technologie zou hebben in het vergaand veranderen van de manier waarop we werken. Daarbovenop is het gebruik van de technologie (ook binnen de publieke sector) al heel wijdverspreid (Eurostat, 2024; TNO, 2024).

De eigenschappen die zijn toe te schrijven aan gender en leeftijd en die volgens Venkatesh et al. (2003) zouden leiden tot een meer of mindere bereidheid tot het gebruik van technologie, spelen onder deze omstandigheden mogelijk een minder grote rol.

5.2 Discussie

5.2.1 Limitaties

Dit onderzoek heeft meerdere sterke punten. Allereerst, heeft dit onderzoek een relatief grote steekproef (n= 278) gegeven het hoofddoel van het onderzoek, namelijk het toetsen van het conceptuele model. De adequate steekproefgrootte draagt bij aan de kans op het vinden van een statistisch significant effect gegeven dat er ook sprake is van een daadwerkelijk effect (Cohen, 2013). Ten tweede, biedt dit onderzoek inzicht in AI-geletterdheid en de acceptatie en het gebruik van AI-systemen in een heel nieuwe context, namelijk de publieke sector. Dit is tot op heden niet tot nauwelijks gedaan. Ook draagt het onderzoek bij aan een verbeterd wetenschappelijk begrip van wat AI-geletterdheid is.

Echter, is geen enkel onderzoek zonder limitaties en dat geldt ook voor dit onderzoek. Allereerst, hoewel dit onderzoek een grote steekproef heeft waarmee kon worden voorzien in het bereiken van het hoofddoel van dit onderzoek, is er door het gebruik van *convenience sampling* nadrukkelijk geen sprake van een representatieve steekproef (Creswell & Creswell, 2017). Vanwege de niet-representatieve steekproef hebben de resultaten van dit onderzoek een beperkte generaliseerbaarheid. Met andere woorden, er kan op basis van de resultaten van dit onderzoek geen algemene conclusie worden getrokken over de publieke sector.

Ten tweede, is er in dit onderzoek mogelijk sprake van *social desirability bias*. Social desirability bias is de neiging van respondenten om sociaal wenselijk gedrag over te rapporteren (Krumpal, 2013). Zeker met betrekking tot de vragen over het gebruik van AI-systemen, kan er sprake zijn geweest van sociaal wenselijke beantwoording. Dit vanwege de op dat moment geldende overheidsbrede visie op generatieve AI, waarin staat vastgelegd dat het gebruik van niet-gecontracteerde generatieve AI-toepassingen (Zoals ChatGPT, Midjourney etc.) niet is toegestaan (Rijksoverheid, 2024). Hierdoor kan een discrepantie zijn ontstaan tussen wat individuen zeggen te doen en wat zij werkelijk doen. Hoewel geprobeerd is dit zoveel mogelijk te minimaliseren door het garanderen van anonimiteit (Ong & Weiss, 2000), is het denkbaar dat dit desondanks wel een rol heeft gespeeld.

Ten derde, kunnen er vraagtekens worden gezet bij de validiteit van de schaal die gebruikt

is om het concept AI-geletterdheid te meten. De in dit onderzoek uitgevoerde factoranalyse levert geen bewijs voor de door Wang et al. (2023) veronderstelde vierdimensionale structuur bestaande uit bewustzijn, gebruik, evaluatie en ethiek. In de factorstructuur van dit onderzoek is alleen de dimensie ethiek duidelijk zichtbaar. De drie andere dimensies zijn niet duidelijk van elkaar te onderscheiden. Dit betekent dat de AILS in de context van dit onderzoek niet precies lijkt te meten wat het bedoelt te meten.

Tot slot, is dit onderzoek beperkt door de keuze om voor de analyse hoofdzakelijk gebruik te maken van PROCESS. Door het gebruik van PROCESS was het niet mogelijk om het gehele conceptuele model in één keer te toetsen en was er een aanvullende lineaire regressieanalyse nodig. Hierdoor kunnen de resultaten van dit onderzoek niet volledig in samenhang worden geïnterpreteerd. Het hoofdmodel en het aanvullende model zoals zichtbaar in figuur 4 en 5 moeten afzonderlijk van elkaar worden beschouwd.

5.2.2 Theoretische implicaties

De bevindingen van dit onderzoek dragen op een aantal manieren bij aan de bestaande literatuur. Allereerst, valideert dit onderzoek het belang van prestatieverwachting en gebruikersintentie bij de acceptatie en het gebruik van technologie. Hoewel bestaand onderzoek prestatieverwachting en gebruikersintentie al langer ziet als belangrijke voorspellers binnen het UTAUT-model (Venkatesh et al., 2003; Williams et al., 2015), breidt dit onderzoek deze inzichten uit naar de acceptatie en het gebruik van een relatief nieuwe technologie (AI-systemen) in een tot op heden nog onderbelichte context (de publieke sector).

Daarnaast bouwt dit onderzoek voort op recent onderzoek dat AI-geletterdheid toevoegt als voorspellende factor aan bestaande technologische acceptatiemodellen als het gaat om de acceptatie van AI-systemen (Al-Abdullatif, 2024; Ma & Lei, 2024; Schiavo et al., 2024). De bevindingen van dit onderzoek onderbouwen het idee dat AI-geletterdheid bijdraagt aan technologische acceptatie, deels doordat mensen die AI-geletterd zijn een hogere verwachting hebben van het nut van de technologie bij wat ze moeten doen. De bevindingen laten verder zien dat dit effect onafhankelijk is van gender of leeftijd. Deze kenmerken werden eerder nog niet onderzocht als moderatoren in de relatie tussen AI-geletterdheid, prestatieverwachting en gebruikersintentie.

Tot slot, reageert dit onderzoek in algemene zin op de oproep van Long & Magerko (2020) en Schiavo et al. (2024) tot meer empirisch onderzoek naar AI-geletterdheid en in het specifiek

op de oproep van Lintner (2024) tot blijvend onderzoek naar de kwaliteit van de bestaande schalen die worden gebruikt om AI-geletterdheid te meten. Hoewel dit onderzoek sterk bewijs levert voor de betrouwbaarheid van de AILS, is er zoals eerder gezegd, geen bewijs gevonden voor de constructvaliditeit van de AILS. De uitgevoerde factoranalyse biedt geen onderbouwing voor de theoretische aanname dat AI-geletterdheid bestaat uit vier dimensies (bewustzijn, gebruik, evaluatie en ethiek), maar wijst juist op een tweedimensionale structuur met ethiek als één van de twee dimensies. De andere dimensie lijkt samengevat te kunnen worden tot praktische/technische kennis van AI-systemen. Dit beeld komt overeen met andere definities en schalen van AI-geletterdheid die, hoewel op aspecten verschillend, met elkaar gemeen hebben dat AI-geletterdheid gaat over zowel technische kennis van AI-systemen als het hebben van ethische overwegingen met betrekking tot het gebruik van AI-systemen.

5.2.3 Praktische aanbevelingen

Dit onderzoek laat zien dat het gebruik van AI-systemen binnen de provinciale organisatie heel wijdverspreid is: bijna driekwart van de respondenten geeft aan op regelmatige basis gebruik te maken van AI-systemen in hun werk. Dit gegeven onderstreept het belang van het bevorderen van de AI-geletterdheid onder werknemers binnen de provinciale organisatie, zoals ook op Europees niveau is vastgelegd in de AI-verordening, waarin onder andere staat dat alle organisaties die AI-systemen ontwikkelen of gebruiken, verplicht zijn om de AI-geletterdheid onder haar werknemers te bevorderen (*Verordening (EU) 2024/1689*, 2024).

Hoewel de AI-verordening organisaties verplicht om de AI-geletterdheid van haar werknemers te bevorderen, zegt de AI-verordening weinig over wat AI-geletterdheid precies is en hoe organisaties dit zouden moeten bevorderen (*Verordening (EU) 2024/1689*, 2024). Dit onderzoek laat op basis van theoretisch en empirisch onderzoek zien dat AI-geletterdheid in ieder geval tweeledig is: het omvat enerzijds een technisch/praktisch aspect en anderzijds een ethisch aspect. Gegeven hoe wijdverspreid het gebruik van AI-systemen al is en de snelheid waarmee dit gebruik toeneemt (TNO, 2024), is het raadzaam om een training ‘AI-geletterdheid’ te ontwikkelen voor alle werknemers, waarin aandacht is voor zowel het het herkennen, gebruiken en evalueren van AI (praktisch/technisch aspect) alsook aandacht is voor het vergroten van het bewustzijn over de verantwoordelijkheden en de risico’s die inherent zijn aan het gebruik van AI-systemen (ethisch aspect).

Tot slot, lijkt de premisse van de AI-verordening, dat het vergroten van de AI-geletterdheid

van individuen, hen in staat stelt AI-systemen geïnformeerd in te zetten (*Verordening (EU) 2024/1689*, 2024), te kloppen. Daarbij blijft het echter aanbevolen om er sterk op toe te zien dat AI-systemen ook daadwerkelijk op een verantwoorde manier worden ingezet. Immers, hoewel geïnformeerd gebruik van AI-systemen logischerwijs voorwaardelijk is voor verantwoorde inzet van AI-systemen, is dit nadrukkelijk niet hetzelfde.

5.2.4 Suggesties voor toekomstig onderzoek

Op basis van de tekortkomingen en theoretische implicaties van dit onderzoek, kunnen er een aantal suggesties voor toekomstig onderzoek worden gedaan. Allereerst, heeft survey-onderzoek net als ieder onderzoeksontwerp zijn beperkingen. Toekomstig onderzoek zou gelijktijdig gebruik kunnen maken van meerdere methoden (triangulatie) zodat voor de beperkingen van de andere methoden wordt gecompenseerd (Thurmond, 2001). Ten tweede, is er vervolgonderzoek nodig naar AI-geletterdheid en de acceptatie en het gebruik van AI-systemen in de publieke sector waarbij gebruik wordt gemaakt van een representatieve steekproef, zodat de onderzoeksbevindingen ook extern valide zijn en niet alleen binnen de kaders van deze steekproef. Ten derde, is er meer onderzoek nodig dat zich richt op de ontwikkeling van betrouwbare en valide schalen om het concept AI-geletterdheid te meten. Hoewel van alle beschikbare schalen om AI-geletterdheid te meten, voor de AILS het meeste robuuste bewijs is gevonden (Lintner, 2024), blijkt uit dit onderzoek een gebrek aan bewijs ten aanzien van de constructvaliditeit van de AILS. Ten vierde, zou toekomstig onderzoek gebruik moeten maken van meer geavanceerde statistische methoden zoals *Structural Equation Modeling* (SEM) waarmee het gehele conceptuele model in één keer getoetst kan worden. Door middel van SEM zou het ook mogelijk zijn om gelijktijdig andere factoren te onderzoeken die naast prestatieverwachting, een rol spelen in het verklaren van de relatie tussen AI-geletterdheid en gebruikersintentie. Ten vijfde, is er gezien het nog beperkte onderzoek naar AI-geletterdheid in relatie tot technologische acceptatie, replicatieonderzoek nodig om vast te kunnen stellen of het gevonden mediatie-effect zich inderdaad voordoet onafhankelijk van gender en leeftijd. Tot slot, zou een interessante onderzoeksrichting zijn of het bevorderen van AI-geletterdheid daadwerkelijk bijdraagt aan verantwoorde adoptie van AI-systemen en niet enkel aan de adoptie van AI-systemen op zichzelf.

Referenties

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J. & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of ai. *International Journal of Information Management*, *60*, 102387. https://doi.org/10.1016/j.ijinfomgt.2021.102387

Al-Abdullatif, A. M. (2024). Modeling teachers' acceptance of generative artificial intelligence use in higher education: The role of ai literacy, intelligent tpack, and perceived trust. *Education Sciences*, *14*(11), 1209. https://doi.org/10.3390/educsci14111209

Al Naqbi, H., Bahroun, Z. & Ahmed, V. (2024). Enhancing work productivity through generative artificial intelligence: A comprehensive literature review. *Sustainability*, *16*(3), 1166. https://doi.org/10.3390/su16031166

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., … others (2020). Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information fusion*, *58*, 82–115. https//-doi.org/10.1016/j.inffus.2019.12.01

Assaf, R., Omar, M., Saleh, Y., Attar, H., Alaqra, N. T. & Kanan, M. (2024). Assessing the acceptance for implementing artificial intelligence technologies in the governmental sector: An empirical study. *Engineering, Technology & Applied Science Research*, *14*(6), 18160–18170. https://doi.org/10.48084/etasr.8711

Baabdullah, A. M. (2024). The precursors of ai adoption in business: Towards an efficient decision-making and functional performance. *International Journal of Information Management*, *75*, 102745. https://doi.org/10.1016/j.ijinfomgt.2023.102745

Banh, L. & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, *33*(1), 63. https://doi.org/10.1007/s12525-023-00680-1

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102$_1$0

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Creswell, J. W. & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319–340. https://doi.org/10.2307/249008

Draper, N. R. & Smith, H. (1998). *Applied regression analysis*. John Wiley & Sons.

Europees Parlement. (2021). *Shaping the digital transformation: Eu strategy explained.* Geraadpleegd op 17 februari 2025, van https://www.europarl.europa.eu/topics/en/article/20210414STO02010/shaping-the-digital-transformation-eu-strategy-explained

Europese Commissie. (2024). *Ai act enters into force.* Geraadpleegd op 17 februari 2025, van https://commission.europa.eu/news/ai-act-enters-force-2024-08-01_en

Eurostat. (2024). *Use of artificial intelligence in enterprises.* Geraadpleegd op 17 februari 2025, van https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises

Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. (2009). Statistical power analyses using g* power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fishbein, M. & Ajzen, I. (1975). *Beliefs, attitude, intention, and behavior: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Gado, S., Kempen, R., Lingelbach, K. & Bipp, T. (2022). Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching*, *21*(1), 37–56. https://doi.org/10.1177/14757257211037149

Gilster, P. (1997). *Digital literacy*. John Wiley & Sons, Inc.

Hayes, A. F. (2012). *Process: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.* University of Kansas, KS.

Het algoritmeregister. (2024). *Remote sensing gebiedsclassificatie op basis van ai beeldherkenning.* Geraadpleegd op 17 februari 2025, van https://algoritmes.overheid.nl/nl/algoritme/remote-sensing-gebiedsclassificatie-op-basis-van-ai-beeldherkenning-provincie-zuidholland/87239212

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Khan, A. N., Jabeen, F., Mehmood, K., Soomro, M. A. & Bresciani, S. (2023). Paving the way for technological innovation through adoption of artificial intelligence in conservative industries. *Journal of Business Research*, *165*, 114019. https://doi.org/10.1016/j.jbusres.2023.114019

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature

review. *Quality & quantity*, *47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Laupichler, M. C., Aster, A., Haverkamp, N. & Raupach, T. (2023). Development of the "scale for the assessment of non-experts' ai literacy"–an exploratory factor analysis. *Computers in Human Behavior Reports*, *12*, 100338. https://doi.org/10.1016/j.chbr.2023.100338

Lintner, T. (2024). A systematic review of ai literacy scales. *npj Science of Learning*, *9*(1), 50. https://doi.org/10.1038/s41539-024-00264-4

Long, D. & Magerko, B. (2020). What is ai literacy? competencies and design considerations. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–16). https://doi.org/10.1145/3313831.3376727

Ma, S. & Lei, L. (2024). The factors influencing teacher education students' willingness to adopt artificial intelligence technology for information-based teaching. *Asia Pacific Journal of Education*, *44*(1), 94–111. https://doi.org/10.1080/02188791.2024.2305155

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. (2021). Conceptualizing ai literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F. & Chu, S. K. W. (2024). Design and validation of the ai literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, *55*(3), 1082–1104. https://doi.org/10.1111/bjet.13411

Ong, A. D. & Weiss, D. J. (2000). The impact of anonymity on responses to sensitive questions. *Journal of Applied Social Psychology*, *30*(8), 1691–1708. https://doi.org/10.1111/j.1559-1816.2000.tb02462.x

Rijksoverheid. (2024). *Overheidsbrede visie generatieve ai.* Geraadpleegd op 23 mei 2025, van `https://www.rijksoverheid.nl/documenten/rapporten/2024/01/01/overheidsbrede-visie-generatieve-ai`

Schiavo, G., Businaro, S. & Zancanaro, M. (2024). Comprehension, apprehension, and acceptance: Understanding the influence of literacy and anxiety on acceptance of artificial intelligence. *Technology in Society*, *77*, 102537. https://doi.org/10.1016/j.techsoc.2024.102537

Sonderen, E. v., Sanderman, R. & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PloS one*, *8*(7), e68967. https://doi.org/10.1371/journal.pone.0068967

Thurmond, V. A. (2001). The point of triangulation. *Journal of nursing scholarship*, *33*(3),

253–258. https://doi.org/10.1111/j.1547-5069.2001.00253.x

TNO. (2024). *Steeds meer kunstmatige intelligentie ingezet door overheid.* Geraadpleegd op 17 februari 2025, van https://www.tno.nl/nl/newsroom/2024/06/kunstmatige-intelligentie-inzet-overheid/

Upadhyay, N., Upadhyay, S. & Dwivedi, Y. K. (2022). Theorizing artificial intelligence acceptance and digital entrepreneurship model. *International Journal of Entrepreneurial Behavior & Research*, *28*(5), 1138–1166. https://doi.org/10.1108/IJEBR-01-2021-0052

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, 157–178. https://doi.org/10.2307/41410412

Verordening (EU) 2024/1689. (2024). Van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels voor kunstmatige intelligentie en tot wijziging van Verordeningen (EG) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 en (EU) 2019/2144 en Richtlijnen 2014/90/EU, (EU) 2016/797 en (EU) 2020/1828 (verordening artificiële intelligentie) (Voor de EER relevante tekst). Geraadpleegd op 17 februari 2025, van https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:32024R1689

Wang, B., Rau, P.-L. P. & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & information technology*, *42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Williams, M. D., Rana, N. P. & Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (utaut): a literature review. *Journal of enterprise information management*, *28*(3), 443–488. https://doi.org/10.1108/JEIM-09-2014-0088

Bijlagen

A Verantwoording literatuuronderzoek

Het theoretisch kader is opgebouwd op basis van een uitgebreid zoekproces. Voor het zoekproces is hoofdzakelijk gebruikgemaakt van Scopus, een wetenschappelijke database met peer-reviewed artikelen. Daarnaast is er gebruikgemaakt van Google Scholar om gericht te zoeken naar bepaalde artikelen. In het selectieproces zijn enkel Engelstalige artikelen geselecteerd. Er is gezocht door gebruik te maken van zowel losse zoektermen als combinaties van zoektermen.

Zoektermen

- AI literacy
- Artificial intelligence literacy
- AI competency
- AI skills
- AI understanding
- Digital literacy
- Public administration
- Public sector
- Government
- AI
- Artificial intelligence
- AI adoption
- AI acceptance
- AI use
- Artificial intelligence acceptance
- Artificial intelligence adoption
- Artificial intelligence use
- Technology acceptance
- Technological acceptance
- UTAUT
- TAM
- Technology acceptance model

Combinaties van zoektermen

- (“ai literacy” OR “artificial intelligence literacy” OR “ai knowledge” OR “ai competency” OR “ai skills” OR “ai understanding”)
- (“ai” OR “artificial intelligence”) AND (“acceptance” OR “adoption” OR “use”)
- (“ai” OR “artificial intelligence”) AND (“acceptance” OR “adoption” OR “use”) AND (“public sector” OR “government” OR “public administration”)
- (“ai literacy” OR “artificial intelligence literacy”) AND (“ai acceptance” OR “artificial intelligence acceptance” OR “technology acceptance” OR “technological acceptance”)

B Poweranalyse

F tests - Linear multiple regression: Fixed model, R² deviation from zero

Analysis:	A priori: Compute required sample size		
Input:	Effect size f^2	=	0.15
	α err prob	=	0.05
	Power (1-β err prob)	=	0 .80
	Number of predictors	=	5
Output:	Noncentrality parameter λ	=	13.8000000
	Critical F	=	2.3205293
	Numerator df	=	5
	Denominator df	=	86
	Total sample size	=	92
	Actual power	=	0.8041921

C Inleidende tekst survey

Beste deelnemer,

Mijn naam is Anne-Fleur Kremer en in het kader van mijn afstudeerscriptie voor de Master Beleid & Politiek aan de Erasmus Universiteit Rotterdam, doe ik onderzoek naar de relatie tussen AI-geletterdheid en de acceptatie en het gebruik van AI-systemen in de publieke sector. Ik voer dit onderzoek uit vanuit het team Digitaal Bestuur bij de Provincie Zuid-Holland onder begeleiding van Ivonne Jansen-Dings.

De survey zal beginnen met een aantal vragen over uw geslacht, leeftijd en de organisatie waarin u werkzaam bent. Daarna krijgt u een aantal vragen over uw gebruik van AI-systemen in uw werk. Tot slot, volgen een aantal stellingen die u kunt beantwoorden op een schaal van 'helemaal niet mee eens' tot 'helemaal mee eens'.

Het onderzoek neemt ongeveer 5 minuten in beslag. Voordat u deelneemt is het belangrijk dat u onderstaande informatie in acht neemt:

- Aan dit onderzoek zijn geen risico's verbonden.
- Uw deelname aan dit onderzoek is vrijwillig en u kunt op ieder moment uw deelname stoppen.
- De verzamelde data zal worden gebruikt voor een geaggregeerde data-analyse en vertrouwelijke informatie of persoonlijke gegevens zullen niet worden gebruikt in de uitkomsten van dit onderzoek.
- De data zal na afronding van het onderzoek voor een periode van 1 jaar bewaard blijven.

Voor deelname aan het onderzoek is het vereist dat u 18 jaar of ouder bent en werkzaam bent binnen het openbaar bestuur of de overheidsdiensten.

Als dank voor het invullen van de survey, kunt u kans maken op een VVV cadeaukaart ter waarde van 25 euro. Wanneer u kans wilt maken op deze cadeaukaart, kunt u uw e-mailadres aan het eind van de survey achterlaten. Uw e-mailadres wordt apart verwerkt en niet gekoppeld aan uw onderzoeksgegevens.

Voor vragen of opmerkingen kunt u met mij contact opnemen via 556748ak@student.eur.nl of a.kremer@pzh.nl.

Ik heb bovenstaande informatie begrepen en bevestig mijn deelname.

○

D Demografische vragen

Wat is uw geslacht?

- ○ Man
- ○ Vrouw
- ○ Anders

Wat is uw leeftijd (in jaren)?

_____

Bij welke provincie bent u werkzaam?

- ○ Provincie Zuid-Holland
- ○ Provincie Groningen
- ○ Provincie Friesland
- ○ Provincie Drenthe
- ○ Provincie Overijssel
- ○ Provincie Gelderland
- ○ Provincie Flevoland
- ○ Provincie Utrecht
- ○ Provincie Noord-Holland
- ○ Provincie Zeeland
- ○ Provincie Noord-Brabant
- ○ Provincie Limburg
- ○ Ik werk binnen een andere organisatie in de publieke sector

E Aanvullende vragen over gebruik van AI-systemen

Maakt u in uw werk weleens gebruik van een interne chatbot (speciaal ontwikkeld voor de organisatie waarin u werkzaam bent)?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van een openbare chatbot (zoals ChatGPT, Microsoft Copilot of vergelijkbare tool)?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van een afbeeldingsgenerator (zoals DALL-E, Midjourney of vergelijkbare tool)?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van een schrijfhulpmiddel (zoals Grammarly, DeepL translator of vergelijkbare tool)?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van een progammeerassistent (zoals GitHub Copilot of vergelijkbare tool)?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van een hulpmiddel om spraak automatisch om te zetten in tekst?

- ○ Ja
- ○ Nee

Maakt u in uw werk weleens gebruik van de suggesties die Outlook doet bij het beantwoorden van een e-mail?

- ○ Ja
- ○ Nee

F Resultaten factoranalyse

Tabel 10 Totale verklaarde variantie

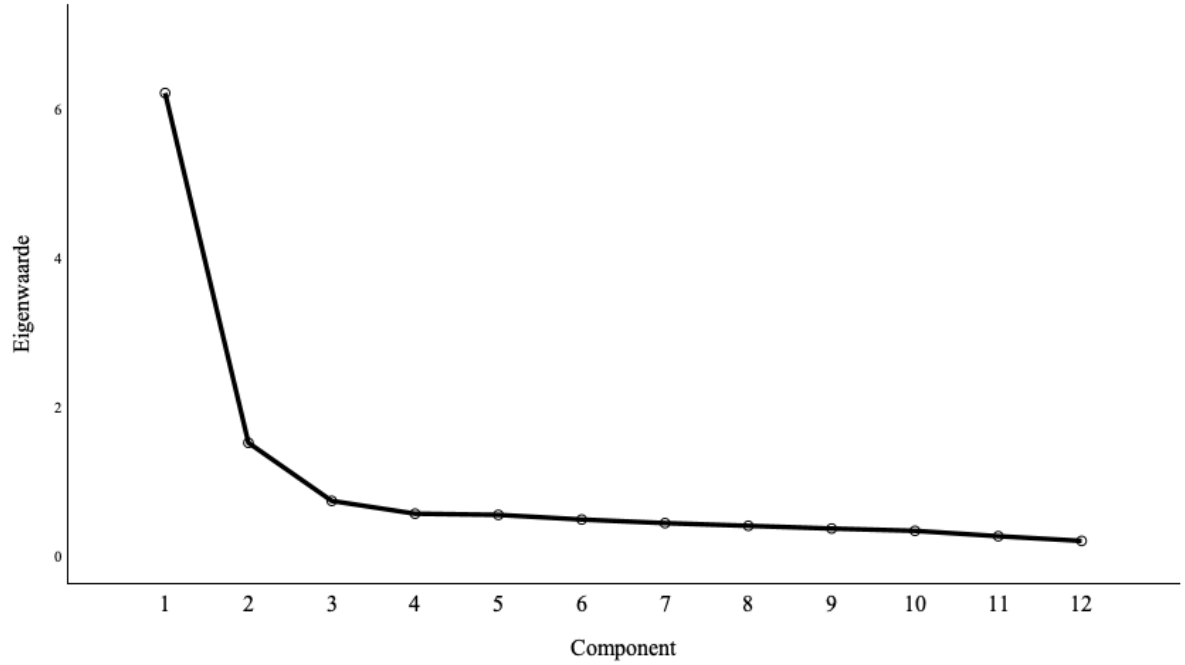
Component	Eigenwaarde	% van de variantie	Cumulatief %
1	6.26	52.13	52.13
2	1.47	12.25	64.38
3	.72	5.97	70.36
4	.56	4.68	75.03
5	.55	4.54	79.58
6	.48	4.03	83.61
7	.43	3.58	87.19
8	.40	3.31	90.50
9	.36	3.01	93.51
10	.33	2.75	96.26
11	.25	2.11	98.37
12	.20	1.63	100.00

Extractiemethode: Principal Component Analysis.

Tabel 11 Geroteerde componentenmatrix

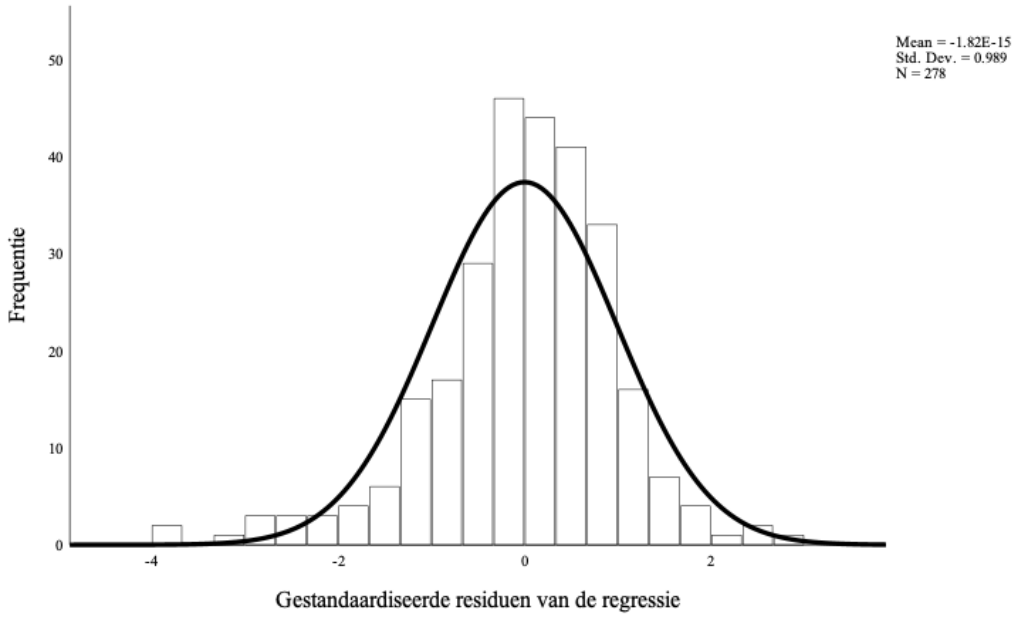
	Component	
	1	2
Ik kan slimme apparaten onderscheiden van niet slimme apparaten	.578	.407
Ik weet hoe AI-technologie mij kan helpen	.762	.358
Ik kan de AI-technologie identificeren die wordt gebruikt in de applicaties en producten die ik gebruik	.680	.297
Ik kan AI-applicaties of producten vaardig gebruiken om mij te helpen met mijn dagelijkse werk	.864	.124
Het is meestal makkelijk voor mij om een nieuwe AI-applicatie of product te leren gebruiken	.741	.267
Ik kan AI-applicaties of producten gebruiken om mijn werk efficiëntie te verbeteren	.768	.125
Ik kan de mogelijkheden en beperkingen van een AI-applicatie of product evalueren nadat ik het een tijdje heb gebruikt	.738	.261
Ik kan een geschikte oplossing kiezen uit verschillende oplossingen die worden aangeboden door een slimme agent	.727	.169
Ik kan de meest geschikte AI-applicatie of product kiezen uit een verscheidenheid voor een specifieke taak	.787	.151
Ik houd altijd de ethische principes in acht bij het gebruik van AI-applicaties of producten	.230	.796
Ik ben altijd alert op privacy- en informatiebeveiligingsproblemen bij het gebruik van AI-applicaties of producten	.186	.851
Ik ben altijd alert op misbruik van AI-technologie	.208	.826

Extractiemethode: Principal Component Analysis
Rotatiemethode: Varimax met Kaiser-normalisatie
Rotatie geconvergeerd in 3 iteraties

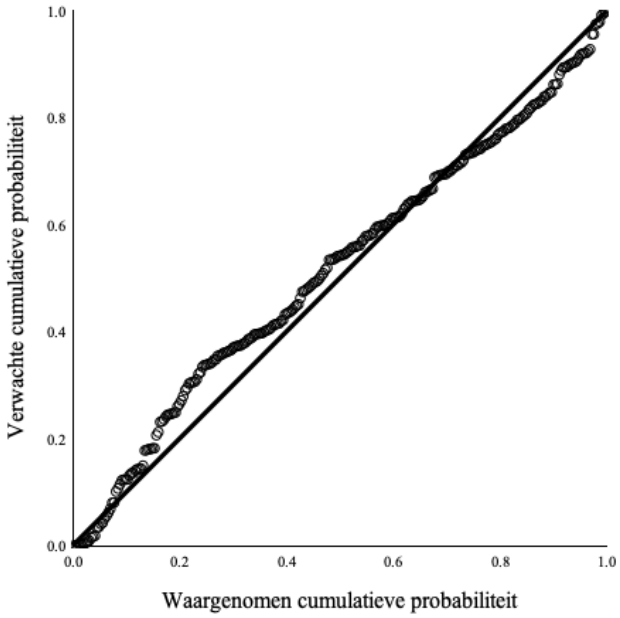


Figuur 6: Screeplot

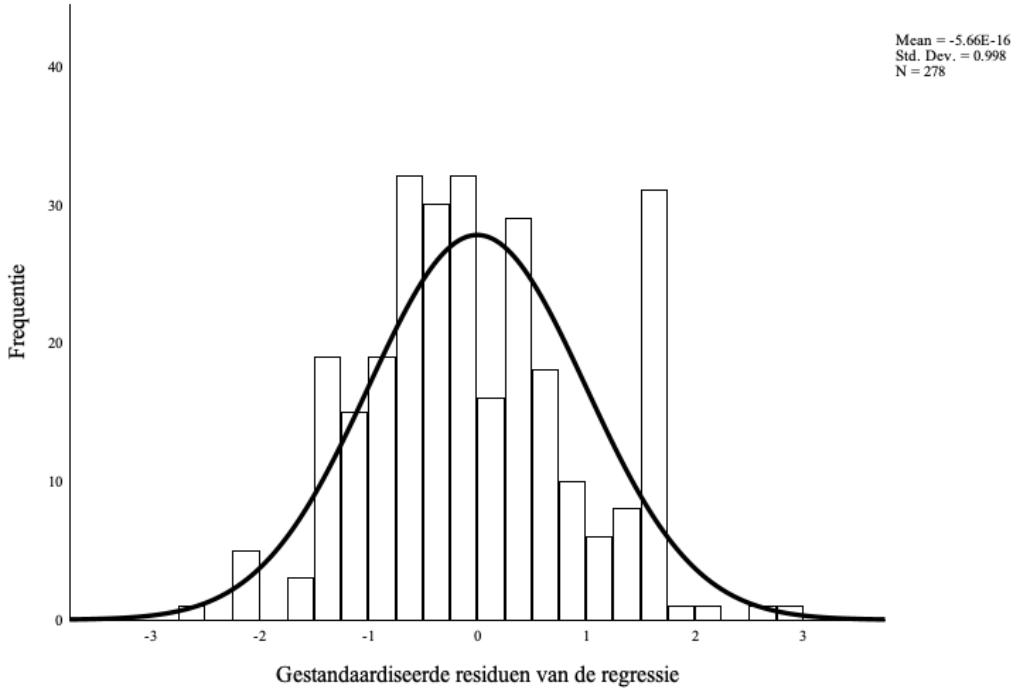
G Resultaten normaliteit van residuen



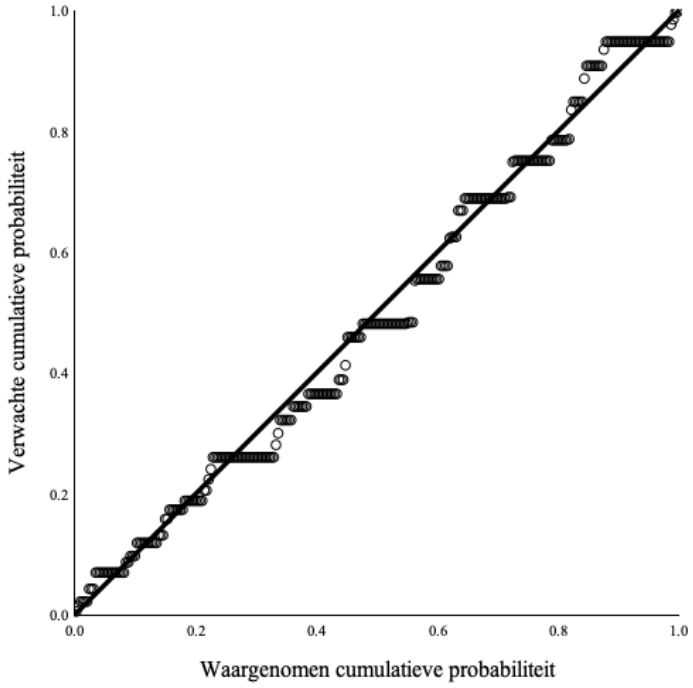
Figuur 7: Histogram van gebruikersintentie



Figuur 8: Normale P-P plot van de gestandaardiseerde residuen van de regressie voor de voorspelling van gebruikersintentie

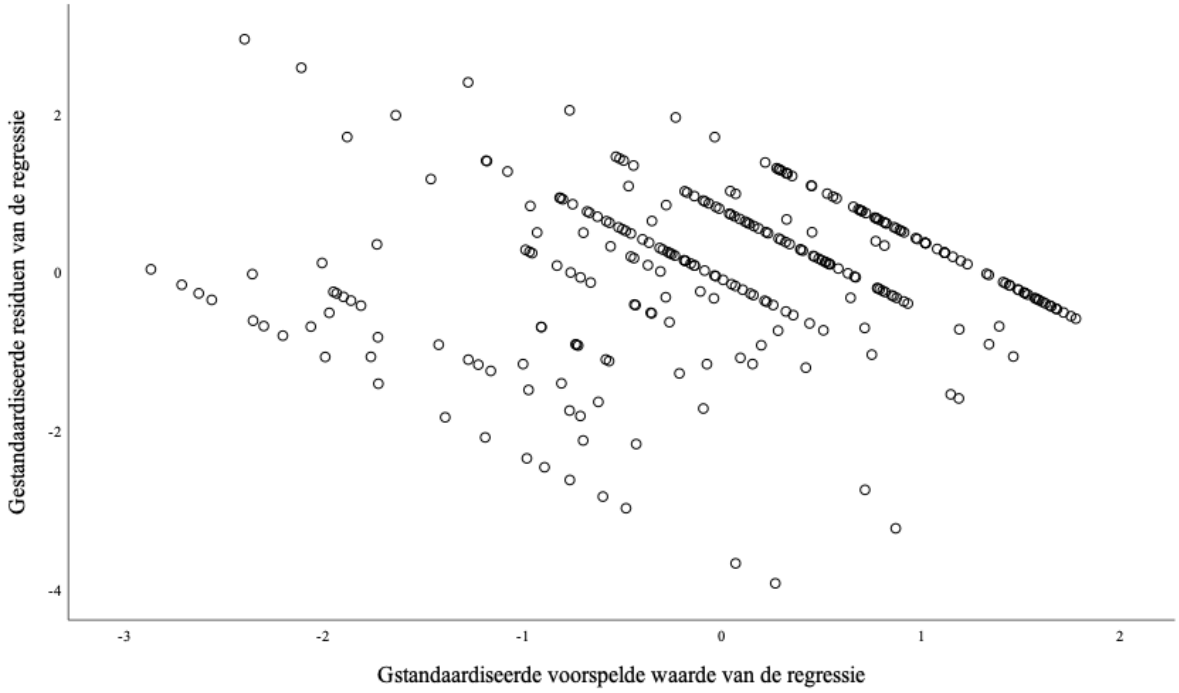


Figuur 9: Histogram van gebruikersgedrag

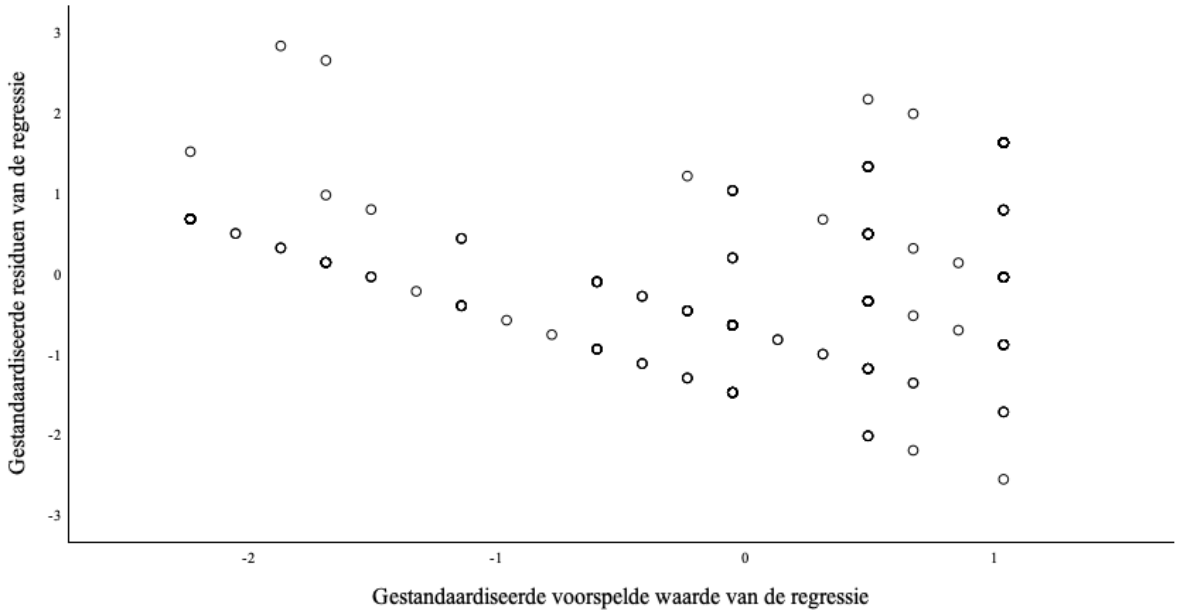


Figuur 10: Normale P-P plot van de gestandaardiseerde residuen van de regressie voor de voorspelling van gebruikersgedrag

H Resultaten homoscedasticiteit

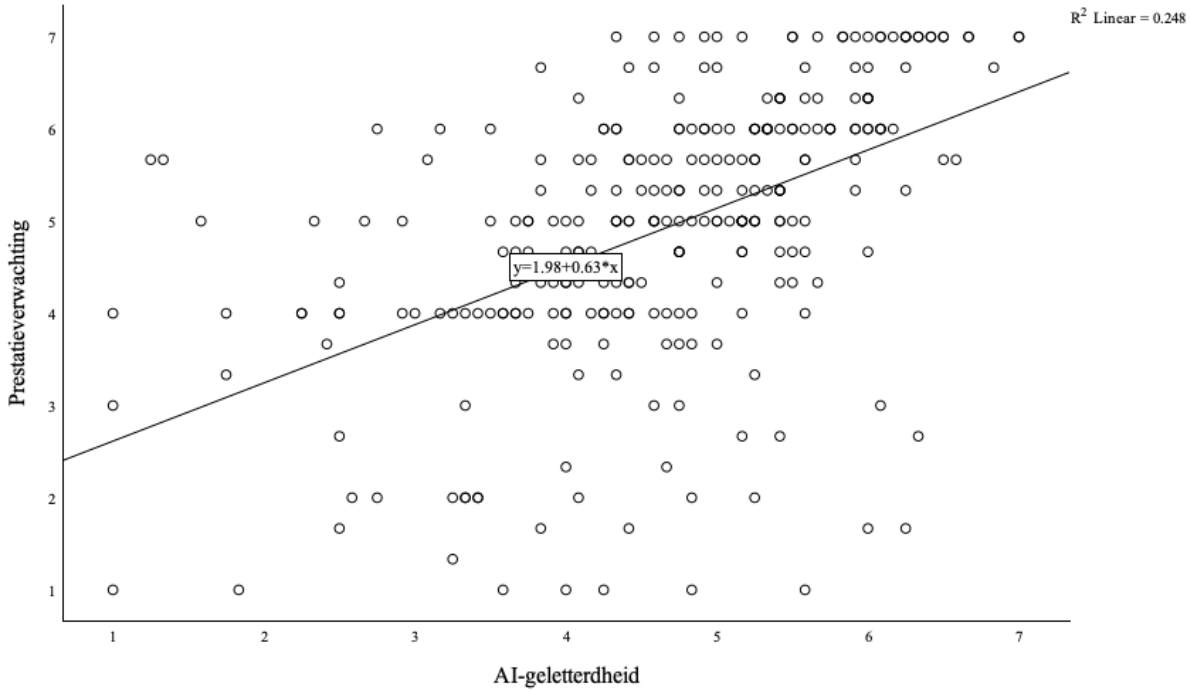


Figuur 11: Scatterplot gebruikersintentie

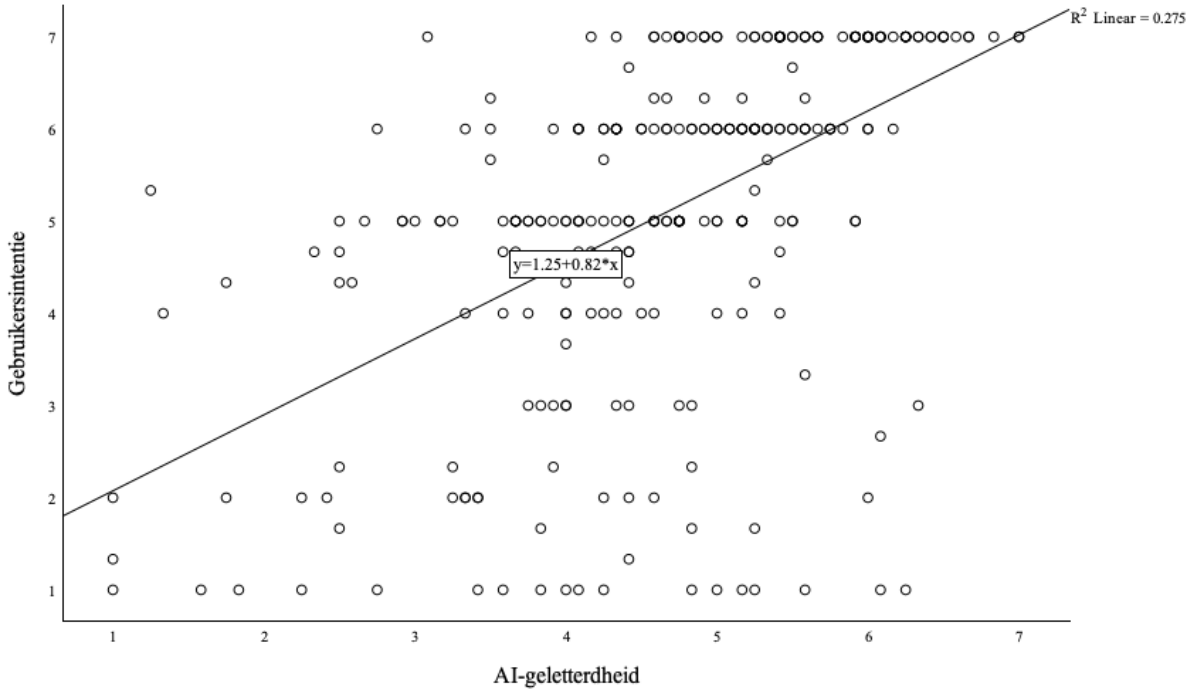


Figuur 12: Scatterplot gebuikersgedrag

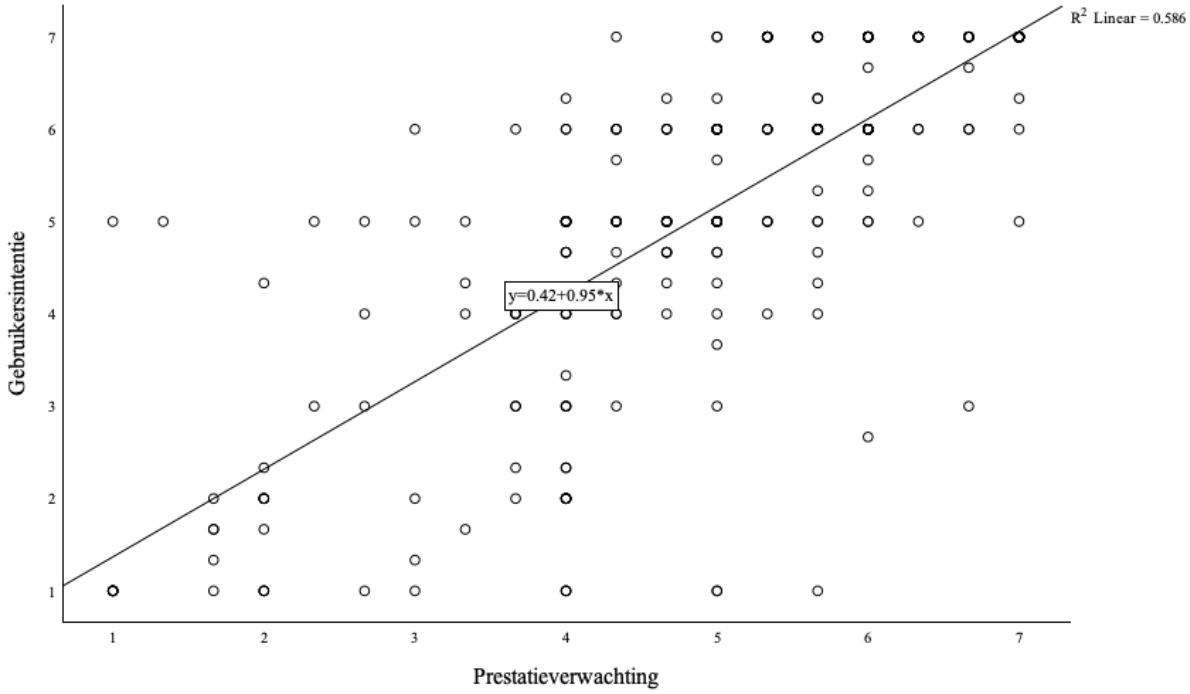
I Resultaten lineariteit



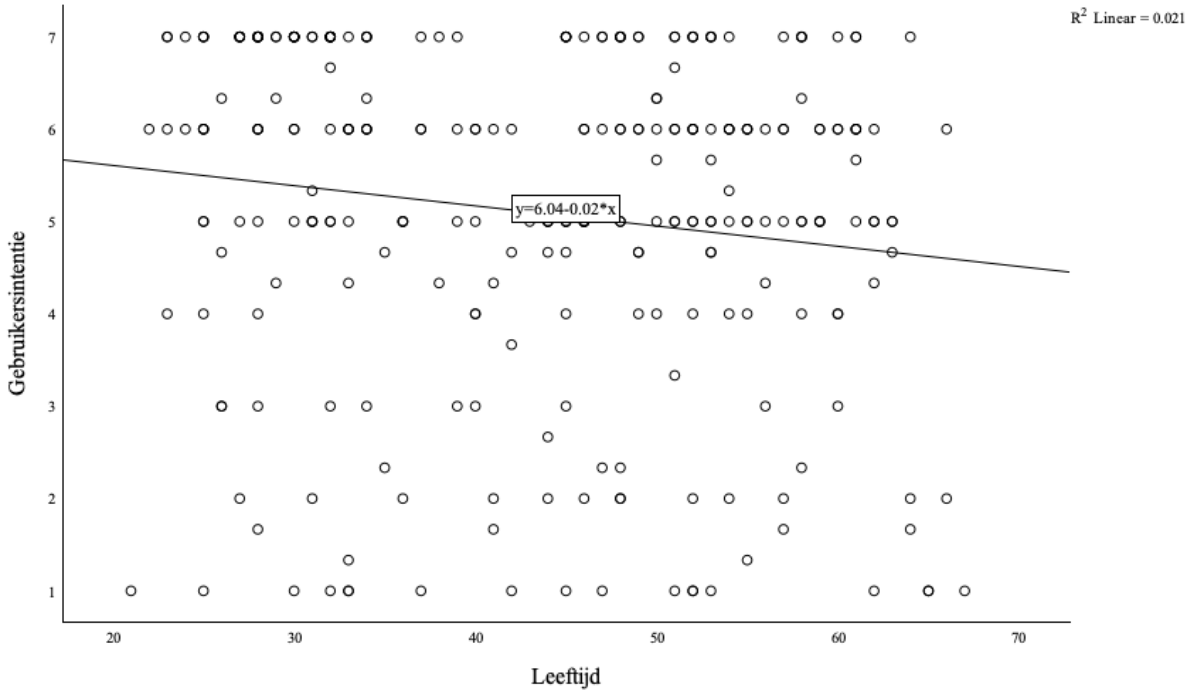
Figuur 13: Scatterplot AI-geletterdheid en prestatieverwachting



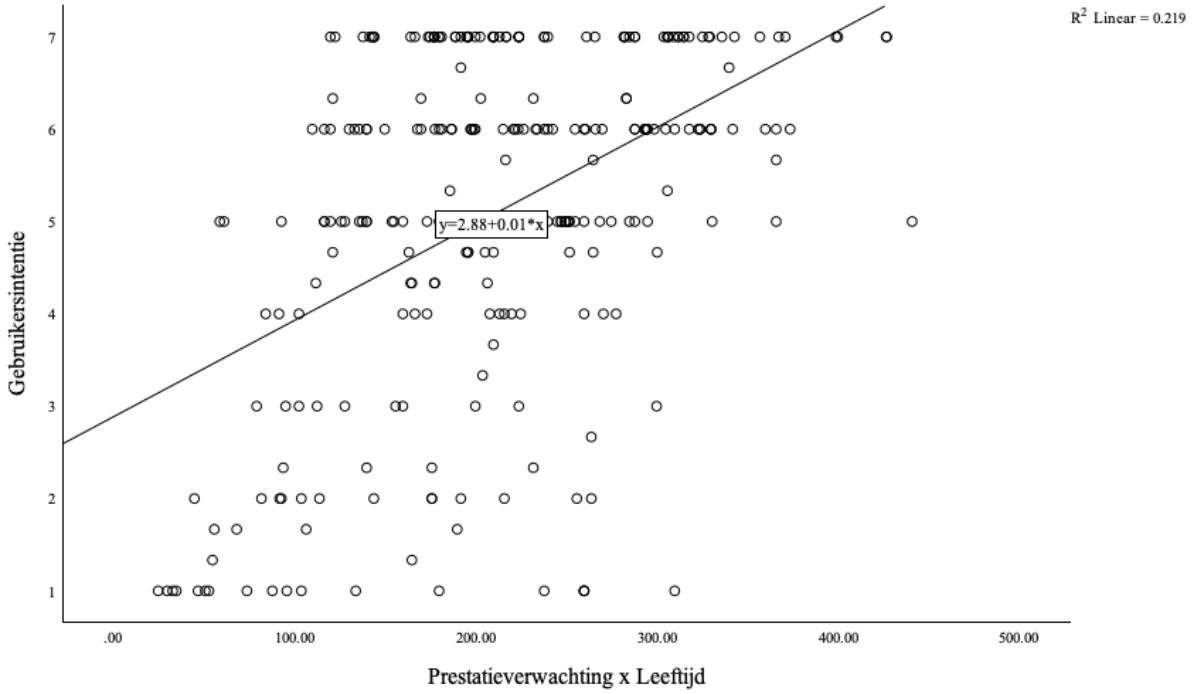
Figuur 14: Scatterplot AI-geletterdheid en gebruikersintentie



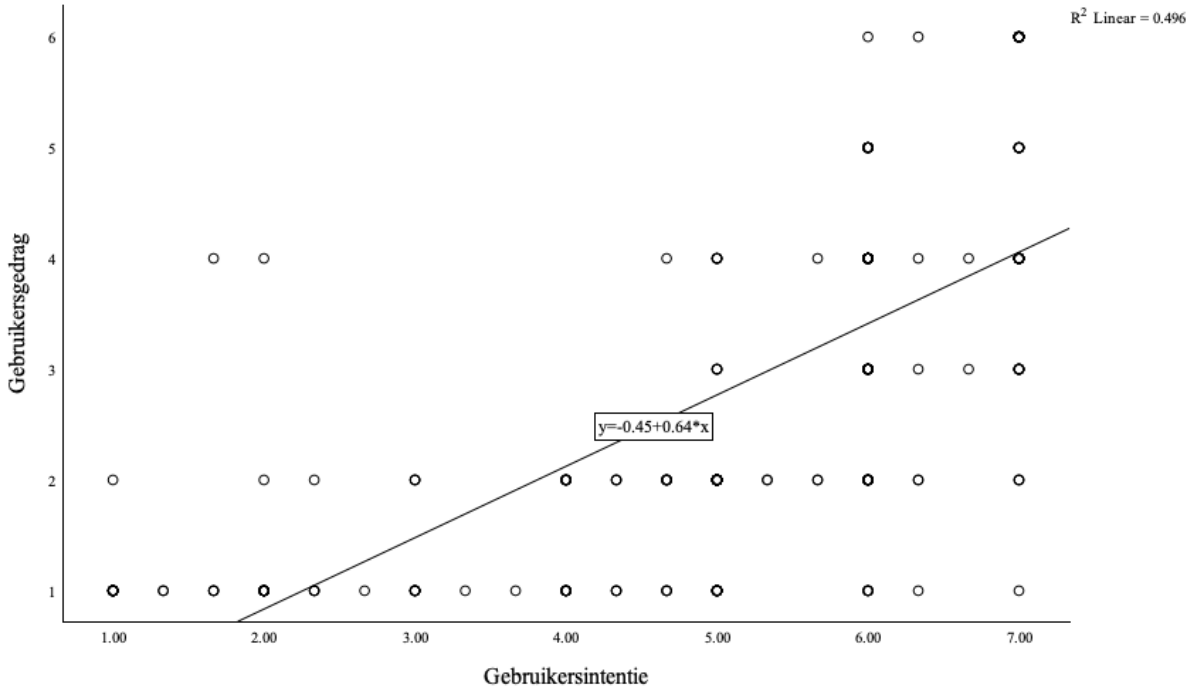
Figuur 15: Scatterplot prestatieverwachting en gebruikersintentie



Figuur 16: Scatterplot leeftijd en gebruikersintentie



Figuur 17: Scatterplot prestatieverwachting x leeftijd en gebruikersintentie



Figuur 18: Scatterplot gebruikersintentie en gebruikersgedrag

J Bootstrapresultaten regressieanalyses

Tabel 12 Bootstrapresultaten multiple regressieanalyse

	B	BootMean	BootSE	BootLLCI	BootULCI
Constante	-2.945	-2.958	.363	-3.699	-2.255
AI-geletterdheid	.632	.635	.074	.487	.785

Afhankelijke variabele: prestatieverwachting

Tabel 13 Bootstrapresultaten multiple regressieanalyse

	B	BootMean	BootSE	BootLLCI	BootULCI
Constante	3.422	3.410	.407	2.583	4.193
AI-geletterdheid	.338	.342	.084	.180	.513
Prestatieverwachting	.824	.819	.093	.631	.993
Gender	.161	.158	.140	-.119	.434
Gender x Prestatieverwachting	-.006	-.004	.101	-.200	.198
Leeftijd	.004	.004	.006	-.008	.015
Leeftijd x Prestatieverwachting	-.007	-.007	.004	-.016	.001

Afhankelijke variabele: gebruikersintentie

Tabel 14 Bootstrapresultaten lineaire regressieanalyse

	B	BootSE	BootLLCI	BootULCI
Constante	-.455	.183	-.841	-.129
Gebruikersintentie	.644	.038	.574	.723

Afhankelijke variabele: gebruikersgedrag