Zien we verschillen tussen wat echt of nep is? De rol van inductief leren en vertrouwen op de accuraatheid in het onderscheiden tussen AI-gegenereerde foto's en echte foto's

Merwe Oosterhof

Studentnummer: s5201632

Afdeling Psychologie, Rijksuniversiteit Groningen

PSB3A-BT15: Bachelor These
Supervisor: Ben Gützkow

Tweede beoordelaar: prof. Dr. Ernestine Gordijn

In samenwerking met: Lisa Bevers, Maaike de Kruijf, Meagan Kane, Moritz Noß, Tess Walvius

30 januari 2026

Een scriptie is een proeve van bekwaamheid voor studenten. De goedkeuring van de scriptie is het bewijs dat de student over voldoende onderzoeks- en rapportagevaardigheden beschikt om af te studeren, maar biedt geen garantie voor de kwaliteit van het onderzoek en de resultaten van het onderzoek als zodanig, en de scriptie is daarom niet per se geschikt als academische bron om naar te verwijzen. Als u meer wilt weten over het in deze scriptie besproken onderzoek en de daarop gebaseerde publicaties waarnaar u zou kunnen verwijzen, neem dan contact op met de genoemde begeleider.

Verklaring AI-gebruik

“Ik verklaar dat ik het vertaalprogramma Deepl Translator (DeepL SE, n.d.) heb gebruikt om materiaal te vertalen. De sectie methode is in dit onderzoek in groepsverband geschreven. Deze is in het Engels geschreven en dit is met deze hulp vertaald naar het Nederlands.”

Can we tell the difference between what is real and what is fake? The role of inductive learning and confidence on accuracy in distinguishing between AI-generated images and real images

Abstract

The rapid development of Artificial Intelligence (AI) has led to the emergence of increasingly hyperrealistic AI-generated images that are difficult to distinguish from real images. In the same time, people tend to overestimate their ability to distinguish between the images which can make people victims of deception or fraud via AI-images. This study investigated whether inductive learning, in the form of exposure to AI-generated and real images can improve distinguishing and whether this correspond with the reported confidence. Participants (N = 270) were randomly assigned to the training or control condition in an online experiment. This resulted in higher accuracy among participants in the inductive learning method, indicating that inductive learning enhances the ability to distinguish AI-generated from real images. Contrary to expectations confidence decreased significantly in the control condition, while in the training condition, it remained largely unchanged. This may suggest that the increased awareness of the complexity of AI leads to more accurate self-assessments. Given the rapid development of AI programmes, further research is needed to examine their effects and consequences. Implications for future research should focus on age effect, educational level, experience with AI, and long-term effects.

Keywords: AI-generated images, image distinguishing, inductive learning, confidence

Samenvatting

De razendsnelle ontwikkeling van Kunstmatige Intelligentie (AI) heeft geleid tot de ontwikkeling van steeds meer hyperrealistische AI-gegenereerde foto's, die moeilijk te onderscheiden zijn van de realiteit. Daarnaast blijkt er sprake te zijn van een overschatting in vertrouwen, wat mensen een slachtoffer van misleiding of oplichting via AI-foto's kan maken. Dit onderzoek onderzocht of inductief leren, in de vorm van blootstelling aan AI-gegenereerde en echte foto's, het onderscheiden kan verbeteren en of dit in lijn staat met het gerapporteerde vertrouwen. De geworven deelnemers ($N = 270$) werden in een online omgeving in de training- of controleconditie geplaatst. Dit resulteerde in een hogere accuraatheid van deelnemers in de inductieve leermethode, wat suggereert dat inductief leren werkt. Het vertrouwen scoorde in tegengestelde richting dan verwacht werd op basis van de literatuur, in de controleconditie daalde het vertrouwen flink en in de trainingsconditie bleef het nagenoeg gelijk. Wat mogelijk suggereert dat mensen kennis over de complexiteit van AI hebben, voor accuratere inschattingen. Door de toenemende ontwikkeling in AI-programma's blijft er onderzoek nodig naar de effecten en gevolgen ervan. Implicaties voor toekomstig onderzoek doelen op leeftijdseffect, opleidingsniveau, ervaring met AI en lange termijneffecten.

Trefwoorden: AI-gegenereerde foto's, foto's onderscheiden, inductief leren, vertrouwen

Zien we verschillen tussen wat echt of nep is? De rol van inductief leren en vertrouwen op de accuraatheid in het onderscheiden tussen AI-gegenereerde foto's en echte foto's

Elke dag scrollen we langs allerlei beelden en verhalen die we geloven, maar wat als wat we zien helemaal niet de echte waarheid is? Zie je ook weleens een beeld voorbijkomen op je telefoon waarbij je twijfels hebt? Of iets wat te perfect lijkt waardoor het nep kan zijn? We kunnen er niet meer omheen: de nieuwe technologie neemt toenemend verschillende aspecten van ons leven over. De opkomst van Kunstmatige Intelligentie (AI) is al jaren bezig, maar met de recente ontwikkelingen breidt het zich nu razendsnel uit. Programma's als Midjourney (Midjourney, 2025) and Dall e (OpenAI, 2025) maken het mogelijk dat we met een korte tekst prompt ineens een realistisch gezicht kunnen creëren. De gangbare toegankelijkheid van de programma's maakt het mogelijk voor vrijwel iedereen om AI-content te genereren. Daarnaast zijn de beelden die gecreëerd worden met AI-programma's met een korte blik tijdens het scrollen nauwelijks te onderscheiden van de realiteit (Karnouskos, 2020) en mensen overschatten hun eigen vermogen in het onderscheiden ervan (Bray et al., 2023). In de maatschappij zijn we nog niet voldoende ontwikkeld en ervaren om met deze AI-afbeeldingen om te gaan (Karnouskos, 2020). De combinatie van deze razendsnelle technologie en onwetendheid in de maatschappij kan een serieus probleem worden. Het brengt vele risico's met zich mee, waaronder fraude, oplichting, misleiding of beïnvloeding in de politiek. Zo kunnen criminelen de programma's inzetten om kwaadaardige software te ontwikkelen. Daarnaast kan het ook gebruikt worden voor manipulatie van verkiezingen of verspreiding van nepnieuws (Wat Is AI en Wat Zijn de Risico's?, 2024). De misleiding ligt op de loer, met name voor mensen die nog geen weet hebben over AI-programma's en haar functies. Dit is een kwetsbare doelgroep voor de risico's van deze nieuwe technologie.

In verschillende onderzoeken varieert het percentage van nauwkeurigheid in het juist onderscheiden tussen de 50% en de 60% (Bray et al. 2023; Cooke et al. 2025; Köbis et al. 2021; Somoray en Miller, 2023), wat duidt op net boven kansniveau. Hoewel het een recent en nieuw onderwerp lijkt, is de opgang al jaren bezig en dat betekent ook dat er vele eerdere onderzoeken naar dit onderwerp zijn gedaan. Het begrip AI-hyperrealisme is de wereld in gekomen, wat inhoudt dat AI-programma's gezichten kunnen genereren die soms eerder als menselijk worden beoordeeld dan zelfgemaakte foto's van echte mensen (Miller et al., 2023). Hier sluit het onderzoek van Cooke et al. (2025) bij aan, die concludeert dat het herkennen van gezichten zelfs nog moeilijker is dan het onderscheiden van AI-gegenereerde dieren, landschappen of voedsel. Het niet kunnen onderscheiden van de foto's vormt een probleem en zal verbeterd moeten worden. Daarnaast worden we tegenwoordig in steeds meer situaties geconfronteerd met AI-gegenereerde foto's en vanwege de bijbehorende risico's is het daarom van belang om dit onderwerp meer aandacht te geven. De onvoldoende capaciteiten en het overschatten van eigen zelfvertrouwen hierdoor suggereert het belang van training in deze context. Training zou kunnen helpen aan een beter vermogen om de AI-gegenereerde foto's van de echte foto's te onderscheiden en kan zorgen voor een realistischere zelfinschatting. Op basis daarvan staat de volgende vraag centraal; kunnen we mensen trainen door blootstelling te geven aan AI-gegenereerde foto's, om zo een onderscheid te leren maken tussen deze foto's? Er wordt gekeken naar of training hierin een effect heeft, of vertrouwen in eigen vermogens een rol speelt en of dit invloed heeft op elkaar.

Inductief leren

Hiervoor kan inductief leren, het leren van een nieuw concept of categorie door het observeren van voorbeelden (Kornell & Bjork, 2008), een mogelijke belangrijke rol spelen. Zonder dat we er zelf bewust van zijn, gebruiken we inductief leren voortdurend in het

dagelijks leven. Een duidelijk voorbeeld hiervan is het herkennen van gezichten bij baby's. Een van de eerste dingen die een baby leert is het herkennen van gezichten. Met het herhaaldelijk blootstellen van een aantal gezichten aan de baby, leert deze hieruit te onderscheiden. Dit wijst ons op een vroeg en onbewust leerproces, in deze context gericht op gezichten. Uit eerder onderzoek is de methode van inductief leren effectief gebleken (Kang en Pashler 2011; Kornell en Bjork 2008) en zou dit een belangrijke rol kunnen spelen bij het leren onderscheiden van foto's. Dus als we kunnen concluderen dat het van nature in mensen zit om met zelfs al heel kleine verschillen te leren om te onderscheiden, bij gezichten in het bijzonder, werkt dit dan ook bij het onderscheiden van AI-gegenereerde foto's van echte foto's? Het doel van de training is dat op basis van inductief leren deelnemers patronen gaan herkennen, die zullen helpen bij het onderscheiden tussen de AI-gegenereerde en echte foto's.

Dit leidt tot de eerste hypothese;

Hypothese 1; mensen die blootgesteld worden aan AI-gegenereerde foto's zullen vervolgens beter zijn in het onderscheiden tussen AI-gegenereerde foto's en echte foto's dan mensen zonder enige blootstelling vooraf.

Vertrouwen

Mensen die het meest vertrouwen hebben in hun eigen oordelen zijn kwetsbaarder als slachtoffer van oplichtingen via AI-foto's. Ondanks de negatieve consequenties bij fouten in onderscheiding blijken mensen nog niet veel beter te zijn dan het kansniveau in het onderscheiden van de foto's (Bray et al., 2023). De resultaten in dit onderzoek wijzen ook op een overschatting in vertrouwen bij de deelnemers. Eerdere onderzoeken ondervonden aansluitend dat de mensen die de meeste fouten maken bij het onderscheiden, ook het meest vertrouwen hebben in hun eigen oordelen (Köbis et al., 2021; Miller et al., 2023). Deze resultaten suggereren dat het gerapporteerde vertrouwen mogelijk niet accuraat

is, maar kan duiden op een overschatting in eigen beoordelingsvermogen. In lijn hiermee concludeerden Lindsey en Nagel (2015) dat de studenten er niet in slaagden alle vragen correct te identificeren die ze ofwel correct, ofwel fout beantwoord hadden. Dit impliceert gelijke bevindingen, wat kan resulteren in een niet-accurate relatie tussen prestatie en vertrouwen. Deze overschatting kan worden verklaard doordat mensen met enige incompetentie tegen twee problemen aanlopen; mensen maken fouten en trekken verkeerde conclusies, terwijl hun onvermogen het inzicht van deze fouten wegneemt (Kruger & Dunning 1999; Dunning et al. 2003). Dit kan leiden tot een verkeerde inschatting in vertrouwen. Anderzijds, wanneer mensen meer kennis bezitten, kunnen ze realistischer beoordelen welke vragen er correct of fout beantwoord zijn. Dit suggereert dat mensen met ook maar een beetje kennis van de complexiteit en vermogens van AI-programma's mogelijk voorzichtiger zijn, omdat ze weten dat ze makkelijk misleid kunnen worden.

Met opmerkingen [A1]: hmm ja. Zou je zeggen dat deze mensen eigenlijk niet voorzichtig zijn? Dat zou je argumentatie misschien iets veranderen, maar is toch wel nauwkeuriger. Gaat het nog over vertrouwen of overmoed? ('overconfidence') Herschrijf niet de hele paragraaf, maar misschien iets om nog even in te werken.

In de context van de AI-gegenereerde foto's, wijst dit op een risicovolle samenhang tussen vertrouwen en misleiding. Mensen die minder goed zijn in het maken van onderscheid tussen AI-gegenereerde en echte foto's hebben meer vertrouwen in eigen kunnen (Köbis et al., 2021; Miller et al., 2023). Volgens dit principe zouden mensen met sterkere capaciteiten voorzichtiger om kunnen gaan met beoordelingen door de mindere mate van vertrouwen. De groep met mindere capaciteiten heeft geen weet van eigen onwetendheid (Dunning et al., 2003) en hierdoor is de kans groter dat ze in de AI-gegenereerde foto's zullen trappen, met negatieve consequenties als gevolg. Idealiter zou dit patroon juist andersom moeten zijn. Gebaseerd op alle meegewogen resultaten verwachten we het volgende;

Hypothese 2; mensen met meer blootstelling aan de verschillen tussen AI-gegenereerde en echte foto's, zullen een afname in vertrouwen laten zien in vergelijking tot

voor de blootstelling. Terwijl de mensen zonder extra blootstelling een overschatting van eigen vermogen zullen laten zien.

Vertrouwen door training?

De relatie tussen blootstelling, accuraatheid en vertrouwen blijkt een complex verband. Enerzijds suggereert inductief leren dat blootstelling positief correleert met accuraatheid (Kornell & Bjork, 2008), wat wil zeggen dat er meer foto's juist worden onderscheiden. Aan de andere kant lijkt het dat vertrouwen niet altijd in één lijn loopt met accuraatheid, te schrijven aan de onwetendheid en onkunde. Daarnaast laat Chandler (1994) een gelijk patroon zien, het laten zien/oefenen met gerelateerde foto's zorgt voor een afname van de accuraatheid maar juist een toename in het zelfvertrouwen. De afname van de accuraatheid kan worden verklaard door een gevoel van bekendheid dat de deelnemers door de training ontwikkelen. Deze herkenning kan een misleidend gevoel geven bij het maken van de test, waarbij tegelijkertijd het vertrouwen toeneemt. Echter, worden er in de 14 meegenomen experimenten voornamelijk resultaten gebaseerd op de herkenning van bepaalde aanwijzingen, terwijl het ontdekken van patronen, zoals bij inductief leren, de leerweg en de accuraatheid kunnen verbeteren (Kornell & Bjork, 2008). Ook Bray et al. (2023) ondervonden dat een hoger vertrouwen niet altijd positief in relatie staat met een hogere accuraatheid.

Hierdoor verwachten we dat er een interactie ontstaat tussen blootstelling, accuraatheid en vertrouwen. Namelijk dat mensen die veel blootgesteld worden, beter zijn in het onderscheiden van de foto's. Tegelijkertijd zullen ze een patroon leren over de kleine, subtiele verschillen die tussen de AI-gegenereerde foto's en de echte foto's te zien zijn en merken ze hoe lastig het is en zal hun vertrouwen in eigen vermogen daarom ook afnemen.

Hypothese 3; de mensen die veel blootgesteld zijn aan AI-gegenereerde foto's zullen een significante afname in vertrouwen laten zien tussen voor de blootstelling van de foto's en erna, voordat de test begint, en zullen ondanks minder vertrouwen een hogere score op de test laten zien. De mensen zonder extra blootstelling zullen een overschatting van eigen vermogen in vertrouwen laten zien.

Methode

Deelnemers

In totaal hebben $N = 291$ deelnemers het onderzoek voltooid. Alle deelnemers moesten minimaal 16 jaar oud zijn om aan dit onderzoek deel te nemen. De meeste deelnemers werden geworven via het SONA-platform van de Rijksuniversiteit Groningen ($n = 242$), voornamelijk eerstejaars psychologiestudenten, die SONA-punten ontvingen voor hun deelname. Andere deelnemers werden geworven via de persoonlijke netwerken van de onderzoekers ($n = 49$); deze personen ontvingen geen vergoeding voor hun deelname. Alle deelnemers ontvingen een toestemmingsformulier, dat zij voorafgaand aan hun deelname aan het onderzoek hebben gelezen en ondertekend. In dit toestemmingsformulier werden hun rechten als deelnemers aan een onderzoek en de wijze waarop hun gegevens zouden worden verwerkt, uitgelegd. Deelname was vrijwillig en de gegevens werden anoniem bewaard.

Om de kwaliteit van de gegevens te waarborgen, werd een verplichte invultijd gehanteerd. Voor de berekening van deze invultijd werd de mediaan gebruikt. In de controlegroep was de mediane voltooiingstijd 1004 seconden. Deelnemers die het onderzoek in minder dan de helft van deze tijd of in meer dan drie keer deze mediane tijd voltooiden, werden uitgesloten van de analyse. Op basis van deze criteria werd één deelnemer uitgesloten, omdat deze te weinig tijd had genomen en werden er acht deelnemers uitgesloten van de analyse, omdat ze te veel tijd hadden genomen. In de

trainingsconditie was de mediane voltooiingstijd 1808 seconden. Onder toepassing van dezelfde uitsluitingscriteria zoals hiervoor werden twaalf deelnemers uit de trainingsconditie uitgesloten, omdat ze te lang over het onderzoek deden. Er werden geen deelnemers uit de trainingsconditie uitgesloten omdat ze te weinig tijd hadden genomen. Na toepassing van deze criteria bleven in totaal n = 270 deelnemers (203 vrouwen, 60 mannen, 5 non-binair, 2 niet bekendgemaakt) over in de uiteindelijke steekproef. Van deze deelnemers bleven n = 148 over in de controlegroep en n = 122 deelnemers bleven over in de trainingsgroep. De meeste deelnemers (n = 231) waren tussen de 18 en 24 jaar oud. De overige leeftijdscategorieën waren als volgt vertegenwoordigd: 25-34 (n = 14), 35-44 (n = 4), 45-54 (n = 11), 55-64 (n = 3) en 65+ (n = 7).

Het onderzoek werd door de richtlijnen van de ethische commissie als laag risico beschouwd en daarom vrijgesteld van verdere ethische toetsing. De gegevensverzameling begon op 6 november 2025 en eindigde op 16 november 2025 onder code PSY-2526-S-0030.

Ontwerp

Dit onderzoek maakte gebruik van een gemengd experimenteel ontwerp en werd online uitgevoerd met behulp van Qualtrics. Het ontwerp omvatte één factor tussen proefpersonen (training versus controle) en drie factoren binnen proefpersonen (emotie, aantrekkelijkheid en type afbeelding). Deelnemers werden willekeurig toegewezen aan ofwel de experimentele conditie (d.w.z. training) ofwel de controleconditie (geen training). De onafhankelijke variabelen omvatten de trainingsconditie, emoties (boos, blij, bang), aantrekkelijkheid (hoog of laag) en beeldtype (door AI-gegenereerd of echte foto). De gemeten afhankelijke variabelen waren accuraatheid, vertrouwen, theory of mind en aandacht voor details. De deelnemers werd gevraagd om de studie online in te vullen, die werd uitgevoerd met behulp van het Qualtrics-platform.

Materialen en procedure

Het experimentele ontwerp werd opgezet om te onderzoeken of inductief leren het vermogen van mensen om door AI-gegenereerde gezichten van echte gezichten te onderscheiden verbetert. Alle deelnemers voltooiden een test waarbij ze moesten bepalen welke foto door AI was gegenereerd en welke foto echt was. Deze afbeeldingen varieerden systematisch in emotie (boos, blij en bang) en aantrekkelijkheid, en waren evenwichtig verdeeld over geslacht en ras. De echte foto's waren afkomstig uit de Chicago Face database (Ma et al., 2021). Deze database werd gekozen omdat alle gezichten waren gelabeld en gecategoriseerd op basis van emoties, geslacht, ras en aantrekkelijkheid, waarbij de laatste was gebaseerd op een rangschikking door een grote steekproef (n = 100). Alle foto's in de database werden gelijkgemaakt, wat betekent dat de achtergrond werd verwijderd en alle sieraden werden verwijderd. De AI-afbeeldingen werden gegenereerd met Midjourney versie 7 (Midjourney, 2025) en Dall-e versie 3 (OpenAI, 2025). Deze modellen werden gekozen omdat ze gemakkelijk toegankelijk zijn en veel worden gebruikt, zodat ze representatieve foto's zouden genereren. Voorbeelden van gebruikte prompts zijn te vinden in de bijlage. De beslissing over welke foto's werden opgenomen, werd genomen op basis van een gezamenlijke overeenkomst (n = 6) op basis van fotorealisme, uitgedrukte emotie en niet overdreven aantrekkelijkheid. De foto's werden gekoppeld aan vergelijkbare afbeeldingen uit de database van Chicago, om relatief te matchen met het geslacht, ras en de aantrekkelijkheid. Beide AI-modellen voegden soms sieraden of een gekleurde achtergrond toe. Het gebruik van prompts om de afbeelding te verwijderen of te wijzigen leidde niet tot afbeeldingen die vergelijkbaar waren met de echte afbeeldingen. Daarom werden de AI-afbeeldingen bewerkt om de achtergrond en sieraden te verwijderen op een manier die vergelijkbaar was met de echte foto's.

De test bevatte 48 niet-gelabelde afbeeldingen, 24 AI-afbeeldingen en 24 echte afbeeldingen. Deze werden verder onderverdeeld in 12 mannen en 12 vrouwen, elk met de 3 soorten emoties. Voor elke emotie (boos, blij en bang) waren er 4 foto's, waarvan de helft als aantrekkelijk werd beschouwd en de andere helft als onaantrekkelijk. In de inductieve leerconditie werden in totaal 96 afbeeldingen getoond met een beschrijving of dit een AI-afbeelding of een echte afbeelding was, waarop de test volgde.

Vertrouwen

Met behulp van een Likertschaal werd het vertrouwen van de deelnemers in hun vermogen om afbeeldingen te classificeren driemaal beoordeeld in de experimentele situatie en tweemaal in de controleconditie. Dit werd gemeten vóór de training (“Hoe zeker bent u dat u door AI-gegenereerde gezichten kunt onderscheiden van echte menselijke gezichten?”), na de training (“Hoe zeker bent u nu, na het voltooien van de training en het bekijken van voorbeelden van AI-gezichten en echte gezichten, dat u door AI-gegenereerde gezichten kunt onderscheiden van echte menselijke gezichten?”) en na het voltooien van de test (“Hoe zeker bent u er na het voltooien van de test van dat u AI-gegenereerde gezichten nauwkeurig van echte menselijke gezichten hebt kunnen onderscheiden?”) door middel van een 6-punts Likertschaal (1 = helemaal niet zeker, 6 = zeer zeker). Er werd gekozen voor een 6-punts Likertschaal in plaats van een standaard 7-puntsschaal om neutrale antwoorden te vermijden.

Demografische kenmerken

De deelnemers werd gevraagd hun geslacht, leeftijdsgroep en voorkeurstaal (Nederlands/Engels) aan te geven. De bijbehorende vraag voor geslacht was: “Wat is uw geslacht?”, met als antwoordmogelijkheden: man, vrouw, non-binair, anders of zeg ik liever niet. Ook werd er naar de leeftijd gevraagd, die werd ingedeeld in groepen: 18-24,

25-34, 35-44, enzovoort, tot 65+. Om ethische redenen werd voor elke optie de mogelijkheid geboden om dit soort informatie niet bekend te maken.

Accuraatheid

Voor de beoordeling of het een AI-gegenereerd gezicht of een echt gezicht is, vroegen we bij elke getoonde afbeelding: “Deze afbeelding is:” met twee antwoordopties. “Door AI gegenereerd” of “Een echte foto” met een toegestane antwoordtijd van 15 seconden. In de inductieve leersituatie voltooiden de deelnemers een trainingssessie waarin ze een reeks gezichten bekeken die correct gelabeld waren als door AI-gegenereerd of een echte foto. Deze werden afwisselend getoond om de deelnemers te helpen de visuele verschillen tussen beide categorieën op te merken. Deelnemers in de controlegroep kregen deze training niet en gingen direct door naar de test.

Na de enquête kregen de deelnemers de gelegenheid om opmerkingen of feedback over het experiment te geven. Ten slotte kregen ze een bericht te zien waarin hen werd bedankt voor hun deelname, waarmee de procedure werd afgesloten.

In de enquête kwamen nog twee andere onderwerpen aan bod; visuele perceptuele vaardigheden en theory of mind. Deze onderwerpen komen niet aan bod in dit onderzoek, maar namen tevens wel deel uit van de enquête.

Resultaten

Hier worden de resultaten van het onderzoek naar het effect van blootstelling op het juist onderscheiden van foto’s en vertrouwen gerapporteerd. In tabel 2 is te zien dat het gemiddelde percentage van juist beantwoordde foto’s hoger ligt in de trainingsconditie (M = 42,0; SD = 4,3), ten opzichte van de controleconditie (M = 35,7; SD = 7,0), dit wijst ons op een positief effect van de training. In figuur 1 is te zien dat de laagste scores van de trainingsconditie vele plaatsen hoger ligger dan de laagste scores in de controleconditie. De

resultaten in hoeveel vertrouwen er in beide condities waren laat interessante uitkomsten zien. Verwacht werd dat het vertrouwen in de trainingsconditie omlaag zou gaan en dit in de controleconditie gelijk zou blijven. Het gemiddelde vertrouwen voor de training ligt lager (*M* = 3,6; *SD* = 1,0), dan het vertrouwen na de training (*M* = 3,8; *SD* = 1,2) en dan na de test (*M* = 3,7; *SD* = 1,2). In de controleconditie is er echter een groter verschil tussen het vertrouwen voor de test (*M* = 3,4; *SD* = 1,1) en na de test (*M* = 2,8; *SD* = 1,1). Hieronder worden enkele correlaties tussen de variabelen weergegeven (tabel 1).

Tabel 1

Met opmerkingen [A2]: Er zijn op dit moment nog decimaalpunten in plaats van komma's!

Pearson'sCorrelations

Variable		Total_Pictures_Correct	Age	Gender	Confidence_pre
1. Total_Pictures_Correct	Pearson's r				
	p-value				
2. Age	Pearson's r	-0,398			
	p-value	< ,001			
3. gender	Pearson's r	0,074	0,099		
	p-value	0,225	0,106		
4. Confidence_pre	Pearson's r	0,107	-0,106	0,266	
	p-value	0,251	0,211	0,004	
5. Confidence_After_Training	Pearson's r	0,330	-0,106	0,217	0,481
	p-value	< ,001	0,243	0,016	< ,001

Hypothesetoetsen

Wat betreft *Hypothese 1;* mensen die blootgesteld worden aan AI-gegenereerde foto's zullen vervolgens beter zijn in het onderscheiden tussen echte foto's en AI-gegenereerde foto's dan mensen zonder enige blootstelling vooraf. Er werd gekeken naar het effect van training op de accuraatheid in het onderscheiden van AI-gegenereerde foto's. Dit werd getest met behulp van een onafhankelijke t-test, waarvoor de assumpties vooraf werden gecontroleerd. Het bracht een significant resultaat (p < ,001), voor zowel de

Shapiro-Wilk test als de Levene's test. Gezien de voldoende grote steekproef (>250) (Field, 2018) en gegeven dat de onafhankelijke t-test robuust is tegen ongelijke varianties, is er gekozen voor de Welch's onafhankelijke t-test. Deze test toont aan dat de deelnemers die blootstelling hebben gehad, beter waren in het onderscheiden van de foto's dan de deelnemers zonder blootstelling vooraf (t = -9,014; p <,001; d = -1,077). Deze resultaten, te zien in tabel 3, ondersteunen de hypothese.

Met opmerkingen [A3]: Kijk uit naar decimaalpunten. Deze moeten wel komma's zijn

Ter opzicht van *Hypothese 2;* mensen met meer blootstelling aan de verschillen tussen AI-gegenereerde en echte foto's, zullen een afname in vertrouwen laten zien in vergelijking tot voor de blootstelling. Terwijl de mensen zonder extra blootstelling een overschatting van eigen vermogen zullen laten zien. Er werd gekeken naar het verschil in vertrouwen tussen voor en na de training bij de deelnemers. Na het controleren van de assumpties is er gekozen voor een gepaarde t-test. Deze vergelijking geeft een ander resultaat dan verwacht, het vertrouwen voor de training (M = 3,6; SD = 1,0) ligt lager dan na de training (M = 3,8; SD = 1,2) (tabel 4). Echter, is dit geen significant verschil (t = -1,4; p = ,162; d = -0,131) (tabel 5). Daarnaast ligt ook tegen verwachting in het gemiddelde na de test in de controleconditie (M = 3,2; SD = 1,3) lager dan voor de test (M = 3,4; SD = 1,1). Dit verschil is wel significant (t = 6,005; p < ,001; d = 0,500). Voor een duidelijker beeld, te zien in figuur 2.

Om te kijken of blootstelling zorgt voor een afname in vertrouwen (*Hypothese 3)*; de mensen die veel blootgesteld zijn aan AI-gegenereerde foto's zullen een significante afname in vertrouwen laten zien tussen voor de blootstelling van de foto's en erna, voordat de test begint, en zullen ondanks minder vertrouwen een hogere score op de test laten zien. De mensen zonder extra blootstelling zullen een overschatting van eigen vermogen in vertrouwen laten zien. Werd er een Repeated Measures ANOVA uitgevoerd. Met tijd van

meting (voor of na de test/training) als binnen-subject factor en conditie (training of controle) als tussen-subject factor.

Het resultaat duidt op een significante verandering in vertrouwen over tijd in beide condities, $F(1, 248) = 21{,}01$; $p < {,}001$ (tabel 6). Echter, is er in beide condities een verschil in tegengestelde richting. Het vertrouwen in de trainingsconditie blijft nagenoeg gelijk, met een lichte stijging. Het vertrouwen in de controleconditie daalt flink. De interactie tussen tijd van meting en vertrouwen is aanwezig (figuur 3), desalniettemin in een andere richting dan eerder verwacht.

Verkennend onderzoek

In een verkennend onderzoek werd gekeken naar of het apparaat dat gebruikt wordt (laptop, tablet, telefoon) invloed heeft op de score die behaald wordt. Het grootste deel (n = 203) (tabel 7) heeft gebruikt gemaakt van een laptop, daarnaast heeft ook een redelijk groot deel de test gedaan op een telefoon (n = 64). Dit betekent dat het onderzoek op een kleiner scherm is uitgevoerd, met een kleinere resolutie die de zichtbaarheid vermindert. Dit zou mogelijk invloed kunnen hebben op de verschillen in de scores. Een ANOVA liet zien dat er een significant resultaat is, $F(2, 267) = 3{,}304$; $p = {,}038$; $\eta 2 = 0{,}024$ (tabel 8). Waaruit blijkt dat het device waar de test op gemaakt is, een significant maar zeer klein verschil maakt, positief voor de laptopgebruikers (figuur 4).

Discussie

In dit onderzoek werd er onderzocht of training een positief effect heeft op het juist onderscheiden van AI-gegeneerde foto's versus echte foto's en of vertrouwen een rol speelt in accuraatheid. De deelnemers werden verdeeld over twee groepen (controle en training) en kregen tijdens de test gelijke foto's te zien waarbij de keuze gemaakt moest worden tussen een AI-gegenereerde foto of een echte foto. Beide groepen kregen voor en na de test

de vraag hoeveel vertrouwen men had in eigen vermogen in het juist onderscheiden van de foto's. De trainingsconditie kreeg deze vraag ook voor en na de training. Deze gegevens werden verzameld en statistisch berekend.

Resultaten en implicaties

De resultaten ondersteunen de eerste hypothese, die stelt dat; mensen die blootgesteld worden aan AI-gegenereerde foto's vervolgens beter zullen zijn in het onderscheiden tussen AI-gegenereerde foto's en echte foto's dan mensen zonder enige blootstelling vooraf. Te zien was dat de groep in de trainingsconditie (M = 42,0) significant hoger scoorde dan de groep in de controleconditie (M = 35,7), wat suggereert dat inductief leren werkt. Dit staat in lijn met de literatuur, die ook bevestigde dat inductief leren een positief effect heeft (Kang en Pashler 2011; Kornell en Bjork 2008). Ondanks de korte training is het effect aanwezig en bevestigt dit dat blootstelling een positief effect heeft op de accuraatheid in het onderscheiden.

De tweede hypothese: mensen met meer blootstelling aan de verschillen tussen AI-gegenereerde foto's en echte foto's, zullen een afname in vertrouwen laten zien in vergelijking tot voor de blootstelling. Terwijl de verwachting was dat mensen zonder extra blootstelling een overschatting van eigen vermogen zullen laten zien, is dit niet bevestigd.

Resultaten wijzen op een effect de andere richting op dan verwacht, tevens was dit effect niet significant. In de controlegroep werd wel een significant effect gevonden, echter ook in de tegengestelde richting. Zoals uit de literatuur bleek, zouden lagere performers een hoger vertrouwen aangeven, door de onwetendheid over eigen kennis (Dunning et al., 2003). Uit resultaten bleek dat juist de controleconditie, die de lagere scores had t.o.v. de trainingsconditie flink daalde in vertrouwen. Een mogelijke verklaring hiervoor is dat tegenwoordig het grootste deel van de bevolking al wel ergens in aanraking

is gekomen met de AI-programma's door de razendsnelle toename ervan. Zodoende hebben mensen al ervaring met AI, waardoor het Dunning-Kruger effect (Dunning et al., 2003) niet van toepassing is, dat namelijk meer gericht is op mensen met extreem lage competentie of met lage scores. Dit kan suggereren dat mensen in aanraking zijn gekomen met de complexiteit van AI, waardoor ze zichzelf niet te hoog inschatten, wat de lagere scores op vertrouwen kan verklaren. Hierop aansluitend kan ook ervaring met AI-programma's een rol spelen. Uit onderzoek van Högemann et al. (2025) blijkt dat mensen met minder ervaring in gebruik van AI-programma's voorzichtiger zijn in het oordelen, in vergelijking tot mensen met meer ervaring. Dit geeft een relatie tussen accuraatheid en vertrouwen die overeenkomt met onze resultaten. Interessant genoeg verwachtte ook Somoray en Miller (2023) op basis van eerdere onderzoeken een negatieve relatie tussen vertrouwen en prestatie te vinden. Echter, werd er zelfs een kleine positieve relatie gevonden. Een mogelijk genoemde oorzaak hiervoor is de zogeheten middle-response bias, die inhoudt dat mensen liever kiezen voor de 'veilige' middeloptie, dan voor een extremer antwoord. Dit zou de relatie in vertrouwen in de trainingsconditie kunnen verklaren, die nagenoeg gelijk blijft. Hoewel er in onze enquête is gekozen voor een 6-punts Likertschaal, waarmee de optie om voor het middelste antwoord te kiezen wegvalt, is er niet uit te sluiten dat de deelnemers hebben gekozen voor de relatief 'veiligere' middelopties in plaats van de extremere antwoorden. Dit kan de kleine verschillen in vertrouwen verklaren.

Er is een interactie tussen de tijd van meting en vertrouwen gevonden, wat suggereert dat er verschillen zijn gevonden in het vertrouwen over tijd van de deelnemers. Dit werd getest in hypothese 3: de mensen die veel blootgesteld zijn aan AI-gegenereerde foto's zullen een significante afname in vertrouwen laten zien tussen voor de blootstelling van de foto's en erna, voordat de test begint, en zullen ondanks minder vertrouwen een

hogere score op de test laten zien. De mensen zonder extra blootstelling zullen een overschatting van eigen vermogen in vertrouwen laten zien. Het effect bleek significant, echter verschilt ook hier de richting ten opzichte van de verwachting. Het vertrouwen in de controleconditie daalt flink, te wijten aan het feit dat deelnemers mogelijk zichzelf voor het begin van de test overschatten, te wijten aan het fenomeen onwetendheid over complexiteit, en tijdens de test is er ervaren dat de AI-gegenereerde foto's moeilijker te onderscheiden zijn dan verwacht. Tevens is er ook een verandering in vertrouwen in de trainingsconditie gevonden, echter verschilt deze minimaal met de meting voor de training. Dit suggereert dat de deelnemers een realistische inschatting van eigen vertrouwen hebben gemaakt. De kleine stijging in vertrouwen kan te wijten zijn aan het feit dat inductief leren werkt (Kang en Pashler 2011; Kornell en Bjork 2008). De deelnemers hebben door het leren een subtiele stijging in vertrouwen opgedaan, maar zijn er ondanks dat bewust van geworden hoe complex de verschillen kunnen zijn, waardoor het vertrouwen niet sterk toe is genomen. Een andere mogelijke oorzaak hiervoor is dat mensen denken dat ze bovengemiddeld presteren bij makkelijke taken (Burson et al., 2006). Deze richting is in het effect van vertrouwen te zien in de trainingsconditie, die vooraf blootstelling heeft gehad. Anderzijds, bij taken die als moeilijk worden ervaren, denken mensen dat ze ondergemiddeld scoren, wat de daling in vertrouwen in de controleconditie kan verklaren.

Als we kijken naar het algemene beeld, dan laten de resultaten zien dat inductief leren werkt. Ook vertrouwen speelt een rol bij zowel de controle- als de trainingsconditie. Er is een interactie tussen vertrouwen en de tijd van meten in deze beide groepen. In de controlegroep daalt het vertrouwen sterk en bij de training stijgt het minimaal.

Praktische implicaties

Er is een razendsnelle ontwikkeling in de opkomst van AI-programma's. Dit betekent dat er steeds meer modellen zullen uitkomen en deze zullen verbeteren. Dit suggereert dat het steeds moeilijker zal zijn om onderscheid te kunnen maken, door toegenomen complexiteit van de modellen. Anderzijds betekent dit ook dat er een toename zal zijn van het aantal AI-foto's dat overal verschijnt, die door mensen gezien worden. Dit kan een positief effect opleveren, zodra men doorheeft dat dit daadwerkelijk AI-foto's zijn. Zodra dit bewustzijn er is, werkt onbewust het effect van inductief leren voort. Men wordt meer blootgesteld aan de foto's en leert zo onderscheid te blijven maken.

Inductief leren is succesvol bewezen, desalniettemin vereist het enige training. Dit impliceert dat herhaalde educatie erg waardevol kan zijn. Hierin zal van belang zijn aan te geven welke foto door AI is gegenereerd en welke echt is, in combinatie met feedback die gegeven wordt om zelf patronen te leren ontdekken.

Beperkingen, sterke punten & suggesties voor toekomstig onderzoek

In dit onderzoek zijn ook een aantal kanttekeningen. De deelnemers zijn grotendeels geworven door het platform SONA, het online platform voor psychologiestudenten van de RUG. Dit heeft als gevolg een selecte steekproef, wat resulteert in nauwelijks variatie in leeftijd en opleidingsniveau.

Tevens is er in dit onderzoek alleen gebruikgemaakt van gezichten, die de deelnemers moesten onderscheiden. De gecreëerde gezichten zijn geprobeerd expres zo moeilijk mogelijk te maken, om te compenseren voor de onjuistheid die AI geeft. AI-programma's kunnen er wat van om onnatuurlijke foto's (te perfecte huid, niet-realistische belichting) te creëren op basis van prompts. Hiervoor moest gecompenseerd worden door

gezichten zo realistisch mogelijk te maken. Dit is gedaan in de twee verschillende programma's Midjourney versie 7 (Midjourney, 2025) en Dall-e versie 3 (OpenAI, 2025). Door de razendsnelle ontwikkeling van AI-programma's is het lastig om hier conclusies aan te verbinden. Dit is het gevolg van de nieuwe ontwikkelingen, die zelfs in de tijd van dit onderzoek uitgevoerd zijn in de AI-programma's. Voor dit onderzoek houdt dat in dat de test niet exact repliceerbaar is, evenals de modellen die op gelijke prompts verschillende afbeeldingen creëren. Aansluitend daarop is de volgende stap in deze ontwikkelingen waarschijnlijk de toename van video's, gecreëerd door AI-programma's. Vandaag de dag komen er steeds meer boven water, daar waar ze nu nog opvallen omdat ze zo duidelijk onderscheidbaar zijn, zullen ze met de tijd alleen maar realistischer worden. Dan stellen wij ons de vraag; blijft inductief leren voldoende om jezelf te beschermen tegen de misleiding die AI ons voordoet?

Anderzijds zijn er ook een aantal sterke punten te benoemen. Er heeft een groot aantal deelnemers geparticipeerd aan het onderzoek en de deelnemers zijn willekeurig in een conditie geplaatst. Daarnaast is het onderzoek via een online platform uitgevoerd. Wat inhoudt dat, ondanks een eventuele ruis (d.w.z. geen controleerbare setting), deelnemers het experiment mochten uitvoeren op hun eigen device naar voorkeur. Met waarschijnlijkheid is dit het device waar deelnemers het meest gebruik van maken, wat in het voordeel heeft kunnen werken, door gewenning aan het device en scherm.

Zowel de gecreëerde foto's als de foto's van de Chicago Face Database (Ma et al., 2021) zijn gestandaardiseerd; voor alle foto's is er gekozen voor een gelijke achtergrond, een gelijke vorm en kleur shirts en geen sieraden. Hiervoor is gekozen om zo de focus echt alleen op het gezicht te kunnen leggen en alle overige contextuele aanwijzingen te verwijderen. Door alle focus alleen op de gezichten zelf te leggen wordt de interne validiteit versterkt. Om deze reden zouden AI-gegenereerde foto's niet herkend kunnen

worden door o.a. aanwijzingen van een achtergrond of een schaduw die net verkeerd is. De focus in dit onderzoek is daardoor gelegd op het kunnen onderscheiden op basis van echt alleen het gezicht, er is gekozen voor een hogere interne validiteit ten nadele van de externe validiteit (d.w.z. een natuurlijke omgeving). Hierop aansluitend zou een suggestie voor toekomstig onderzoek kunnen zijn om dit onderzoek uit te voeren met meer alledaagse foto's. Zo wordt de generalisatie hoger en zijn de resultaten uit te breiden naar andere settingen.

Een andere suggestie voor toekomstig onderzoek is het kijken naar de effecten van inductief leren over lange termijn in plaats van alleen korte termijn, zoals uitgevoerd in dit onderzoek. Dit onderzoek heeft een training van 96 foto's met daarachter gelijk de test, waardoor het resultaat minder zegt over hoe het effect zich verloopt en of het een blijvend effect is. De korte termijneffecten kunnen beïnvloed worden door situatie gebonden effecten (d.w.z. door aandacht of leer- en oefen-effecten). Voor een sterke conclusie zou het een suggestie zijn in toekomstig onderzoek om ook te kijken naar hoe het effect zich ontwikkelt op de lange termijn. Blijft het effect bestaan of werkt dit alleen op de korte termijn? Door dit effect te onderzoeken kan worden vastgesteld of het effect ook generaliseerbaar is naar andere settingen en over een langere periode van tijd.

Daarnaast vertellen de resultaten weinig over de relatie tussen leeftijd, opleidingsniveau en ervaring met AI. Interessant zou zijn om te kijken of de resultaten verschillen in ouderen, die over het algemeen relatief weinig ervaring/kennis met AI hebben, in vergelijking tot jongeren. Een suggestie zou daardoor zijn; het meenemen van leeftijdseffecten, opleidingsniveau en ervaring/kennis over AI. Ook kan er in de toekomst gekeken worden naar langetermijneffecten en het onderscheiden in video's.

Concluderend kunnen we stellen dat inductief leren werkt in de context van het onderscheiden tussen AI-gegenereerde foto's en echte foto's. Dit wil zeggen dat training helpt en belangrijk kan zijn in de toekomst. Met de snelle ontwikkelingen breidt AI zich razendsnel uit en blijft het daardoor van belang om educatie en training te geven om zo misleiding en oplichting te kunnen voorkomen. Nadat mensen hebben ervaren hoe complex de AI-foto's gegenereerd kunnen worden, daalt het vertrouwen in het kunnen onderscheiden flink. Dit staat in een positieve lijn ten opzichte van voorzichtigheid, waardoor mensen minder snel een slachtoffer van misleiding en oplichting zullen worden. Suggesties voor toekomstig onderzoek zijn het meenemen van leeftijd, opleidingsniveau, ervaring en training op lange termijn, om de generaliseerbaarheid te vergroten.

Referenties

Bray, S. D., Johnson, S. D., & Kleinberg, B. (2023). Testing human ability to detect 'deepfake' images of human faces. *Journal of Cybersecurity, 9*(1). https://doi.org/10.1093/cybsec/tyad011

Burson, K. A., Larrick, R. P., & Klayman, J. (2006b). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal Of Personality And Social Psychology*, *90*(1), 60–77. https://doi.org/10.1037/0022-3514.90.1.60

Chandler, C. C. (1994). Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition test. *Memory & Cognition, 22*(3), 273–280. https://doi.org/10.3758/bf03200854

Cooke, D., Edwards, A., Barkoff, S., & Kelly, K. (2025). As good as a coin toss: Human detection of AI-generated content. *Communications of the ACM*. https://doi.org/10.1145/3729417

DeepL SE. (n.d.). *DeepL Translator* [Software]. https://www.deepl.com/translator

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Field, A. (2018). *Discovering statistics using IBM SPSS Statistics*. SAGE Publications Limited.

Högemann, M., Betke, J., & Thomas, O. (2025). What you see is not what you get anymore: a mixed-methods approach on human perception of AI-generated images. *Frontiers in Artificial Intelligence*, *8*, 1707336. https://doi.org/10.3389/frai.2025.1707336

Kang, S. H. K., & Pashler, H. (2011). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26*(1), 97–103. https://doi.org/10.1002/acp.1801

Karnouskos, S. (2020). Artificial intelligence in digital media: The era of deepfakes. *IEEE Transactions on Technology and Society, 1*(3), 138–147. https://doi.org/10.1109/tts.2020.3001312

Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience, 24*(11), 103364. https://doi.org/10.1016/j.isci.2021.103364

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories. *Psychological Science, 19*(6), 585–592. https://doi.org/10.1111/j.1467-9280.2008.02127.x

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Lindsey, B. A., & Nagel, M. L. (2015). Do students know what they know? Exploring the accuracy of students' self-assessments. *Physical Review Special Topics - Physics Education Research*, *11*(2). https://doi.org/10.1103/physrevstper.11.020103

Lyons, B. A., Montgomery, J. M., Guess, A. M., Nyhan, B., & Reifler, J. (2021). Overconfidence in news judgments is associated with false news susceptibility. *Proceedings of the National Academy of Sciences, 118*(23). https://doi.org/10.1073/pnas.2019527118

Ma, D. S., Kantner, J., & Wittenbrink, B. (2021). Chicago Face Database: Multiracial expansion. *Behavior Research Methods, 53*(3), 1289–1300. https://doi.org/10.3758/s13428-020-01482-5

Midjourney. (2025). *Midjourney v7* [AI image-generation model]. https://www.midjourney.com

Miller, E. J., Steward, B. A., Witkower, Z., Sutherland, C. A. M., Krumhuber, E. G., & Dawel, A. (2023). AI hyperrealism: Why AI faces are perceived as more real than human ones. *Psychological Science, 34*(12), 1390–1403. https://doi.org/10.1177/09567976231207095

OpenAI. (2025). *DALL·E 3* [AI image-generation model]. https://openai.com/dall-e

Somoray, K., & Miller, D. J. (2023b). Providing detection strategies to improve human detection of deepfakes: An experimental study. *Computers in Human Behavior, 149*, 107917. https://doi.org/10.1016/j.chb.2023.107917

Wat is AI en wat zijn de risico's? (2024, 18 november). https://veiliginternetten.nl/wat-is-ai-en-wat-zijn-de-risicos/

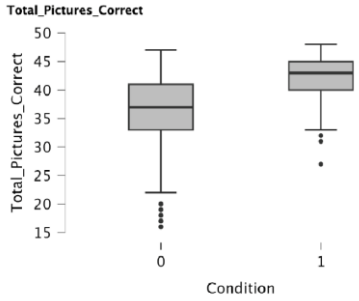
Tabellen en figuren

Tabel 2

Descriptive Statistics

	Total_Pictures_Correct	
	0	1
Valid	148	122
Missing	0	0
Mean	35.703	41.975
Std. Deviation	7.038	4.273
Minimum	16.000	27.000
Maximum	47.000	48.000

Figuur 1



Tabel 3

Independent Samples T-Test

	t	df	p	Cohen's d	SE Cohen"s d
Total_Pictures_Correct	-9.014	248	< .001	-1.077	0.137

Note. Welch's t-test

Tabel 4

Descriptives

	N	Mean	SD	SE	Coefficient of variation
Confidence_pre	117	3.641	0.987	0.091	0.271
Confidence_After_Training	122	3.770	1.191	0.108	0.316
Conf_control	144	3.347	1.060	0.088	0.317
Conf_3	270	3.196	1.274	0.078	0.399

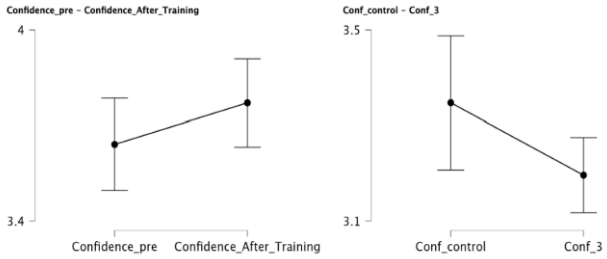
Tabel 5

Paired Samples T-Test

Measure 1		Measure 2	t	df	p	Cohen's d	SE Cohen's d
Confidence_pre	-	Conf_After_Training	-1.408	115	0.162	-0.131	0.095
Conf_control	-	Conf_after Test	6.005	143	< .001	0.500	0.097

Note. Student's t-test.

Figuur 2



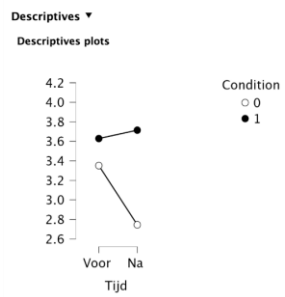
Tabel 6

Within Subjects Effects

Cases	Sum of squares	df	Mean square	F	p
Tijd	8.271	1	8.271	11.897	< .001
Tijd + Condition	14.607	1	14.607	21.011	< .001
Residuals	172.411	248	0.695		

Note. Type III Sum of squares

Figuur 3



Tabel 7

Descriptive Statistics

	Total_Pictures_Correct		
	Laptop	Phone	Tablet
Valid	203	64	3
Missing	0	0	0
Mean	39.034	37.281	31.667
Std. Deviation	6.485	7.095	9.018
Minimum	16.000	17.000	23.000
Maximum	48.000	47.000	41.000

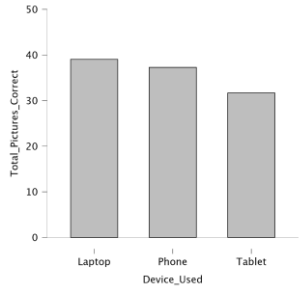
Tabel 8

ANOVA Total_Pictures_Correct

Cases	Sum of squares	df	Mean square	F	p
Device_Used	292.767	2	146.383	3.304	0.038
Residuals	11828.363	267	44.301		

Note. Type III Sum of squares

Figuur 4



Bijlage

Prompts voor het creëren van AI-foto's

In de gebruikte modellen, Midjourney versie 7 (Midjourney, 2025) en Dall-e versie 3 (OpenAI, 2025), zijn Engelse prompts gebruikt om de foto's te creëren. Deze staan hieronder uitgeschreven, met een aantal voorbeelden van de AI-gegenereerde foto's die de programma's gaven.

Females

Angry (Attractive)

Make a realistic photo of a 35 year old white woman. She has an angry expression, wearing a grey t-shirt (round neck) against a pure white background. Natural skin texture, minimal makeup, unstyled hair, and realistic lighting.



Angry (Unattractive)

Make a frontal photograph of an ugly-looking black woman (45 years old) with an irritated expression, wearing a grey t-shirt (round neck) against a white background. Uneven skin texture, tired eyes, unstyled long hair, no makeup, realistic lighting and colors, round, chubby face.

Happy (Attractive)

Make a front-facing photograph of a middle aged attractive black woman with a happy expression (closed mouth), wearing a grey t-shirt (round neck) against a white background. Minimal make-up, no jewelry and realistic lighting.



Happy (Unattractive)

Create a frontal photograph of an ugly-looking young black woman with a happy expression (open mouth), wearing a grey t-shirt with a round neck against a completely white background. Her face must be round and full, she should not be wearing make up and her teeth should be a little crooked.

Fear (Attractive)

Create a picture of an attractive middle-aged white woman which could be used for a database for emotions. The picture should have a white background and she should be wearing a grey shirt with a round neck. She should look scared, but realistically. She should have her hair up and have no jewelry. Her shirt and shoulders should be visible.

Fear (Unattractive)

Make a picture of an unattractive adolescent white woman with acne (23 years old) who is considered ugly exhibiting the emotion fear, she has to be terrified. Completely white background. Gray shirt (round neck). Should not be too close up (shirt and shoulders visible). Realistic human with facial asymmetry, and add some blemishes in the skin. Not too pretty looking. Not too much light and reduce shininess on the face.



Males

Angry (Attractive)

Make a realistic photo of a 40 year old white man. He has an angry expression, wearing a grey t-shirt with a round neck against a pure white background. Natural skin texture and realistic lighting.

Angry (Unattractive)



Create a frontal photograph of an unattractive adolescent white man with an angry/irritated expression, wearing a grey t-shirt with a round neck against a completely white background. He is overweight and has skin blemishes. Shirt and shoulders should be visible.

Happy (Attractive)



Make a front-facing picture of an attractive white man (25 years old) who is smiling, exhibiting the emotion of being happy. Emotion recognition task theory of mind database. Grey shirt with a white neck and pure white background. Not too zoomed in, his shirt and shoulders should be visible. No additional shine on the face.

Happy (Unattractive)

Produce an image of an ugly middle-aged white man who is happy (smiling with mouth open). He should have facial asymmetry, eye bags, skin blemishes, a grey shirt with a round neck and a fringe. The background needs to be white. Should be indistinguishable from reality.

Fear (Attractive)

Image of an attractive black man (30 years old) looking scared against a white background. The picture should be usable in a database for emotion recognition. Gray shirt with a rounded neck. Shoulders should be visible. Realistic lightning and reduce shininess on the face.

Fear (Unattractive)

Make a front facing image of a black man (50 years old) who is considered ugly and unattractive. He should be looking frightened. Give him facial asymmetry. He should be

wearing a grey shirt with a round neck and his shoulders should be in the picture. Do not add any jewelry or additional facial shininess.

