universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Your world, Your ads: How knowing about deepfake personalization shapes consumer purchase intentions Deepfakes in Advertising

verfasst von | submitted by

Octavia Ronja Maria Lazarov B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 915
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Betriebswirtschaft
Betreut von | Supervisor:	Univ.-Prof. Dr. Martin Eisend M.A.

Abstract (English)

Background: Deepfake advertising has emerged as a new AI technology enabling hyper-personalization. Personalized AI ads have been shown to increase persuasion outcomes, but prior research shows that disclosures of deepfake advertising can reduce these effects. While disclosure gives clear insights about the nature of the ad, it has not yet been thoroughly examined how consumer knowledge about the technology, awareness of deepfakes, affects purchase intention. Disclosure and consumer knowledge operate through different mechanisms: disclosures provide explicit cues, whereas prior knowledge may trigger internal evaluations, especially when the presence of a deepfake is uncertain.

Purpose: This thesis therefore investigates (1) how knowledge about deepfakes in advertising influences purchase intention, and (2) whether deepfake-powered personalized ads are effective.

Methods: Using the framework proposed by Campbell et al. (2021) we built hypotheses and used a quantitative experimental design. The design included a 2x2 factorial design with personalization and awareness of deepfakes as the two manipulated variables. In an online survey, 133 participants viewed deepfake ads that were personalized to their stated preferences.

Results: (1) Findings show that increased awareness of deepfakes reduces the realistic visual aspect and consistency of the ad (portrayal of realism). This reduction in portrayal of realism subsequently increased awareness of falsity, which in turn negatively influenced purchase intention. (2) Further, results showed that deepfake-based hyper-personalization is ineffective for unknown brands.

Conclusion: This work contributes to the literature in several ways: the framework from Campbell et al. (2021) has been extended by incorporating awareness of deepfakes as an additional antecedent variable. This work suggested to separate the variable perceived realism within this framework into two new variables: perceived likelihood of realism and portrayal of realism. Then results suggest that (1) with increasing knowledge about deepfakes the technical quality of a deepfake directly shapes purchase intention. Notably, awareness of deepfakes selectively impacted only portrayal of realism, leaving perceived likelihood of realism unaffected. (2) This work identified that hyper-personalization offers little value for unfamiliar brands. Future research should examine whether these effects differ when brand familiarity is high.

Abstract[1] (Deutsch)

Hintergrund: Deepfake-Werbung hat sich als neue KI-Technologie etabliert, die eine Hyper-Personalisierung ermöglicht. Es hat sich gezeigt, dass personalisierte KI-Werbung die Überzeugungskraft steigern; frühere Untersuchungen legen jedoch nahe, dass die Offenlegung von Deepfake-Werbung diese Effekte abschwächen kann. Während eine Offenlegung klare Einblicke in die Natur der Anzeige gewährt, wurde bislang noch nicht gründlich untersucht, wie sich das Wissen der Verbraucher über die Technologie und ihr Bewusstsein für Deepfakes auf die Kaufentscheidung auswirken. Offenlegung und Verbraucherwissen wirken über unterschiedliche Mechanismen: Offenlegungen liefern explizite Hinweise, während Vorwissen interne Bewertungen auslösen kann, insbesondere wenn das Vorhandensein eines Deepfakes ungewiss ist.

Zweck: Diese Arbeit untersucht daher (1) wie das Wissen über Deepfakes in der Werbung die Kaufabsicht beeinflusst und (2) ob personalisierte Deepfake-Anzeigen wirksam sind.

Methoden: Unter Verwendung des von Campbell et al. (2021) vorgeschlagenen Model haben wir Hypothesen aufgestellt und ein quantitatives experimentelles Design verwendet. Das Design umfasste ein 2x2-Faktordesign mit Personalisierung und Bewusstsein für Deepfakes als den beiden manipulierten Variablen. In einer Online-Umfrage sahen sich 133 Teilnehmer Deepfake-Anzeigen an, die auf ihre angegebenen Präferenzen zugeschnitten waren.

Ergebnisse: (1) Die Ergebnisse zeigen, dass ein gesteigertes Bewusstsein für Deepfakes den realistischen visuellen Aspekt und die Konsistenz der Werbung (Realismusdarstellung) verringert. Diese Verringerung der Realismusdarstellung führte in der Folge zu einem gesteigerten Bewusstsein für die Fälschung, was wiederum die Kaufabsicht negativ beeinflusste. (2) Darüber hinaus zeigten die Ergebnisse, dass Deepfake-basierte Hyperpersonalisierung bei unbekannten Marken unwirksam ist.

Fazit: Diese Arbeit leistet in mehrfacher Hinsicht einen Beitrag zur Literatur: Das Rahmenkonzept von Campbell et al. (2021) wurde durch die Einbeziehung des Bewusstseins für Deepfakes als zusätzliche Vorläufervariable erweitert. Diese Arbeit schlug vor, die Variable „wahrgenommener Realismus“ innerhalb dieses Rahmenkonzepts in zwei neue Variablen zu unterteilen: „wahrgenommene Wahrscheinlichkeit von Realismus“ und „Darstellung von Realismus“. Die Ergebnisse legen nahe, dass (1) mit zunehmendem Wissen über Deepfakes die technische Qualität eines Deepfakes die Kaufabsicht direkt beeinflusst. Bemerkenswert ist, dass sich das Bewusstsein für Deepfakes selektiv nur auf die Darstellung von Realismus auswirkte, während die wahrgenommene Wahrscheinlichkeit von Realismus davon unberührt blieb. (2) Diese Arbeit hat gezeigt, dass Hyper-Personalisierung für unbekannte Marken nur einen geringen Mehrwert bietet.

[1] *Übersetzt mit DeepL.com (kostenlose Version)*

Zukünftige Forschung sollte untersuchen, ob sich diese Effekte unterscheiden, wenn die Markenbekanntheit hoch ist.

Table of Contents

List of Figures

Figures - Main Part

Figures - Appendix

List of Tables

Tables - Main Part

Tables - Appendix

List of Abbreviations

AI	Artificial Intelligence
AD	Awareness of deepfake
AF	Awareness of falsity
PR	Perceived Realism
PI	Purchase intention
PoR	Portrayal of realism
PLR	Perceived likelihood of realism
P	Personalization
CE	Celebrity endorsement
BA	Brand attitude
CV	Control variable
DV	Dependent variable
IV	Independent variable
H	Hypothesis

I. Introduction: Deepfakes as an emerging opportunity in advertising

Imagine watching an advertisement with your favourite actor, Bruce Willis. He appears to promote a mobile phone operator while trying to defuse a bomb. Now imagine learning that Bruce Willis never took part in the filming. The person you saw was actually an unknown Russian actor. The advertisers replaced his face with Bruce Willis's using digital technology. How would you feel? Deceived?

This ad really existed in 2021, a company called Deepcake created an advertisement for the Russian mobile phone operator Megafon, it featured Bruce Willis, even though he never appeared on set (Wilson, 2022). The company used a technology called deepfake.

1.1 Problem statement

Authors claim that Deepfakes have the potential to disrupt the marketing sector (Campbell et al., 2022; Kietzmann et al., 2021). One reason is deepfakes can significantly decrease costs, because advertisers do not need to hire models or celebrities (Kietzmann et al., 2021). Another reason is: The ability to create highly personalized advertisements. and enable hyper-personalization, meaning one-to-one advertising (Davenport, 2023). Campbell et al., 2021 (2022) suggests that advertisers could tailor ads to a customer's ethical background or body shape, but he also draws attention to the problems when using deepfakes to achieve hyper-personalization. He formulates the question on how the use of deepfake powered personalized ads impact the advertisements efficiency (Campbell et al., 2022). The author also draws attention to the problems when the users awareness about this use of deepfakes in ads increases, which might result in increase of privacy concerns (Campbell et al., 2021). Potential negative reactions when knowing that the ad is created using deepfakes brings us to ask the question on: how awareness about deepfakes and their use in advertisements affect the consumer purchase intention?

It is important for marketers to understand how consumers respond to hyper-personalized deepfake ads once they know they exist. Eventually, consumers will become aware of their use (Kharvi, 2024). If reactions are neutral or positive, marketers could use deepfakes in advertisements and thus benefit in two ways: reducing production costs and achieving hyper-personalization.

Some studies have looked at consumer reactions to deepfake advertising (Arachchi & Samarasinghe, 2024; Sivathanu & Pillai, 2022), and other research has studied the effects of how disclosure of deepfakes impacted purchase intentions (Karpinska-Krakowiak & Eisend, 2024). But very few have examined how awareness of deepfakes [2] affects purchase intention. Further studies on hyper-personalized deepfake ads are even rarer. Because of this, it is important to ask whether using deepfakes

[2] It refers to the knowledge people have generally about the technology (see appendix 12 for definition) while the disclosure is a direct indication in the advertisement content that it has been produced using AI.

for hyper-personalization has a positive effect on purchase intention. Thus this work tries to answer the following question:

Research Question:

How does consumer awareness of deepfake-powered hyper-personalization influence purchase intention?

Study Objective:

This study examines whether consumer awareness of personalized deepfakes affects their willingness to purchase a product. To explore this, an adapted form of the Consumer Response to Manipulated Advertising framework from Campbell et al. (2021) has been applied and tested it through an experiment.

Motivation:

Deepfake is a new disruptive technology with interesting advantages for advertisers: cost reduction and hyper personalization. But the technology is also presenting risks and questions on how consumers could react. Knowledge that consumers have about deepfakes and their use might impact the advertisements efficiency. In order to well use deepfakes for producing advertising content it is important for marketers to know what factors can impact the efficiency of the advertisement.

1.2 Scope of this study

In the following we will take a closer look at the scope of this study and bring in a short explication on what deepfakes actually are.

1.2.1 What are deepfakes? Technical background

The term *deepfake* combines "deep" and "fake." The word "deep" refers to the artificial intelligence technology involved, while "fake" denotes the alteration or synthesis of media by merging real and artificial elements (Nguyen et al., 2022). Like in the Megafon advertisement, where the movements of a real actor were combined with the digitally added face of Bruce Willis.

a. Technological aspect

The first part of the Term "Deep" refers to the technological side of deepfakes, specifically deep learning. Deepfakes make the use of Machine learning techniques, deep learning, which uses highly complex neural networks: Neural networks (defined in Appendix 12) are a type of AI systems inspired

by the human brain, but deepfakes rather use deep learning models, which have at least three layers, but often hundreds or thousands, which makes them much more computationally intensive than traditional models but also makes them more efficient (Holdsworth & Scapicchio, 2024). Deepfakes can be created using two main types of machine learning models: the autoencoder and the generative adversarial network (Seow et al., 2022).

Autoencoder

The autoencoder is a model frequently applied in the creation of deepfakes (Campbell et al., 2021; Kietzmann et al., 2021)**.** More broadly, autoencoders are used to remove noise from images or to detect irregularities in data, for instance in fraud detection (Seow et al., 2022) An autoencoder consists of two neural networks: an encoder and a decoder (Kietzmann et al., 2021). In facial processing, the encoder compresses the features of a face into approximately 300 core characteristics. This reduction is called the bottleneck. The decoder reconstructs the original face from these 300 characteristics. Noise is removed because only the essential features are transmitted through the bottleneck; the reconstruction excludes irrelevant visual information. If two autoencoders are trained on two different faces, their decoders can be exchanged. The encoder for face 1 records its distinctive movements and features, such as head orientation. Then these encoded features are passed to the decoder but the one trained on face 2, thus the output is a reconstruction of face 2 performing the movements of face 1. (Seow et al., 2022) This mechanism is an example for the face-swapping capability of deepfakes, however deepfakes can also do puppet mastery, moving the body, or voice imitations (Kietzmann et al., 2021; Seow et al., 2022).

Generative Adversarial Networks (GANs)

The second technology used is Generative Adversarial Networks (GANs). Some authors treat GANs as distinct from deepfakes (Kietzmann et al., 2021). While others note that many applications offering pretrained deepfake models, such as Zao, are based on GANs, also numerous academic research on deepfakes relies on models built with GANs, so Seow et al. (2022). GANs can be applied both to create and to detect deepfakes. The model architecture consists of two neural networks. The first network, the generator, attempts to create synthetic content, such as an artificial image of a face. The second network, the discriminator, is trained to distinguish between real and generated content. These networks are trained in competition with one another. When the discriminator correctly tags the content as fake, then the generator adjusts its model parameters to improve the content generated realism. On the other side, when the discriminator fails to detect a fake, it adapts its own parameters to improve the detection. This adversarial process continues until the generated content becomes increasingly difficult to distinguish from genuine data. (Seow et al., 2022)

b. *Synthetic aspect*

The “fake” component refers to creating convincing synthetic media from existing material. Initially, this term was used to describe face-swapping in videos or photographs (Kleine, 2022) just like seen in the first example with Bruce Willis. Kietzmann et al. (2021) later proposed a classification of deepfakes into photo, video, audio, and combined audio-video categories. The technology is not restricted to only switch faces in a video, the author adds also the use of Lip-sync , Puppet mastery (faking the whole body movements) and voice imitating to the potential uses of Deepfakes. Seow et al., (2022) expand these possibilities further and describe techniques for generating entirely new faces or merging two faces to create a new one. They also include face attribute manipulation, such as altering the shape of a person’s nose.

1.2.2 Deepfakes in advertising

Deepfake research began in 2017 and gained increasing attention across fields like computer science, arts, and the humanities (Kleine, 2022). Kharvi (2024) highlighted the ability of deepfakes producing convincingly fake speeches of politicians and spread misinformation, an especially concerning issue during elections. It is important to understand that deepfakes were initially associated with negative meanings and thus using them in advertising could be risky for advertisers. Since consumers might transfer the negative image of deepfakes to the brand or product being advertised. Looking at these concerns, some researchers proposed protection measures. For example, Kietzmann et al. (2020) suggested a legal framework to regulate deepfakes. This could involve mandatory disclosures and technologies that detect fake content. Other authors suggested to improve media literacy, which could help consumers and companies defend themselves against fake news generated by deepfakes (Mustak et al., 2023).

But we will rather focus on the context of advertising, where questions arise concerning consumers reaction on Deepfakes, disclosures of Deepfakes and also what long terms effects the use of deepfakes have on consumers trust and scepticism (Campbell et al., 2021; Kietzmann et al., 2021; Whittaker et al., 2021). While the analysis of a legal framework is important too, this work will rather focus on the aspect of media literacy. One of these legal implications is the use of disclosure, which has already been much researched. This study rather aims to examine the yet not well researched aspect on whether consumers’ knowledge about the use of deepfakes in advertising influences their purchase decisions. Reasons are that legal restrictions vary from country to country, some disclosure might be too small for the consumer to recognize or see. As seen by Karpinska-Krakowiak and Eisend (2024) simple disclosures are not as efficient as extended disclosures for deepfake advertisements. In this work this knowledge will be referred as “awareness of deepfakes”, which refers to the overall knowledge consumer have about deepfakes and their use in advertisement, including their use to achieve hyper- personalization. The

concept awareness of deepfake is only covering a part of media literacy, however due to restricted capacities this study will not be able to implement a whole media literacy intervention, like Austin et al. (2007); Austin and Johnson (1997) and Scull et al. (2014) were able to do.

Further deepfakes also have the potential to open up for new creative uses (Campbell et al., 2021), even though this is an interesting opportunity for marketers this work will not analyse the creative aspect but rather focus on the perceptual side. Including creativity would make the analysis too complex, so it will not be addressed further.

Thus this work will investigate how awareness of deepfakes affects consumers' purchase intentions when viewing a deepfake advertisement. Secondly the interaction between personalization of the advertisement content and awareness of deepfakes on the resulting advertisements purchase intention will be examined.
In order to investigate these questions, an online experiment has been conducted, in which consumers will be faced with a fictive advertising situation. Their level of awareness about deepfakes has been manipulated and they were exposed to either personalized or non-personalized deepfake advertisements. Before outlining the research method in detail, we will review existing literature about deepfakes in advertising. Identifying key use cases for marketers and examining some examples of personalized deepfake ads already in use. We will then present the framework proposed by Campbell et al. (2021), which we adapted for this study to explain consumer behaviour when confronted with deepfakes ads.
Finally, we will describe the experimental design and present the results. Last but not least we will discuss the experimental results and give some suggestions for marketers.

II. Literature review and identified gap

In the context of Deepfakes the amount of studies from social sciences, especially from political sciences, journalism, and ethics where much more prevalent than from other areas (Vasist & Krishnan, 2022; Whittaker et al., 2023). According to Kharvi (2024) the social aspect and the danger that deepfakes brings with them were often topic in research. As for many new technologies, the advantages Deepfakes could contribute to advertisers were not immediately identified**.** Even though past research mostly focused on harmful use case of this technology, more and more studies see the ambivalence of this technology, recognizing both the risks and opportunities of deepfake technology (Vasist & Krishnan, 2022).

In marketing research, researchers found promising use cases. Several academic papers discuss how deepfakes could be used in content creation and editing (Arachchi & Samarasinghe, 2024). These studies also examine the pros and cons of deepfakes for both consumers and businesses (Campbell et al., 2022; Kietzmann et al., 2021; Kwok & Koh, 2021; Mustak et al., 2023; Whittaker et al., 2021).

2.1 The ambivalence of deepfakes in advertising

The ambivalent aspect of deepfakes in marketing has been well described by Whittaker et al. (2021)**.** They describe deepfakes as offering new opportunities for marketers but also on bringing serious risks, especially concerning the spread of fake news. The authors point out that deepfakes can appear very realistic and thus they are very effective in convincing consumers. This creates threats for companies, for example malicious actors could produce deepfakes of a company's spokesperson making harmful statements or competitors could create fake product reviews that highlight negative product features (Whittaker et al., 2021)**.** These could risk destroying the companies brand image and mislead consumer decision making (Mustak et al., 2023)**.**

It becomes harder for consumers to tell whether media content is real or fake. Decision-making becomes more difficult, especially because many consumers still lack the media literacy needed to recognize fake deepfake generated content (Mustak et al., 2023). Knowing that deepfakes are used for harmful purposes in advertising, can increase consumer scepticism (Kietzmann et al., 2021) and people may become more cautious and critical when viewing ads. Consumers might start looking for signs in the video, voice or photo to identify deepfakes, or further assume that ads are using deepfakes, even if the content is actual real. This could reduce the effectiveness of advertising.

However, researchers have also identified several positive effect that the new technology has in marketing and advertising. For example in advertising, an important advantage is the reduction of production costs. Deepfakes allow advertisers to reduce the working hours of actors, makeup artists, and designers, advertisers can also reuse existing video footage in multiple advertising campaigns

(Kietzmann et al., 2021). Then companies can replace the face of an unknown actor with that of a celebrity, this removes the need for the celebrity to be physically present and lowers travel costs (Kwok & Koh, 2021). Deepfakes can also show a celebrity doing things they would not normally do, such as dancing or performing stunts (Chitrakorn, 2022).

Further the realism of deepfakes makes them especially useful in advertising. Realistic images are often more appealing than animations or illustrations (Kim et al., 2019; Whittaker et al., 2021). This technology also allows for the creation of highly realistic virtual brand ambassadors much more realistic than traditional 3D models, advertisers also gain more control over the ambassador than when involving celebrities or models, which can have uncontrollable behaviour (Mustak et al., 2023).

Since deepfakes definitely can bring interesting benefits to advertisers, we should closely examine how consumers respond to deepfake ads. If the response is positive, advertisers can fully benefit from lower costs and greater control over brand representation. Some researchers have already examined consumers response to deepfake advertisements and have shown that using deepfakes in advertising can positively influence purchase intention (Sivathanu & Pillai, 2022). Arachchi and Samarasinghe (2024) found that the use of deepfakes in advertisements can positively impact brand credibility, because the new technology increases interest and relevance of the ad. The authors showed that the use of deepfakes can increase customer interest, because of the innovative nature of Deepfakes. Deepfake videos can attract attention and offer new creative ways to do advertising (Campbell et al., 2021). One creative uses of deepfakes include letting consumers swap their faces with a celebrity performing a dance, which shifts the role of the viewer from a passive audience member to an active co-creator (Kietzmann et al., 2021). Companies can also use deepfakes to recreate historical figures (Kwok & Koh, 2021), Edeka for example used the deepfake of Einstein to advertise their new soft drinks (EDEKA, 2025), making the ad highly interesting and eye catching.

Further Arachchi and Samarasinghe (2024) found that deepfakes make the advertisement feel more relevant. For instance, marketers can overcome language barriers by dubbing videos using a deepfake of the person's voice and matching their lip movements (Mustak et al., 2023). This approach has already been used in 2019, a malaria awareness campaign featured David Beckham speaking nine different languages (Zero Malaria Britain, 2019). This example is good to show how the use of deepfakes increases the relevance for the viewers, by speaking the viewers language the viewer will feel more personally addressed. Therefore advertisement relevance can be increased with deepfakes thanks to the ability of deepfakes to achieve high levels of ad personalization. This is one of the aspects we want to take a closer look in this work.

2.2 Advertising personalization as an opportunity enabled by deepfakes

2.2.1 Personalization definition

Davenport (2023) explains that personalization originates from the concept of market segmentation, in which consumers are grouped according to sociographic or geographic criteria. He explains that advertising personalization is a concept in which the content and message of an advertisement is adapted to the individual characteristics of the consumer. This includes the use of the consumer's name or tailoring the ad to their preferences (Baek & Morimoto, 2012). The main goal is to better match the ad to the consumer's specific needs, thus companies collect and analyse data such as demographics, psychographics, purchase history, location, lifestyle, or interests in order to then be able to personalize the ad to the specific customer groups (Baek & Morimoto, 2012; Semeradova & Weinlich, 2019).

There are various applications of personalization, such as product recommendations on platforms like Amazon, personalized content suggestions on services like Netflix, personalisation of product attributes or services, and finally personalized advertisement fitting the interests and needs of the consumer (Davenport, 2023).

2.2.2 Hyper-personalization definition

Personalization can exist on different levels. White et al. (2007) describe the different degrees of personalization with the example of an email: A low degree of personalization only includes the recipient's name, but a higher degree of personalization would also include their profession and location.

A very high degree of personalization would be hyper-personalization. Davenport (2023, p. 30) refers to this as the "true personalization". According to him, hyper-personalization means creating one-to-one marketing and is made possible through machine learning systems analysing large datasets to detect patterns and make buyer predictions. However the author only looks at one aspect of hyper-personalization: the need to identify the preferences, needs and characteristics of the consumer. He does not explore how the actual ad content can be adapted to better match the consumer, such as skin tone, height, or gender of a model. Without Deepfakes and AI it is difficult to achieve true personalized ad content, since creating personalized ads for each individual using traditional tools can be extremely expensive and time-consuming (Kietzmann et al., 2021). This is where generative AI and deepfakes become important. Deepfakes can quickly modify an existing video. Thus the emergence of deepfakes make it possible for companies to even consider hyper-personalization as a possible advertising strategy.

The fashion industry is great example to picture the use of deepfakes for hyper-personalized ads. The fashion industry could easily use customer data collected from social media to create tailored deepfake ads to match the consumer's ethnicity, body shape, age, hobbies, or gender (Campbell et al., 2021;

Kietzmann et al., 2021). Another use case could be to enable consumers to create their own avatars or switch their faces with those of fashion models to see how outfits might look on them. Whittaker et al. (2021) explains that this can boost customer engagement.

Even though the use of deepfakes to achieve hyper-personalization may seem futuristic, some companies have already been able to test it. One example comes from fashion retailer Zalando. According to Chitrakorn (2022) Zalando used deepfakes to create a Facebook campaign featuring Cara Delevingne. A version of the campaign video can be find on YouTube (Wonderlandmovies, 2018). She was made to say the names of cities where Zalando services were available. According to Chitrakorn (2022), in total 290 000 different ads were created and the campaign run with great success through 12 different countries. Hearing the name of their own city made the ad more relevant to the consumer, which helped create a stronger connection with the audience.

However, this technology is still novel. Many consumers therefore may not know that the ad was tailored to them using deepfakes. When Cara Delevingne mentioned the name of their city in the Zalando ad. Did the viewers believe it was just a coincidence? What would have been the consumers reaction when they knew that Cara Delevingne never said their cities name but it was a trick used to capture their attention using deepfake? Even if they don't know exactly that Zalando used a Deepfake. What if the consumer is just generally aware of the existence of Deepfakes and becomes suspicious of the ad just by the model mentioning the small cities name? These questions raise concerns and uncertainties. In the next section, we will examine what researchers have already found in response to these issues.

2.2.3 AI literacy, awareness of deepfakes definition

One core question addressed in this work is how consumers react when they become aware of the use of deepfakes in the advertisement. As stated in part 1.2, explaining the scope of this work we will not take a look at disclosure but rather on awareness of deepfake.

As outlined earlier, *awareness of deepfakes* in this study refers to the consumers' understanding of deepfake technology and its use in advertising. For this, we draw on the construct of *media literacy* as defined by Aufderheide (1993, p. 9), who describes it as *"the movement to expand notions of literacy to include the powerful post-print media that dominate our informational landscape, helping people understand, produce, and negotiate meanings in a culture made up of powerful images, words, and sounds"*. Media literacy thus refers to the ability to critically analyse, understand, and evaluate media.

According to (Aufderheide, 1993), in order to have media literacy, three aspects needs to be considered. The first one is the knowledge on how medias are produced, for example understanding what deepfakes are, how they are created, and what they are capable of. The second aspect he mentions, is being able to decode the aesthetic, recognize the codes and evaluate the messages conveyed. Third, understanding

how audiences interpret media, including their commercial or political implications. Aufderheide (1993) explains that a higher media literacy can be developed through the persons own experience, self-driven questioning, or teachings. Recent discussions on literacy have also introduced the concept of *AI literacy*. Long and Magerko (2020) explain that AI literacy is a group of competencies, which has at it's core media and digital literacy, understanding how to use computers, being able to criticise the technology and use it effectively and reflect on its societal implications. The Authors found that AI literacy includes five aspects: *"What is AI? What can AI do? How does AI work? How should AI be used? and How do people perceive AI?"*(Long & Magerko, 2020, p. 3) and further define seventeen competencies : For example understand how AIs are programmed, recognize AI, understand ethical concerns and risks related to AI and acknowledging the essential human role in designing, training, and guiding AI systems toward specific objectives. Awareness of deepfakes only covers the first two aspects of AI literacy: knowledge of what deepfakes are and what they can do, including, in this context, the potential for hyper-personalization.

2.3 Identified literature gap

2.3.1 Consumers reaction to deepfake advertisements

Several researchers who have studied deepfakes in advertising have published conceptual papers outlining, as we have already discussed, the risks and opportunities they see in this technology(Campbell et al., 2021, 2022; Kharvi, 2024; Kietzmann et al., 2020, 2021; Kwok & Koh, 2021; Mustak et al., 2023; Whittaker et al., 2021). More difficult to find are empirical studies (Arachchi & Samarasinghe, 2024; Cochran & Napshin, 2021; Sivathanu & Pillai, 2022) or experimental studies (Karpinska-Krakowiak & Eisend, 2024; Powers et al., 2023) examining how consumers actually react to deepfakes in advertising.

Arachchi and Samarasinghe (2024) found that deepfake content can be more interesting for consumers and can have more relevance than traditional produced advertisings. Cochran and Napshin (2021) took a look at how consumers perceive Deepfakes and their dangers, further they analysed if consumer think platforms are responsible in disclosing deepfakes.

2.3.2 The effect of awareness of deepfakes on consumer behaviour

Karpinska-Krakowiak und Eisend (2024) and Powers et al. (2023) have used experimental settings and analyzed how consumer react on disclosed deepfake advertisement. Disclosure reduces ad source credibility and thus reduces purchase intention (Powers et al., 2023). Powers et al. (2023) further found that only perceived trustworthiness, one of the three sub dimension of source credibility, is an important mediator, that is reduced with disclosure and increases purchase intention. However, rare has been those authors analysing the aspect of AI literacy or awareness of deepfakes on consumers reaction on deepfake ads. Sivathanu and Pillai (2022) was one of the few researchers that somehow included this aspect of awareness in their work. All participants were told that the video they were shown was a deepfake . In

that sense the authors asked about the participants perceived deception and found that perceived deception reduces purchase intentions. However as mentioned Sivathanu and Pillai (2022) did not use an experimental design to test how consumer react in simulated real life situation. AI Literacy and awareness of deepfakes have not been well analysed until now, this is where our work will bring new insights.

2.3.3 The effect of personalization on consumer behavior

Deepfakes can be an important technology for achieving hyper-personalization. Researchers often point out that personalization has a strong positive impact on consumer behaviour toward advertisements (Guo & Jiang, 2023). Personalization can be very beneficial for advertisers, since consumers will be more fixated to the personalized ad then generic ads (Bang & Wojdynski, 2016), it increases the ads relevance (Guo & Jiang, 2023) and usefulness (Bleier & Eisenbeiss, 2015), the ad is more relatable, hence consumers will develop higher emotional attachment (Shanahan et al., 2019). However, many researchers have also raised concerns about personalization, particularly regarding the question of when personalization becomes excessive (Davenport, 2023).

At present, the use of deepfakes to achieve hyper-personalization remains mostly conceptual. Only a few practical cases exist, and the level of personalization is still relatively limited, for example, Zalando's advertisement. The ideas discussed by several authors concern much higher degrees of personalization including body adaptation of the model to the consumers preferences, celebrity choices fitting perfectly the consumers preferences (Campbell et al., 2021; Kietzmann et al., 2021; Whittaker et al., 2021). Nevertheless, a key question raised by researchers such as Kietzmann et al. (021) is how consumers would react when faced with hyper-personalized advertisements based on deepfakes.

So far, there is literature available on generative AI (see definition appendix 12) and personalization. Guo & Jiang (2023) investigated how AI-generated personalized texts in Chinese online advertising for famous car brands (BMW and Mazda) affect consumers' click-through rates. They found that AI-personalized text can significantly increase click-through rates and also enhance emotional engagement compared to personalized ads written by professional copywriters.

Matz et al. (2024) tested how large language models such as ChatGPT can generate personalized advertising texts tailored to consumer personalities. The authors showed that AI was able to effectively adapt advertisement texts to consumers personalities with simple prompts, which increased persuasiveness. The study also found that ai-powered personalized ads did not always result in higher willingness to pay. For example, a personalized ad for a vacation in Rome performed significantly better than a generic version, but for Nike shoes, personalization did not lead to stronger results.

Kumar and Kapoor (2024) went further and tested AI-generated personalized advertising videos. They compared personalized AI video ads with both personalized traditionally produced videos and generic videos. Their results showed that AI-generated personalized videos performed significantly better, with click-through rates 6–9% higher than the two control groups, non-personalized and human produced personalized video content. As we can see quite some research was performed on generative AI and personalization, specific studies focusing on deepfake technology however are still rare.

Sivathanu and Pillai (2022) investigated personalization with deepfakes. They tested whether deepfake-based personalized advertisements positively influence consumers' online shopping intentions and confirmed this hypothesis. In their study, participants first learned what deepfakes are and then viewed deepfake-based ads. The study was rather empirical and respondents were asked general questions about whether the deepfakes they see on social media fit their interests and if they would consider purchasing the product. The study provides an initial indication that deepfakes in personalized advertising can be beneficial. Yet the lack of real life situation simulation limits insights into how consumers might react in actual situations. An experimental approach that investigates how consumers respond to hyper-personalized deepfake ads and how they act with or without knowledge about deepfakes could provide valuable additional insights.

2.3.4 Interaction between personalization and awareness of deepfakes

Last but not least the specific link between deepfake awareness and personalized or hyper-personalized advertising has rarely been researched. The more consumers use social media, the more they are likely to learn about deepfake technology (Cochran & Napshin, 2021).Thus when marketers use Deepfakes, consumers will have more and more knowledge on what deepfakes are, the positive effects from personalization might be mitigated.

Until now research has been scarce investigating this issue, and if marketers want to make use of this easy and cheap way of achieving hyper personalization it is important to investigated on how AI literate consumers react when confronted with deepfake-based hyper-personalized ads. Looking at current literature we can identify a gap: lack in analysing how awareness of deepfakes and personalization interacts in advertising. A gap that has also been recognized by another master student. Selent (2022) from Technische Hochschule Brandenburg, also highlighted the need to examine hyper-personalized deepfakes more closely and in her thesis, she conducted a small empirical study (N=22). She designed a setup in which the subjects were shown deepfake videos of people they knew personally (friends). Afterwards, she asked the participants about their attitudes toward these deepfake ads. The participants were aware that they were being shown personalized deepfakes of their friends. However, because of the sample size, statistical analysis was incomplete. Nevertheless, the experimental setup, while difficult to realize, demonstrates the importance of testing how consumers respond when confronted with hyper-

personalized deepfakes. Building on this idea, this work aims to fill this gap and design an experiment in which consumer with and without basic knowledge about what deepfakes are and how they are used are exposed to personalized deepfakes.

In the studies from Sivathanu and Pillai (2022) and Selent (2022), participants were informed that they were viewing deepfakes. Research has shown that disclosure can reduce purchase intention. AI disclosure in personalized advertisement also has been analysed by Matz et al. (2024). In their study, they tested two types of disclosure: In the first case, participants were only told that AI had been used to generate the text. The second disclosure was more detailed and they were informed that ChatGPT had been used to personalize the advertisement text. However, the authors did not find any significant differences. This might be because they only applied simple disclosure, and as Karpinska-Krakowiak and Eisend (2024) have shown, simple disclosure does not strongly affect consumers' perception of an ad.

This work however rather wants to focus on AI literacy, awareness about the technology than direct disclosure. Disclosure has already been studied extensively, it is interesting to also consider the role of AI literacy, since consumers increasingly encounter deepfakes, with it their knowledge and awareness of this technology will likely increase and shape their responses. Consumers may view ads differently once they gain more knowledge about deepfakes and their existence. Increasing consumer awareness is inevitable, whether through mandatory disclosure or through rising AI literacy over time. It remains an open question how this awareness affects the effectiveness of deepfake advertising. The goal of this work is to address this gap by investigating consumer reactions to hyper-personalized deepfake advertisements in situations where consumers are aware of the existence and potential use of deepfakes by advertisers.

III. Conceptual framework and hypothesis development

To investigate the identified research gap, we will develop hypothesis, using the framework by Campbell et al. (2021) as our conceptual foundation.

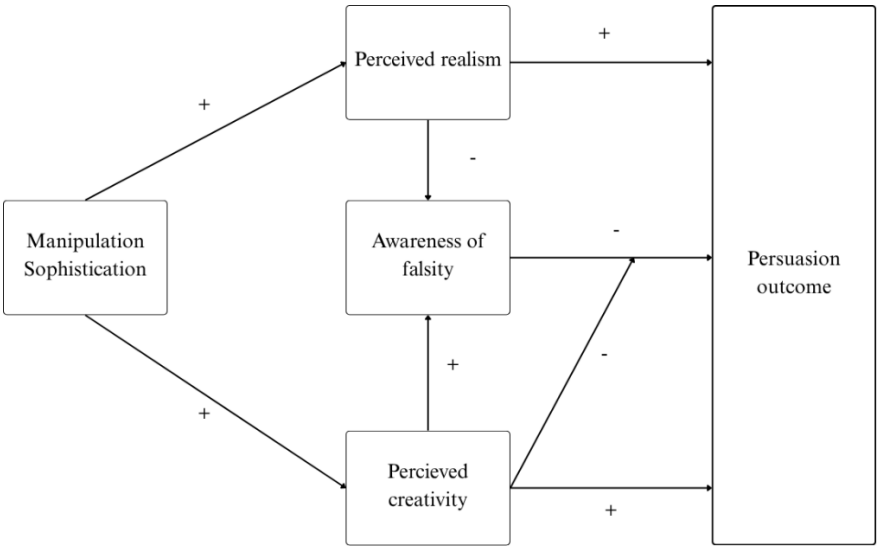
3.1 The consumer response to manipulated advertising framework by Campbell et al. (2021)

In order to better understand how consumer react when confronted with deepfake advertisements we decided to use the framework proposed by Campbell et al. (2021), which is represented in Figure 1. The Framework is useful for understanding how the different advertising manipulations, such as deepfakes, are able to result in higher or lower persuasion outcomes. Figure 1 shows that the idea of manipulation sophistication stands at the beginning of the framework and positively influences the realistic look of the ad and creativity.

The authors outline the evolution of ad manipulation techniques over time, which he classified in three degrees of manipulation sophistication. At first advertisers used analogue methods, for example they could use makeup to make a model's skin appear smoother. Then with the development of digital tools like Adobe Premiere, After Effects, or Photoshop, advertisers could edit images. They could remove skin irritations or irregularities to achieve a flawless skin. Lastly Campbell et al. (2021) describe a third phase: synthetic manipulation. This includes AI-based technologies like generative AI and deepfakes.

According to the authors, as the sophistication of manipulation increases, from analogue, to digital, to synthetic, so does the ad's perceived realism, which they refer to as "perceived verisimilitude" and is the degree on how realistic the advertisement looks. For instance, CGI allows for the creation of highly realistic monsters, which would be difficult using analogue methods like makeup. Additionally, Campbell et al. (2021) argues that higher manipulation sophistication also increases creative potential. An example of synthetic advertising is the campaign #NotJustACadburyAd. During the 2020 pandemic, Mondelez India build a deepfake model of a famous Indian actor. This digital version was made available to small local shops so they could personalize their ads and promote their stores during lockdown (Cadbury Celebrations, 2021). The campaign received the black elephant award from Kyoorius in 2021, which is given for advertisings with outstanding creativity (Kyoorius, 2021). As mentioned before in 1.2 the scope of our study while creativity plays a key role in (Campbell et al., 2021)' s framework, this study focuses specifically on the impact of deepfake manipulated ads on the consumers perceived realism. However, it is important to acknowledge it's the importance of creativity in this construct.

Figure 1: The consumer response to manipulated advertising framework by Campbell et al. (2021).

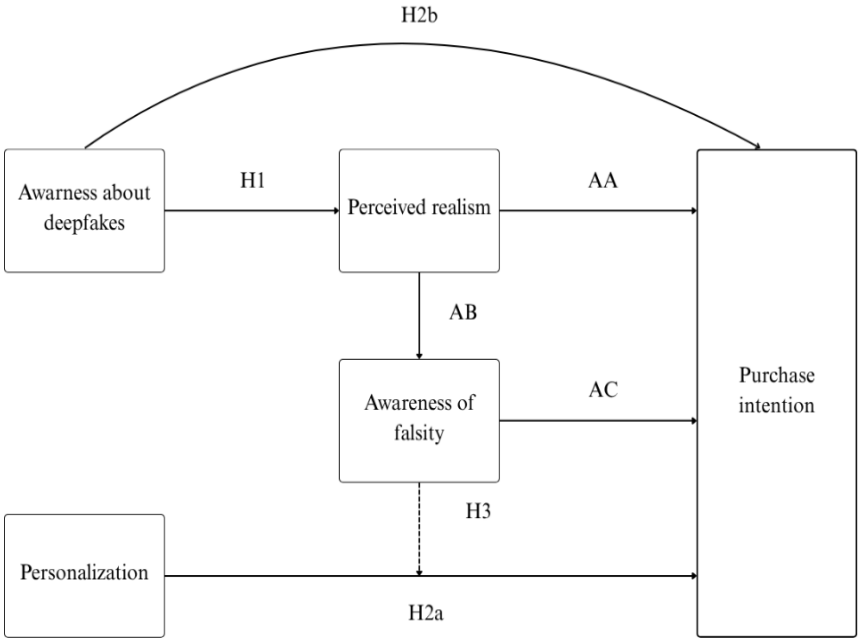


Note. *Reproduced from Figure 3 in Campbell et al. (2021, p.6). The arrows represent the proposed relationships and the +/ - signs indicate the direction of the relationship.*

3.2 Extending the Campbell et al. (2021) framework

Figure 2 presents the final proposed framework, which is based on the model by Campbell et al. (2021)**.** We adapted the framework to the use of deepfake technologies in hyper personalized advertising. The framework aims to explain how personalized deepfake ads influence consumers' purchase intention. We thus included the variable awareness of deepfakes which we believe to affect perceived realism, awareness of falsity, and ultimately purchase intention. The model also illustrates how awareness of deepfakes may influence the relationship between personalization and purchase intention via the mediating role of awareness of falsity. We will now further define the main variables perceived realism and awareness of falsity and then develop the hypothesis represented in Figure 2.

Figure 2: Conceptual framework building upon the model by Campbell et al. (2021)



Note. *The proposed Hypothesis are indicated with the abbreviation H and the assumptions with the abbreviation A.*

3.3 Main variables definition

3.3.1 Perceived realism

As you can see in Figure 2 perceived realism is a central element of the construct.

Campbell proposes the following definition: "*how true, real, or convincing a manipulated ad appears to a consumer*" (Campbell et al., 2021, p. 7). Further, Karpinska-Krakowiak and Eisend (2024, p. 3) define it as "*a subjective evaluation of the degree to which an image, text, or any other type of content reflects the real world.*". However, research also showed that the concept of perceived realism is multidimensional. It cannot be explained by a single aspect alone. Hall (2003), identified, using qualitative methods, six dimensions of perceived realism; these are: plausibility, typicality, factuality, narrative consistency, perceptual quality, and involvement. In contrast, Cho et al. (2012) proposed a model with only five dimensions. They argued that involvement should not be seen as a core part of perceived realism, but rather as an outcome of it and therefore excluded the involvement dimension from the definition. The five dimensions of perceived realism play an important role in advertising, since, as shown by Kim et al. (2019), unrealistic images have been shown to weaken the intended message, therefore if advertisers aim to highlight the positive aspects of a product, realistic representations are better.

Since perceived realism is a key element in Campbell's framework, we will now briefly outline its dimensions.

a. *Plausibility*

Hall (2003) defines plausibility, as the potential that a story told can happen. This idea is not only about avoiding paranormal elements, such as witches or vampires. It also includes avoiding overly idealized situations, like families that appear too perfect. For example, EDEKA recently created a TikTok advertisement showing Albert Einstein promoting a new energy drink (EDEKA, 2025), a situation that is clearly implausible, since the individual is long deceased.

b. *Typicality*

Hall (2003) defines typicality as the possibility that a situation could occur in everyday life. He adds that for something to be typical, it must be plausible. However, not all plausible events are typical. In the context of deepfakes, typicality can become problematic. A deepfake can show someone doing something they would not normally do. For example, a deepfake video of Obama insulting Trump shows behaviour that is clearly untypical for the former president of the United states (BuzzFeed Video, 2018).

c. *Factuality*

Hall (2003) describes *factuality* as the relationship between media representations: how a group of people or an event is portrayed in media; compared to how it really is in the real world. He argues that films based on real events are considered factual. In such cases, factuality is judged based on external sources rather than the viewer's personal experience, like it is done for plausibility or typicality. A recent example is the Netflix series *Inventing Anna*, a story based on a real imposter Anna Sorokin. In this series, the representation was highly factual, for example, the actress wore the same dress as Anna Sorokin did during her real-life court appearances (Ramos-Clemente, 2022). This illustrates the concept of factual accuracy in media. Deepfakes, however, present a major challenge to factuality. Since the person shown in a deepfake never actually performed the actions depicted, factuality is fundamentally absent (Karpinska-Krakowiak & Eisend, 2024).

d. *Perceptual Quality*

Hall (2003) identifies a concept called *perception persuasiveness*, which Cho et al. (2012) later on refers to as *perceptual quality*. This concept denotes to how realistic the visual appearance of media seems. According to Hall's participants, the dinosaurs in Jurassic Park appear very realistic, even though they are not real. Their visual presentation is convincing. Deepfakes can also achieve believable realistic content (Kietzmann et al., 2021), as long as a Deepfake is produced with high quality and does not show visible errors, such as unnatural movements, the viewer will likely perceive it as realistic.

e. *Narrative Consistency*

Hall (2003) also identifies narrative consistency, as the term implies, it refers to how coherent and logically structured the storyline is.

3.3.2 Awareness of falsity

Campbell et al. (2021, p. 9) define *awareness of falsity* as the knowledge that an advertisement has some false elements included, they mention three main elements that can be recognized as false "*product-related claims, product-unrelated claims, and the presented reality*". They further explains that with increasing sophistication in ad manipulation, it becomes more difficult for viewers to distinguish what is real from what is not, thus perceived realism increases. As a result, viewers are more likely to believe that product claims and the presented reality are genuine, which lowers awareness of falsity.

3.4 Development of theoretical assumptions

In order to test the framework from Figure 2, we need to confirm that the part we included from the framework from Campbell et al. (2021) is valid. These assumption are already well tested in research, however we need to make sure that in our experimental setting these assumptions are also true. Thus we have formulated Assumption A, B and C .

3.4.1 Impact of perceived realism on purchase intention

Perceived realism has been shown to influence consumer attitudes. The five dimensions of perceived realism affect three important aspects: (1) consumer identification with the character in the ad, (2) emotional involvement with the narrative, and (3) message evaluation, meaning how convincing the story appears to be; these factors, in turn, have been found to positively influence consumer attitudes (Cho et al., 2012). Karpinska-Krakowiak and Eisend (2024) showed that the dimensions of factuality, perceptual quality, and narrative consistency positively affect consumer attitudes toward ads and, consequently, their choice behaviour. Thus research has shown that perceived realism positively impacts persuasion outcomes, in order for our final construct to work we still need to show that perceived realism is impacting purchase intention positively as expected. Thus we formulate the first assumption as:

Assumption A: *Higher perceived realism will increase consumers' final purchase intentions.*

3.4.2 Impact of perceived realism on awareness of falsity

Further the authors Campbell et al. (2021) explain that viewers are very sensitive towards false information's. He argues that with increasing sophistication in ad manipulation, it becomes more difficult for viewers to distinguish what is real from what is not, thus perceived realism increases. As a result, viewers are more likely to believe that product claims and the presented reality are genuine, which the authors say lowers awareness of falsity.

Assumption B: *Higher perceived realism will decrease consumers' awareness of falsity.*

3.4.3 Impact of awareness of falsity on purchase intention

Further Aljarah et al. (2024) tested experimentally if awareness of falsity influences purchase intention in AI created ads. They manipulated awareness of falsity and showed that greater awareness of falsity reduced online brand engagement. The authors explained this effect through the concept of cognitive dissonance (see definition Appendix 12): the ad created inconsistencies, as the ad attempted to convey a positive message in the context of corporate social responsibility while simultaneously being entirely fabricated thanks to AI and thus this triggered cognitive dissonance and lead to reduced brand engagement. Research has therefore shown that high awareness of falsity decreases purchase intention, this should also be that case in our work. Therefore we formulate the assumption following Campbell's propositions:

Assumption C: *Higher awareness of falsity will decrease consumers' purchase intentions.*

3.5 Hypothesis development

We will now take a look at the main hypothesis that will help us answer our research question.

3.5.1 Awareness of deepfakes and its impact on perceived realism

Figure 2 shows that the first main connection we want to investigate is the connection between awareness of deepfakes and perceived realism. Research has shown that increasing media literacy can reduce viewers' perceived realism (Austin & Johnson, 1997; Scull et al., 2014). This effect can be explained through the Message Interpretation Processing (MIP) model. We apply the version developed by Austin et al. (2007), which offers a simplified explanation of the earlier model introduced by Austin and Meili (1994) in their study on children's responses to alcohol advertisements.

a. The MIP model

The MIP model suggests that message receivers evaluate two dimensions before responding: a rational (cognitive) dimension and an emotional dimension (Austin et al., 2007). Consumers apply heuristics, simple rules of thumb, in both dimensions, but these can be influenced through teaching or media literacy interventions (Arnett, 2007). The rational dimension asks how realistic the portrayed media appears (perceived realism) and how similar it is to the viewer's own life (perceived similarity) (Austin & Meili, 1994). This notion resembles Hall's (2003) concept of perceived realism. In particular, perceived similarity in the MIP model aligns with the dimension of typicality, though Austin and Meili, (1994) framed it more directly as comparing the media with one's own family life, for example, asking children whether alcohol consumption in the media resembles that in their household. The emotional dimension, in turn, asks how desirable the portrayed person or behaviour in the media is. Both the

rational and emotional dimensions shape the level of identification viewers feel with the media. As Austin et al.(2007, p. 486) argue, the central question becomes: *"Do I want to be like this?"*. Positive expectancies then increase the likelihood of adopting the behaviour depicted in the media.

b. *Media literacy and perceived realism*

Austin and Johnson (1997) showed that among children exposed to alcohol advertising with increased media literacy perceived realism reduced. Scull et al. (2014) found similar results but in the context of sexual health education. After exposing adolescents to a new media literacy program, he found that perceived realism decreased and that scepticism of media messages increased. Austin et al. (2007) later examined how media literacy impacts the two dimensions of the MIP model, meaning the rational and the more emotional aspects. Austin et al. (2007) showed that after media literacy interventions, the desirability of the ad had less influence on decision-making, while logical aspects had greater impact. We believe this outcome is highly relevant for deepfake advertising. Consumer may look at more logical aspects and may question whether the endorser's behaviour is typical, whether the perceptual quality of the ad appears realistic, and whether the actions depicted actually happened (the dimension of factuality within perceived realism). Therefore, we argue that increased awareness of deepfake technology and its use in advertising will reduce overall perceived realism.

Hypothesis 1: *A higher level of consumer awareness about the use of deepfake technology in advertising will decrease perceived realism.*

c. *The impact of awareness of deepfakes on the 5 sub dimensions of perceived realism*

Looking back at the studies from Austin and Johnson (1997) and Scull et al. (2014) both studies did not differentiate between the different dimensions of perceived realism. Knowing that the construct perceived realism includes 5 dimensions, we want to closer look at these different dimensions. Thus we assume the following hypotheses for the different dimensions of perceived realism:

If consumers know that advertisers use deepfakes, they may judge the plausibility of the person being in the video differently. There is thus possibility that the celebrity was never present in the advertisement and thus it could reduce the concept of perceived plausibility.

Hypothesis 1a: *Awareness of deepfakes reduces the plausibility dimension of perceived realism.*

If consumers are aware of the use of deepfakes, they will pay closer attention to whether the portrayed person's actions are typical of their normal behaviour. Thus, awareness of deepfakes may reduce the dimension typicality.

Hypothesis 1b: *Higher awareness of deepfakes reduces typicality .*

In line with Karpinska-Krakowiak and Eisend (2024), we assume that with greater knowledge about deepfakes, respondents will evaluate factuality more critically. If they know deepfakes could have been used, they may be less likely to believe that the celebrity was actually present in the video.

Hypothesis 1c: *Higher awareness of deepfakes* reduces *factuality.*

Karpinska-Krakowiak and Eisend (2024) found that in the case of disclosure perceptual quality and narrative consistency did not significantly decrease. However we argue that disclosure and awareness of deepfakes, as part of AI-literacy, have different impacts on the depend variables: if consumers are aware of the deepfake they also have higher knowledge on how to recognize the deepfake and thus following the MIP model the consumer will take more attention on the quality of the video and the consistency. Thus we developed the following 2 hypothesis.

Hypothesis 1d: *Higher awareness of deepfakes reduces perceptual quality* .

Hypothesis 1e: *Higher awareness of deepfakes reduces narrative consistency.*

3.5.2 Hyper personalization and its impact on purchase intention

a. Impact of personalization on purchase intention

Looking back at the proposed framework from Campbell et al. (2021) in Figure 1, which we hold at the core of this work, we see that the framework does not consider how hyper-personalization interacts within it. Even though the authors mention the core importance of this opportunity for marketers, they do not propose any impact that hyper-personalization could have in this construct.

Guo and Jiang (2023) explicitly tested how consumers react when confronted with AI-created hyper-personalized text advertisements They found that this kind of personalization increased both click-through rates and the time viewers spent analysing the ad. The authors explained that personalization raises relevance, as the ad is better adapted to the consumer's needs and preferences.

Personalization of advertisements has been shown to bring several advantages, such as increasing sales and conversion rates (Davenport, 2023). Ad personalization, like adapting the ad content to consumers' preferences, purchase history, or gender, especially in social media, has also been shown to increase brand loyalty and perceived quality, these effects are often mediated by brand attachment and customer engagement, as the ad feels more relatable to the customer (Shanahan et al., 2019). Personalized ads are also more interesting to consumers: Bang and Wojdynski (2016) have shown that personalized ads result in longer fixations. They showed that this is especially true when the respondent is engaged in high cognitive activity. In such cases, the personalized ad was fixated on longer than the non-personalized one. The authors suggests that personalized ads can break through the cognitive load consumers experience in everyday life, while generic ads cannot.

However, several authors have found that personalization can reach a saturation point, where it's benefits diminish. White et al. (2007) found that when using personal data, there is a threshold: If personalization goes beyond that point the effectiveness of the personalization is, reduced because of reactance. Reactance stems from consumers' fear of losing parts of their freedom. As a result, they use coping strategies to prevent this loss of freedom (Brehm, 1989). This fear can also be triggered when marketers attempt persuasion. If consumers recognize the attempt, they may fear that their freedom to behave in a certain way is being influenced or manipulated. Consequently, they respond with coping strategies to resist the persuasive attempt (Fransen et al., 2015). However, White et al. (2007) used personal data such as names and locations, thus reactance might not apply in the context of deepfakes. Semeradova and Weinlich (2019) also found that mid-level personalization performs best, not only compared to low levels, but also to higher levels of personalization. This suggests that too much personalization can be counterproductive. Nevertheless Semeradova and Weinlich (2019) did not adapt the ad content as expected in hyper-personalization. Deepfake powered hyper-personalization would have much more relevance from the content perspective for the viewer than in the case of Semeradova and Weinlich (2019). This raises the question of what happens when the content itself is personalized, such as with the use of deepfakes.

In our case, the use of deepfakes may make ads more interesting to consumers. Personalization through matching the consumer's body type, or using a favourite actor or singer, can increase relevance and usefulness. We therefore argue that hyper-personalization using deepfakes will lead to higher purchase intention.

H2a: *The use of hyper personalized deepfake ads will lead to higher purchase intentions.*

b. Hyper personalization and awareness of deepfakes interaction

The Persuasion Knowledge Model is a model proposed Friestad and Wright (1994), explaining that consumers have a set of knowledge about existing persuasion tactics and persuasion copying methods, which they are recalling to when confronted with a persuasion attempt. The authors explain that this knowledge can be build over time through experience, including knowledge about the advertisers creating the ad, the persuasive tactics used and how we can avoid those persuasion tactics. For example one of these knowledge could be: knowing that marketers use deepfakes to tailor ads to personal preferences.

Following this theory then the increase in awareness of deepfakes could be an increase in persuasion knowledge. Having previously learned that marketers use deepfakes in their persuasion attempts, may lead the consumer to adopt coping strategies to resist persuasion. This can include for more inexperienced consumers rigid beliefs (all ads are bad) or avoiding the ad, and for more experiences

consumers for example selective attention, controlling the attention towards the ad when the ad is presenting a needed product or questioning its message (Friestad & Wright, 1994)**.**

Hypothesis 2b: *Higher awareness of deepfakes reduces purchase intention.*

We thus argue that with increased awareness of deepfakes the positive effect of personalisation might be mitigated. We believe this is because of the increased persuasion knowledge the consumer has about the potential use of deepfakes in order to create highly personalized content . Thus we argue that there is an interaction effect between awareness of deepfakes and personalization: when deepfakes are used in personalized ads, consumers with high awareness will show lower purchase intention.

Hypothesis 2c: *There is an interaction effect between awareness of Deepfakes and hyper personalization such that if the user is aware of the use of deepfakes in hyper personalized ads, then the purchase intention will be lower than if the consumer is not aware.*

We further explain this interaction by arguing that the awareness of falsity moderates the effect of personalization on purchase intention.

We believe that when consumers are aware of, their perceived realism decrease (as suggested in hypothesis 1) and then awareness of falsity increases. This awareness of falsity is a key factor that triggers resistance to persuasion. Once an ad is identified as false, consumers are likely to activate their persuasion coping knowledge and use coping strategies to resist persuasion. As a result, the positive effect of personalization on purchase intention is reduced. We argue that the higher the awareness of falsity, the weaker the positive impact of personalization on consumers' purchase intention.

We therefore propose the following hypothesis:

Hypothesis 3: *Advertisement falsity will negatively moderate the relationship between the use of hyper personalization and customers' purchase intention.*

3.5.3 Perceived likelihood of realism and portrayal of realism

Awareness of falsity is an important variable and we further want to investigate how this variable is impacted.

We assume that the dimensions of perceived realism can be grouped into two main clusters. The authors Karpinska-Krakowiak and Eisend (2024) divided perceived realism into two key dimensions: perceived factuality and portrayal of realism. They define portrayal of realism as:

"*Realistic portrayal is about production quality (how convincingly a depiction reflects reality) and consistency of depiction (the extent to which the integral elements remain coherent).”*(Karpinska-Krakowiak & Eisend, 2024, p. 4). This definition aligns well with the definition we referenced for

narrative consistency and perceptual quality part 3.3.1. Therefore narrative consistency and perceptual quality act similar and can be regrouped as one construct , as proposed by Karpinska-Krakowiak and Eisend (2024).

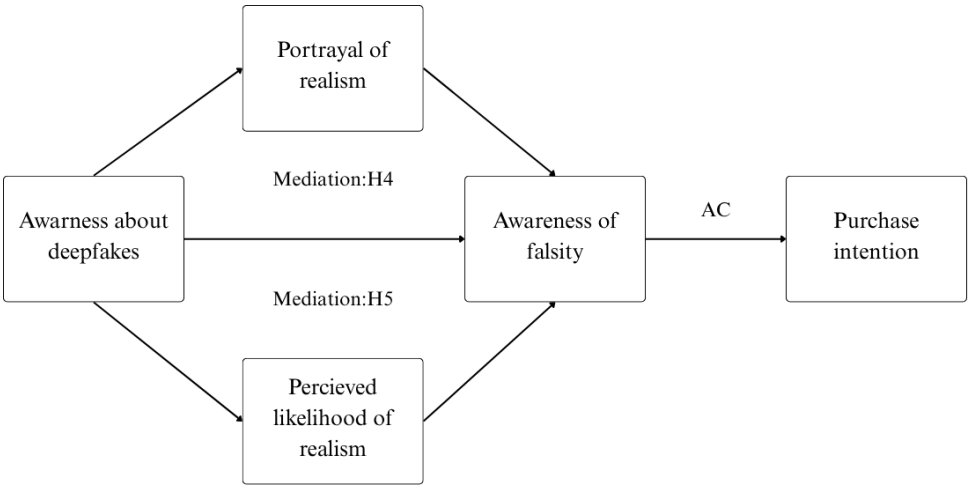
According to Hall (2003), plausibility and typicality are closely related. Hall also explained that typicality depends on plausibility, as something cannot be typical if it is not also plausible. While consumers when assessing plausibility and typicality are comparing the scene they see with their every day experiences, as we have seen in part 3.3.1, when considering factuality consumers are comparing to external sources. All three concepts involve comparing the media content with viewers' prior knowledge or real-world experiences. In contrast, narrative consistency and perceptual quality primarily refer to cues within the media content itself rather than to comparisons with external knowledge sources.Thus these concepts are all quite similar and thus we grouped plausibility, typicality, and factuality into one variable called Perceived Likelihood of Reality.

As illustrated in Figure 2, awareness of deepfakes may reduce perceived realism, which in turn may increase awareness of falsity. We therefore have reasons to believe that perceived realism is a mediator between awareness of deepfakes and awareness of falsity. However, perceived realism is a multidimensional construct, because both dimensions capture different ways in which consumers evaluate the realism of media content, each may explain how awareness of deepfakes influences awareness of falsity. Thus, we argue that the two dimensions of perceived realism: portrayal of realism and perceived likelihood of reality, will each act as a mediator just like shown in Figure 3. We thus formulated the two following hypothesis:

Hypothesis 4: *Portrayal of realism mediates the relationship between awareness of deepfakes and awareness of falsity, such that higher awareness of deepfakes reduces portrayal of realism, which in turn increases awareness of falsity.*

Hypothesis 5: *Perceived likelihood of reality mediates the relationship between awareness of deepfakes and awareness of falsity, such that higher awareness of deepfakes reduces perceived likelihood of reality, which in turn increases awareness of falsity.*

Figure 3: Conceptual model proposed for Hypothesis 4 and 5



Note. *The Hypothesis are presented under the abbreviation H*

IV. Methodology

4.1 Scientific approach

4.1.1 Research type

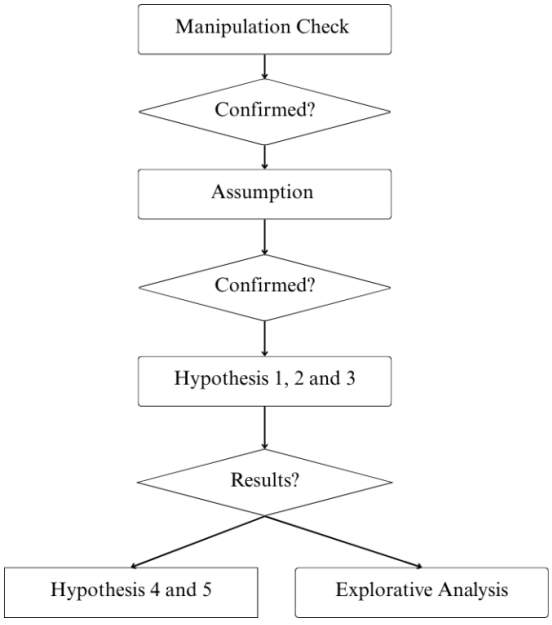
To address our research question (*How does consumer awareness of deepfake-powered hyper-personalization influence purchase intention?*), we followed a deductive approach aimed at testing the previously formulated hypotheses. Our goal is to examine whether the extended framework from Campbell et al. (2021) showed in figure 2 can be applied to real-world scenarios.

While Selent (2022) only performed a rather qualitative research approach, we decided to perform a quantitative research in order to test the proposed framework. By using an experimental design, we intend to identify a cause-and-effect relationship between deepfake-powered hyper-personalized ads and purchase intention. The experimental design we developed is intended to offer deeper insights into this issue by simulating a real life situation.

4.1.2 Research flow chart

As visible in figure 4 this work has followed several step: As the first step is the manipulation check, if the manipulation is confirmed we tested the three assumptions defined in the framework by Campbell et al. (2021). After confirming them, we tested the three main hypothesis 1, 2 and 3. In the last step we tested Hypothesis 4 and 5 and our final explorative analysis.

Figure 4: Flow chart showing step by step how this research was conducted



Note. *If the manipulation check or the theoretical assumptions had not been supported, the experimental design would have required revision. However, both the manipulation check and the assumptions were supported. The assumptions refer to propositions derived from Campbell et al. (2021)'s framework.*

4.2 Data collection

To conduct the experiment, we developed an online survey using SoSci. The online format allowed us to quickly and easily reach test subjects. We chose the platform SoSci, since it provided tools to create an anonymized survey and randomly assign participants to different groups. The questionnaire was in English and the participants took around 5 to 7 minutes to complete it. Respondents were informed about the academic purpose of the study but were not told about the use of Deepfakes, they were only informed that the goal of this research was to explore consumer responses to new technologies in advertising. The data collection period was set to two weeks, so we let the participants answer the questionnaire between the 12.01.2025 and the 26.01.2025.

4.3 Study procedure

4.3.1 The overall design

We decided to use a factorial design, as White et al. (2007) did to test reactance to personalized email marketing. This shows that such designs are commonly used to analyse personalization and it's impact on consumer behaviours. This design allows us to use two categorical variables as independent variables, which we can analyse simultaneously, examining both their main effects and their interaction effects (Hayes, 2017, p. 223,293). Since our research goal is to examine how awareness of deepfakes influences the effectiveness of hyper-personalized deepfake ads, a 2x2 factorial design with two main independent variables is most suitable (awareness of deepfakes: high vs. low × personalization: present vs. absent).

The first independent variable is awareness of deepfakes. This variable is binary: aware vs. not aware. Some respondents were shown a text that informed them about deepfake technology and its use in advertising (see appendix 9). This text was written with the help of ChatGPT. For simplicity we chose a binary variable because we did not want to test varying levels of awareness. Testing different degrees would have gone beyond the scope of this study.

The second independent variable is **hyper-personalization**, also measured as a binary variable. Some respondents saw hyper-personalized ads, while others did not.

In this experiment the test subjects were randomly assigned to one of four groups:

Group 1:	***Group 2:***	***Group 3:***	***Group 4:***
No awareness - No personalization	***Awareness - No personalization***	***No awareness - Personalization***	***Awareness - Personalization***
The control group. These participants did not read the awareness text and were shown a non-personalized deepfake video.	*These participants read the awareness text and were not exposed to hyper-personalized deepfake ads.*	*Participants in this group saw hyper-personalized deepfakes but did not read the awareness text.*	*Participants saw the hyper-personalized ad and also read the awareness text.*

To achieve hyper-personalization, we used celebrity endorsement. The video stimulus featured each subject's preferred celebrities. We chose this method because it offers a simple way to implement hyper-personalization. Adapting ethnicity or body shape would have been difficult to build up technical.

4.3.2 Brand selection

For the stimulus we chose an existing brand, because The stimulus needed to be kept realistic, any other editing than the Deepfake in the video can result in a decrease of one of our main variables: perceived realism. This study used an unknown brand, in order to not disturb our framework. Brand attitude, brand familiarity, brand credibility, perceived fit between celebrity and brand or brand loyalty could impact the purchase intention (Belch & Belch, 2003) and thus distort our results. Therefore we selected a relatively unknown Chinese brand called Flower Knows. It specializes in makeup products with very fashionable, fairy-style packaging. In China, the brand is well known but in Europe, the brand is not yet available in major retail stores but can be purchased through online platforms (Nan, 2025). It seems plausible that such a large Chinese brand is aiming to enter the European market and is using social media advertising to do so. In the questionnaire, we did introduce the brand (see text in appendix 3), but we did not mention that it is a Chinese company. This was done to avoid potential bias or negative attitudes toward its country of origin.

4.3.3 The stimulus

The stimulus chosen was a video that had to meet several criteria: First, it needed to match the brand, *Flower Knows* aesthetic. Second the actor's face needed to be clearly visible and third was the availability of training material. More details about the stimulus video and the chosen celebrities can be found in appendix 11. We selected a video from a fashion influencer Yulia Fomenko (2024), known as @fomenkojulli on Instagram. She was so kind to agree, that we used her video in the context of this master thesis. We added product claims and the *Flower Knows* logo. We then applied a deepfake, replacing her face with those of selected celebrities. The final included celebrities are featured in appendix 11. For the non-personalized deepfake, we chose a celebrity not widely known in the West: Indian actress Neha Sharma.

To create the deepfake video, we chose to used a program called DeepFaceLab. We selected this software after testing several other deepfake apps, most of which produced either low-quality results or deepfakes in which the celebrity was barely recognizable (e.g., FaceSwapper). Although DeepFaceLab is more complex and time-consuming, and requires greater computing power (we used an NVIDIA RTX 4070 graphics card), the results were significantly more satisfactory. DeepFaceLab offers different training methods. After testing several options, we chose the "Quick 96" training method, which is the fastest and simplest, requiring relatively low GPU resources. It produced acceptable results after a short training period. It was important that the overall video quality remained consistent across the celebrities.

We could observe minor inconsistencies (e.g., Lady Gaga's eye colour changed slightly), on average, the visual differences between the videos were negligible (visible in table a16 in appendix 3).

It is also important to note that the videos were only accessible through the questionnaire. They were not publicly available and will not be published, in order to protect the personal data of the celebrities involved.

4.3.4 The questionnaire construction

In the following, we outline the concrete structure of the questionnaire, the full questionnaire can be seen in appendix 9.

First we introduced the purpose of the study, and then all participants were presented with an introduction to the brand, *Flower Knows*. A randomizer was then used to assign each participant to one of four groups. Groups 2 and 4 received a text intended to influence their awareness level, while Groups 1 and 3 proceeded directly to the next step. Then the participants were given a list of celebrities. Each entry included the celebrity's name along with a short description. Neha Sharma, the celebrity for the control group, was not included in this list. All participants were asked to choose their favourite celebrity from the list. Afterwards, participants were shown a stimulus video: For Groups 3 and 4, the video featured their previously chosen favourite celebrity, while for Groups 1 and 2, the video was a deepfake of Neha Sharma.

We tested the questionnaire (set of 3 people) and received feedback, thus we thought it was necessary that all participants required to watch the video twice. The videos were very short (13 seconds), this corresponds to the typical length of videos on social media platforms. But we wanted to ensure that the stimulus had enough impact on the participants.

After viewing both videos, participants were informed that the videos were identical. All groups were then presented with the same set of questions.

4.3.5 Sample characteristics

The original sample consisted of N = 293 participants. We excluded those who did not finish the questionnaire, failed the attention checks, were underage, or over 70 years old (for more details, see Annex 1). The final sample included 133 participants (Women = 99 ; Mage = 27.16; SDage = 6.22).

Due to limited resources, we used a non-probabilistic sampling method. We decided to combine convenience and purposive sampling.

Convenience sampling, since we have no possibilities to access large databases in order to make a probabilistic sampling we chose to reach as much participants by distributing the online survey via direct

messaging through close contacts and posted it on LinkedIn, Instagram, and WhatsApp. We also used platforms like SurveyCircle and SurveySwap to reach more participants.

Since the aim of the study was to test the impact of personalized advertising, we needed a sample that matched the brand and product. Including many participants with little interest in the product would not have made sense, thus we also decided to use purposive sampling. According to Similarweb (2025), 78% of Flower knows customers are women, and 70% are between 18 and 34 years old. We therefore mentioned on SurveyCircle and SurveySwap that the study was for women and/or people interested in makeup. Since these platforms mostly feature young students looking for study participants, we could recruit a sample similar to the brand's target audience. The final sample matched well: 74.4% of participants were women, 25.6% men or non-binary. Additionally, 90.9% were aged 18–34, aligning with the brand's main target group.

Further in our sample 82.7% had at least a bachelor's degree or higher. Most participants were located in the European Union and only 6.9% were from outside the EU (e.g., China, UK, USA, Australia). However, the German-speaking countries are mainly represented in this sample with 87.3% (Germany 62.4%, Austria 22.6%, Switzerland 2.3%).

We then examined whether the sample characteristics as size, age and gender were similarly distributed across the four experimental groups (G1, N = 33; G2, N= 30; G3, N = 35; G4, N= 35) and found that age and gender were similarly distributed in the four groups. For more details about the tests see appendix 1.

4.4 Variables scale development

4.4.1 Controlled variables

We also defined some variables that needed to be controlled. The first is creativity: as Campbell et al. (2021) noted, degrees of creativity can influence the awareness of falsity and, in turn, purchase intention. To prevent creativity from affecting the setup, we decided to use the same video for all groups and celebrities. Only the face was changed; all other aspects of the ad remained the same.

The second variable we needed to control was celebrity endorsement [3]. The endorsement effect across different celebrities should be similar in all videos. To assess the effectiveness of the endorsement, we used the "Appropriateness of the Celebrity" scale developed by Edwards et al. (2009, as cited in Pfiffelmann, n.d.). We believe this scale can indicate whether the viewer thinks the celebrity is a good fit for the brand. It is important that, even when different celebrities are shown, there are no significant differences between the groups. We later tested whether there is an different degree of celebrity

[3] Celebrity endorsement is when a marketer chooses a celebrity as a spokesperson for their brand, so that the celebrity's positive image will transfer to the brand. However, in order to achieve increased purchase intention the marketers must ensure that the celebrity is a good fit for the brand. (Peter & Olson, 2010)

endorsement between the seven celebrities and found that the celebrity and found that this differences was not significant. However celebrity endorsement has an impact on our depended variable purchase intention and thus we need to control this variable when purchase intention is tested. (see appendix 3a)

The third variable is brand attitude[4]. We decided to add this measure because, even though the brand is unknown, it is important to ensure that the brand introduction text itself does not influence brand attitudes. To measure brand attitude, we used a scale developed by MacKenzie et al. (1986, as cited in Pfiffelmann, n.d.). Brand attitude had no impact between the groups but as expected it influences purchase intention and thus needs to be controlled (see appendix 3b)

4.4.2 Main variables

a. Perceived realism

To measure perceived realism, we used the scale developed by Cho et al. (2012), which assesses five dimensions of perceived realism we presented earlier in part 3.3.1. The overall scale showed high reliability with a Cronbach's α of 0.90 (see appendix 2 table a7 for the internal consistency of the sub-dimensions).

b. Awareness of falsity

Since we did not find any scale to measure awareness of falsity, we decided to developed a new scale based on the three aspects proposed by Campbell et al. (2021, p. 9): (1) product-related claims, (2) product-unrelated claims, and (3) presented reality. We adapted the scale celebrity endorsement by Kim, Ratneshwar, and Thorson (2017, as cited in Bruner II, 2019), then we adapted the semantic differential scale by Eisend et al. (2014) and the scale from Pfiffelmann et al. (2023) for presented reality (see appendix 6).

c. Purchase intention

To measure purchase intention we used the Likert scale developed by Park et al. (2021, as cited in Pfiffelmann, n.d.). We chose a Likert scale instead of a simple binary question. Because we wanted to obtain quantitative results. Quantitative data enabled us to use parametric tests, that offer greater statistical power than non-parametric tests, especially important when comparing means (Bortz & Schuster, 2011).

4.5 Manipulation check

In order to ensure that our manipulation worked as intended we performed a manipulation check. We tested whether the initial awareness text was able to significantly increase the respondents AI literacy

[4] Brand attitude refers to the consumer's overall evaluation of a company's brand and has been shown to influence purchase intention (Peter & Olson, 2010).

and if the perceived personalization increased when the respondents were shown the chosen celebrities. In order to see a more detailed description of the manipulation check please take a look at appendix 5 and appendix 10 to see the full questionnaire of the manipulation check.

4.6 Data analysis overview

We will now take a short look at the statistical methods we used to test our hypotheses. We used a confidence level of 95% (corresponding to a significance level $\alpha = 0.05$). For our analysis, we used the program SPSS. Concerning the statistical tests used, we built table 1, representing the hypotheses, the statistical tests applied, and the assumptions tested.

Table 1: Overview of the hypothesis and the respective statistical tests.

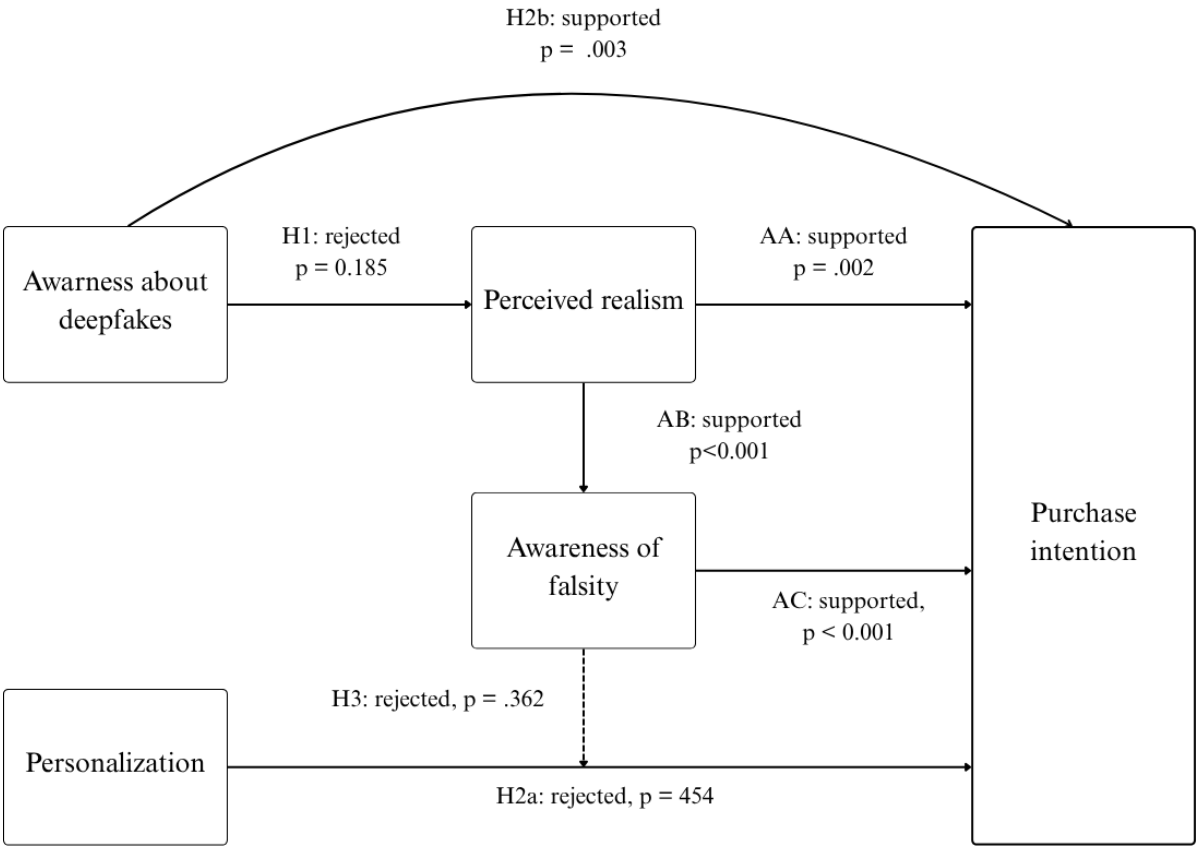
Hypothesis	IV	DV	Moderator / Mediator / Covariate	Statistical Test
AA	PR	PI	CV: CE, BA	Multiple Linear Regression
AB	PR	AF	-	Linear Regression
AC	AF	PI	CV: CE, BA	Multiple Linear Regression
H1	AD	PR	-	Independent Samples t-Test
H2	AD, P	PI	-	Two-Way ANOVA
H3	P	PI	Moderator: AF	Moderation (PROCESS Model 1, Hayes 2017)
H4	AD	AF	Mediator: PoR	Mediation (PROCESS Model 4, Hayes 2017)
H5	AD	AF	Mediator: PLR	Mediation (PROCESS Model 4, Hayes 2017)

Note. *PR: Perceived realism; AF: Awareness of Falsity, AD: Awareness of deepfakes, P: Personalization, PI: purchase intention, PoR: Portrayal of realism, PLR: Perceived Likelihood of Realism. CV are the control variables including CE, celebrity endorsement and BA brand attitude. See detailed table in appendix 4.*

V. Results

In the following, we present the results of our data analysis. Our goal was to identify how awareness of deepfake-powered hyper-personalization influences consumer purchase intention. To do this, we developed a framework based on Campbell et al. (2021). The framework explains consumer reactions to deepfake advertisements. As already shown previously in Figure 2 the framework considers the variables perceived realism of the advertisement and the awareness of falsity. The Figure 5, shows a summary of the results, it shows the hypothesis that have been rejected and supported.

Figure 5: Results of the build-up concept from figure 2



Note. *The Hypothesis are presented under the abbreviation H, the figure indicates whether the Hypothesis have been rejected and the p-values. Hypothesis 7, the interaction effect between awareness of deepfakes and personalization has not been represented. Hypothesis 7 has been rejected, p = .919 .*

5.1 Manipulation check

For the manipulation check we had a sample of 48 respondents, see appendix 5 for more detail results: We first tested whether there is a measurable difference in AI literacy between Group 2a (N = 20, M = 5.53, SD = .72), who read the deepfake introduction text, and Group 1a (N = 28, M = 5.02, SD = 1.48),

who did not. A t-test was conducted and yielded a significant result (t(41.45) = -2.20, p = .034). The effect size was on a medium level (d = -0.58), indicating as expected a higher level of AI literacy in Group 2. We can therefore conclude that the deepfake awareness manipulation was effective. We performed the same analysis for the variable personalisation. We performed a Welch-t-test between the two groups(Group 1b: without personalisation / Group 2b: with personalisation). We therefore performed a Welch T-test, which turned out to be significant (t(29.33) = -3.52, p = .001). The effect size was large (d = -1.11), showing that the group that was shown the preferred celebrity (Group 2b) had a higher perceived personalisation than Group 1b.

We then tested whether the quality of the videos were high and found that both groups' perceived realism scores were relatively low (M group no awareness = 3.09, M group with awareness = 2.94, on a Likert scale from 1 to 7 , where 7 is the highest possible perceived realism). In our manipulation check, the measured AI literacy of both groups was at a higher level (Mean group no awareness = 5.02, Mean group awareness = 5.73). We also assessed whether the awareness text was perceived as a disclosure. Participants rated two items: “I perceived the text at the beginning as a disclosure about the ad” and “The introductory text appeared to provide general information about the upcoming video.” Responses were measured on a 7-point Likert scale. After reverse coding the second item, the two items showed acceptable internal consistency (Cronbach’s α = 0.68). The mean score was M = 3.82 (SD = 0.77, N = 20). Given that the midpoint of the scale is 4, this result indicates that the introductory text was perceived as relatively neutral and not clearly as a disclosure.

5.2 Assumption testing

5.2.1 Assumption A: Perceived realism as a predictor of purchase intention

In Assumption A we stated that perceived realism will have a positive impact on the consumer purchase intention. To test this, we chose to conduct a multiple linear regression. Since we had two quantitative variables and assumed a linear relationship between perceived realism and purchase intention.

However, we also identified that purchase intention (PI) may be influenced by celebrity endorsement and brand attitude. Therefore, these two variables needed to be controlled.

Table 2 shows the result of the multiple linear regression analysis with perceived realism, brand attitude, and celebrity endorsement as predictors. The regression was significant (F(3,129) = 8.26, p< .001) and the model explained 14% of the variance. Table 2 shows that perceived and brand attitude had a significant positive relationship towards purchase intention. However, Celebrity Endorsement did not have a significant implication in the model. The relationship between Perceived realism and Purchase intention was significant and ß was positive we can reject the Null Hypothesis. Since the relationship can be explained theoretically we can assume that there is causality and we can thus support assumption A.

Conclusion: Assumption A is supported and an increase in perceived realism increases purchase intention.

Table 2: Multiple linear regression predicting purchase intention from the variable perceived realism, controlled by the variables brand attitude and celebrity endorsement

Effect	Estimate (B)	SE	95% CI		t	p
			LL	UL		
Intercept	.233	.761	-1.272	1.737	.306	.760
Perceived Realism	.422	.132	.160	.684	3.191	.002
Celebrity Endorsement	-.136	.085	-.303	.032	-1.60	.111
Brand Attitude	.303	.132	.042	.564	2.29	.023
Model summary	R^2	Adjusted R	df1	df2	F	P
	.161	.142	3	129	8.264	<.001

Note. *Dependent variable: Purchase intention.*

5.2.2 Assumption B: Perceived realism and its effect on awareness of falsity

We then wanted to identify whether there is a relationship between perceived realism and awareness of falsity. Since we have here two quantitative variable and assume a linear negative relationship, we decided to use a simple linear regression. The control variables do not need to be applied here.

Table 3 shows that the overall Model is significant ($F(1,131) = 95.84$, $p<.001$) and explained 42.3% of the variance. The predictor variable was perceived realism and it had a significant impact on the dependent variable awareness of falsity, in figure 6 we can clearly see that the relationship is negative.

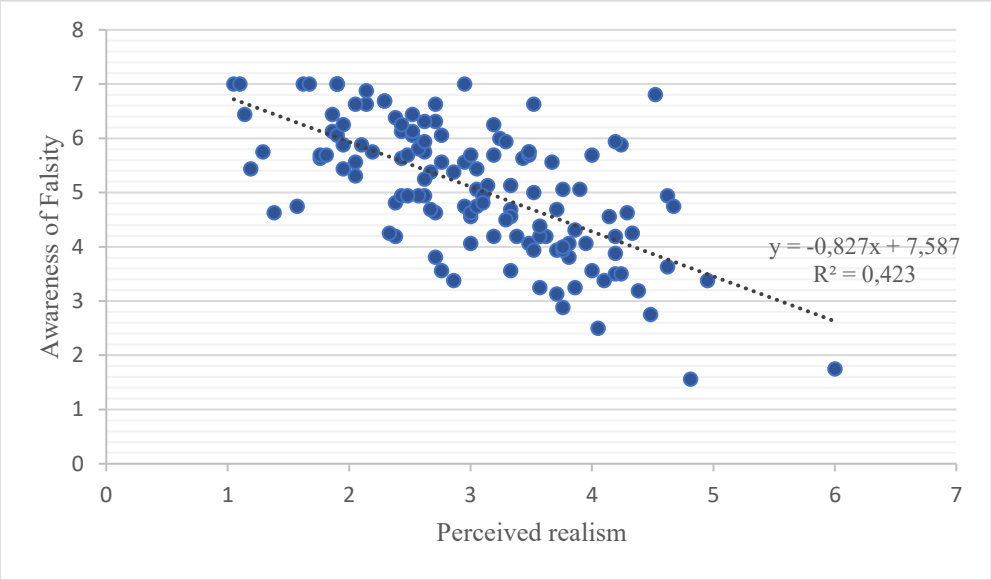
Conclusion: Assumption B is supported and an increase in perceived realism decreases awareness of falsity.

Table 3: Simple linear regression predicting awareness of falsity from perceived realism

Effect	Estimate (B)	SE	95% CI		β	t	p
			LL	UL			
Fixed effects							
Intercept	7.587	.267	7.06	8.11	-	28.43	< .001
Perceived Realism	-.827	.084	-.99	-.66	-.650	-9.79	< .001
Model summary	R^2	Adjusted R^2	df_1	df_2	F		p
	.423	.418	1	131	95.84		< .001

Note. *Dependent variable: Awareness of Falsity. Predictor: Perceived Realism.*

Figure 6: Linear regression analysis of assumption B



Note. *The Graph shows a negative linear relationship between perceived realism and awareness of falsity*

5.2.3 Assumption C: Awareness of falsity as a predictor of purchase intention

Afterwards we tested if awareness of falsity has a negative impact on the consumers purchase intention. Just like before we decided to perform a regression analysis, since both variables are ordinal. However, purchase intention is impacted by celebrity endorsement and brand attitude, therefore we also need to control both of these variables. We thus performed a multiple regression analysis with three predicators: awareness of falsity, brand attitude and celebrity endorsement. Table 4 shows that the model is significant ($F(3,129) = 13.70$, $p < .001$) and it explains 24.2% of the overall variance. Celebrity endorsement appeared to be not significant, but brand attitude the second control variable is significant. Our main variable Awareness of falsity is significant. As we can see from table 4 $\beta = -.53$, is negative and we can conclude that the relationship between purchase intention and Awareness of falsity is negative.

Conclusion: Assumption C is supported and an increase in awareness of falsity results in a decreases in purchase intention.

Table 4: Multiple linear regression predicting purchase intention from awareness of falsity, brand attitude, and celebrity endorsement

Effect	Estimate (B)	SE	95% CI		β	t	p
			LL	UL			
Fixed effects							
Intercept	3.863	.797	2.286	5.440	-	4.846	< .001
Awareness of Falsity	-.527	.106	-.736	-.318	-.432	-4.993	< .001
Brand Attitude	.282	.125	.034	.530	.176	2.256	.026
Celebrity Endorsement	-.019	.086	-.189	.151	-.019	-0.220	.826
Model summary	R^2	Adjusted R^2	df_1	df_2	F	-	p
	.242	.224	3	129	13.70	-	< .001

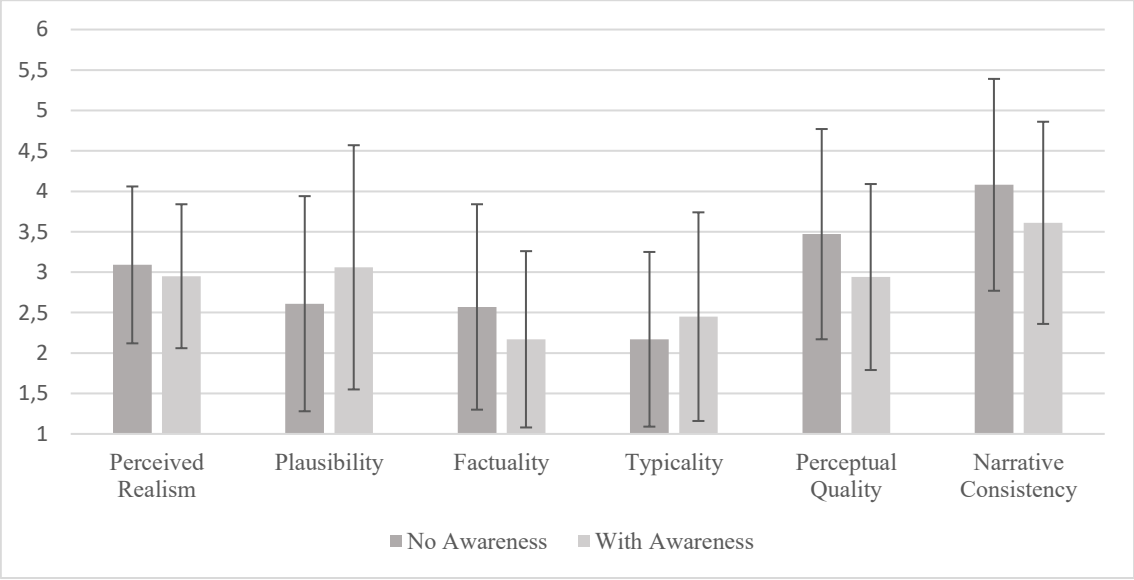
Note. *Dependent variable: Purchase Intention. Predictors: Awareness of Falsity, Brand Attitude, and Celebrity Endorsement.*

5.3 Main experimental results

5.3.1 The Effect of awareness of deepfakes on perceived realism and its underlying mechanisms

a. Hypothesis 1: Impact of deepfake awareness on perceived realism

Figure 7: Compared mean values of perceived realism and the 5 sub dimension between changed awareness of deepfakes



Note. *Means and SD of the tested groups are represented here visually. We can see that perceptual quality and narrative consistency showed significant differences between the two groups.*

In the first hypothesis, we examined whether manipulating respondents' awareness of deepfakes affected their perceived realism. Awareness of deepfakes was treated as a binary variable (N Group A (not aware) = 68, N Group B (aware) = 65). Thus, we aimed to test whether there was a significant

difference in the means between the two groups. We therefore conducted a t-test, visible in table 5. As mentioned in appendix 4, the assumptions for the t-test were all met.

The t-test showed that the difference in the Means were not significant ($t(131) = 0.90$, $p = 0.371$). We can see in Figure 7 that the means difference between Group A and Group B for perceived realism are not very much different. Respondent unaware of what deepfakes are ($M = 3.09$,$SD = 0.97$) did not show significant difference in perceived realism compared to those aware of deepfakes ($M = 2.94$, $SD = 0.89$). We thus need to reject Hypothesis 1.

Conclusion: Hypothesis 1 could not be supported.

Table 5: Results of independent samples t-test of awareness of deepfakes on perceived realism and the MANOVA on perceived realism subdimensions

T-test	Group A		Group B		t	df	p	Cohen's d
	M	SD	M	SD				
Perceived Realism	3.09	0.97	2.95	0.89	0.90	131	.371	0.16
MANOVA	Group A		Group B		F	df	p	Partial η^2
	M	SE	M	SE				
MANOVA	-	-	-	-	5.85	5,127	< .001	.187
Plausibility	2.61	.17	3.06	.18	3.30	1,131	.072	.025
Factuality	2.57	.14	2.17	.15	3.76	1,131	.055	.028
Typicality	2.17	.14	2.45	.15	1.91	1,131	.169	.014
Perceptual Quality	3.47	.15	2.94	.15	5.97	1,131	.016	.044
Narrative Consistency	4.08	.16	3.61	.16	4.40	1,131	.038	.032

Note. *SD: standard deviation; SE: standard error. The independent t-test tested Hypothesis 2. The Manova tested the impact awareness of deepfakes has on the 5 sub dimensions of perceived realism Hypothesis 2.1 to 2. Group A is the unaware group and Group B is the one were respondents are ware of deepfakes.*

b. *Hypothesis 1a-1e: Impact of awareness of deepfake on the dimensions of perceived realism*

However, perceived realism consists of five different dimensions. Therefore, we examined whether awareness of deepfakes had an impact on one or several of these dimensions individually.

We then wanted to analyse more deeply whether there are differences between the dimensions of perceived realism. To do this, we decided to use a MANOVA, which is also represented in table 5. The MANOVA allows us to test whether a categorical independent variable influences several dependent variables at the same time, without increasing the Type I error that would occur when performing multiple t-tests.

We therefore conducted a multivariate ANOVA with awareness of deepfakes as the categorical independent variable and five dependent variables representing the five dimensions of perceived realism: perceptual quality, narrative consistency, plausibility, typicality, and factuality.

The goal is to identify whether the means between Group A (not aware of deepfakes) and Group B (aware of deepfakes), show any differences across the 5 dimensions in perceived realism. The assumptions needed were met (See appendix 4).

The MANOVA revealed highly significant effects of awareness of deepfakes on the combined dependent variables (Wilks' Lambda = .813, $F(5,127) = 5.85$, $p < .001$, partial $\eta^2 = .187$). Indicating that 18.7% of the variance in the 5 subdimension of perceived realism can be explained by awareness of deepfakes. We then performed post-hoc analyses with univariate ANOVAs for each dependent variable to test Hypotheses 1a to 1e. In Figure 7 we can well see that perceptual quality and narrative consistency seem to be the only two dimensions that shows visible differences in means between the two awareness groups.

Hypothesis 1a: Plausibility

To test whether Awareness of deepfake has an impact on plausibility we looked at our post hoc analysis using an univariate ANOVA. Result show that the differences in the Means are not significant ($F(1,131) = 3.30$, $p = .072$, partial $n^2 = .025$). Group A has a lower perceived plausibility ($N = 68$, $M = 2.61$, $SD = 1.33$) then the Group B ($N = 65$, $M = 3.06$, $SD = 1.52$). But the difference is not significant and we thus need to reject Hypothesis 2.1

Conclusion: Hypothesis 1a was rejected.

Hypothesis 1b: Typicality

We then tested whether *awareness of deepfakes* had a significant impact on the dimension of ***typicality***. We also used the univariate ANOVA from our post-hoc test of Group A ($M = 2.17$, $SD = 1.08$) and Group B ($M = 2.45$, $SD = 1.29$). The test turned out to show no significance ($F(1,131) = 1.91$, $p = .169$, partial $n^2 = .014$). Thus we needed to reject Hypothesis 2.2.

Conclusion: Hypothesis 1b was rejected.

Hypothesis 1c: Factuality

Finally we tested whether factuality got reduced with increasing awareness of deepfakes, also using the results from our follow up ANOVAs. The results showed no significant differences ($F(1,131) = 3.76$, $p = .055$, partial $n^2 = .028$) Group A that was not aware ($M = 2.57$, $SD = 1.27$) had a higher factuality than

Group B that was aware (M = 2.17, SD = 1.09), however the difference is not significant and Thus we reject Hypothesis 2.3.

Conclusion: Hypothesis 1c was rejected.

Hypothesis 1d: Perceptual Quality

Hypothesis 1d stated that awareness of deepfakes will decrease perceptual quality. The assumptions for the test have been met, we tested these using the Levene test for the homogeneity of the variances (F(1,131) = 0.73, p = .394) and the Shapiro wilk test to test for normal distribution (WGroupA = .96, pgroupA = .45; WgroupB = .97 pgroupB = .150) was normally distributed. In Group A, even though the Shapiro Wilk test did show significant results , the QQ plot of figure a13 in appendix 4 looked acceptable, thus we assume that normality is given.

We used the post hoc analysis from the previous MANOVA to identify whether the differences between the groups in the dimension perceptual quality were significant. Results showed that for perceptual quality the differences were significant (F(1,131) = 5.97, p = .016, partial n^2 = .044). Thus there might be a significant difference between the mean of Group A (not aware) (N = 68, M = 3.47, SD = 1.30) and Group B (aware) (N = 65, M = 2.94, SD = 1.15) . Thus we can say that awareness of deepfakes reduces perceived quality of the video material.

Conclusion: Hypothesis 1d was supported.

Hypothesis 1e: Narrative Consistency

To test our hypothesis 1e, we tested whether they are significantly different through the post hoc of the MANOVA. Results showed to be significant (F(1.131) = 4.34, p = .038, partial n^2 = .032). The non aware Group A (M = 4.08, SD = 1.31) and the aware Group B (M = 3.61, SD = 1.25) are not equal. Narrative consistency seems to be reduced with higher awareness of Deepfakes.

Conclusion: Hypothesis 1e was supported.

5.3.2 The effect of personalization and awareness of deepfakes on purchase intention

a. Hypothesis 2a-2c: Investigating interaction between personalization and awareness

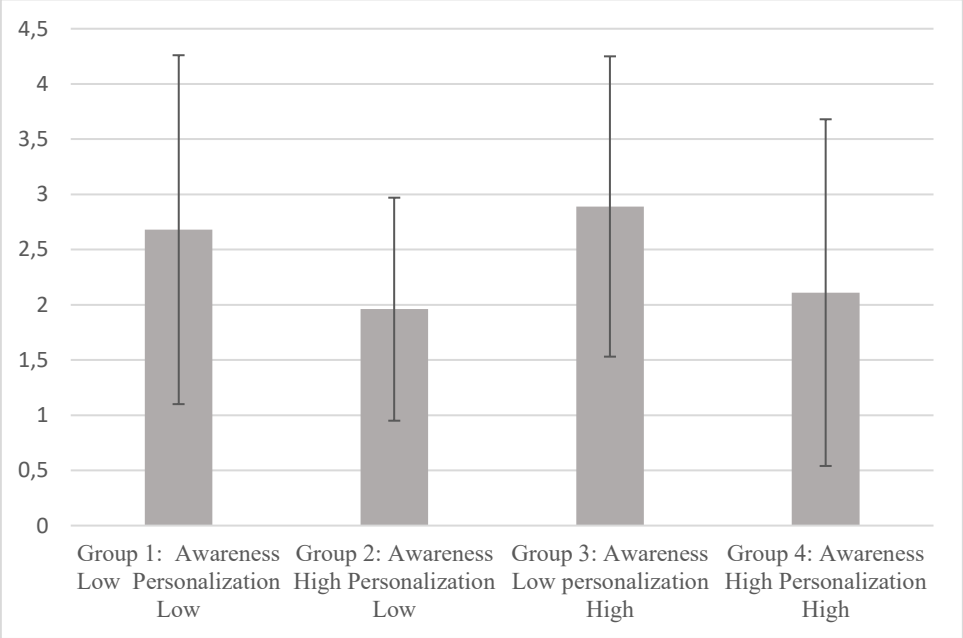
We will now take a look at Hypothesis 2a, 2b and 2c. The Hypothesis state that hyper personalization increases purchase intention (H2a), that deepfake awareness will decrease purchase intention directly (H2b), and that there is an interaction effect between the two variables awareness of deepfakes and hyper personalization (H2c). In Figure 8 we can see that Group 1 and Group 2, Group 3 and Group 4 have clear differences. Indicating that Awareness of Deepfakes reduces purchase intention. However Group 1 and Group 3, Group 2 and Group 4 seem to be very similar showing that with increased personalization the purchase intention might not change.

To test the following 3 hypothesis we performed an Two WAY ANOVA, which can be seen in detail in table 6. Since we have two factors (awareness of Deepfake and personalization) that influence one

dependent variable (purchase intention). The two way ANOVA is most suitable in testing the two main effects from hypothesis 2a, 2b and in testing the interaction effect of Hypothesis 2c (Hayes, 2017). The assumption from the ANOVA are met, the Levene test showed homogeneity in the variances ($F(3,129) = 1.60$, $p = .192$) and the assumption of normality is confirmed by looking at the QQ plots from Purchase intention (see appendix 4, figure a14-a 17).

The descriptive results for awareness of deepfakes are: Group 1 (N1 = 33, M = 2.68, SD= 1. 58) and Group 2 (N2 = 30, M =1.96, SD = 1.01) , Group 3 (N1 = 35, M = 2.89, SD = 1.36) and Group 4 (personalized) (N2 = 35, M = 2.42, SD = 1.45) . These differences can also be well seen in figure 8. In the following Table 6 you can see the cross descriptives with the 4 different groups showing Personalized X Awareness of deepfakes.

Figure 8: Means of respondents purchase intention depending on the different experimental Groups



Note. *The graph shows that there is an impact of awareness of deepfakes on purchase intention but not an interaction between awareness of deepfakes and personalization since the differences between Group 1 and Group 3, Group 2 and Group 4 are not big enough.*

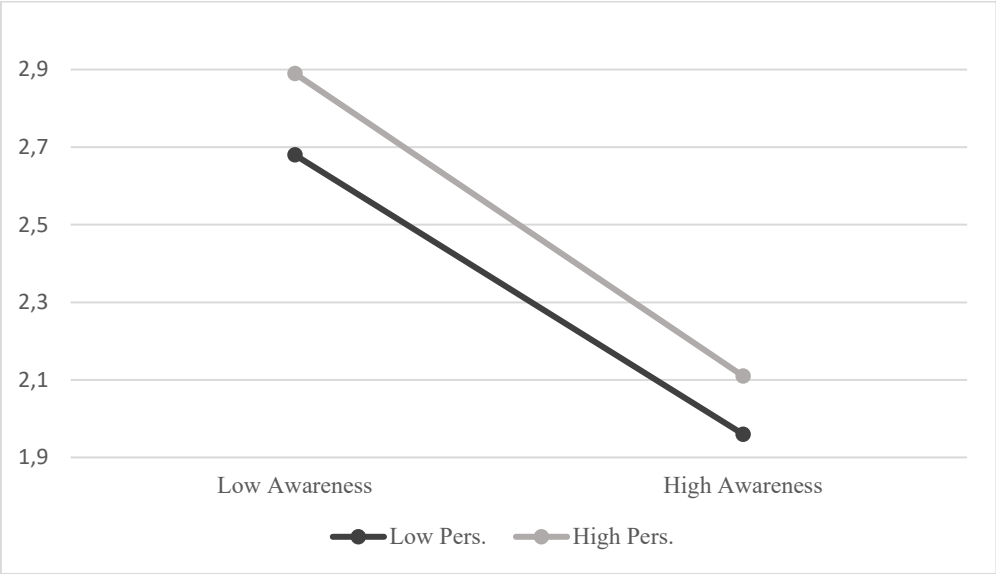
After performing the Two way ANOVA we saw that the overall model was significant ($F(3,129) = 3.27$, $p = .023$). As visible in table 6 the first main effect, Hypothesis 2a, the relationship between personalization and purchase intention were not significant ($p = 0.454$). Which means that we reject Hypothesis 2a.

Looking at the second main effect for Hypothesis 2b on table 6, we can see that the main effect of awareness of deepfake on purchase intention is significant ($p = 0.003$). We can thus support Hypothesis 2b.

Further when analysing the interaction effect we also could not find a significant interaction between the variables awareness of deepfakes and personalization (partial $\eta^2 = 0.00$, $F(1,129) = .011$, $p = .919$).

Figure 9 shows the interaction effect, it shows two parallel lines indicating no interaction between the two variables, we therefore also need to reject Hypothesis 2c.

Figure 9: Interaction between deepfake awareness and personalization on purchase intention



Note. *The Graph shows that there is no interaction between awareness of deepfakes and personalization*

Conclusion: Hypothesis 2b has been supported. However Hypothesis 2a and 2c have been rejected.

Table 6: Two-way ANOVA examining the effects of awareness and personalization on purchase intention

Source	Sum of squares	df	Mean Square	F	p	Partial η^2
Awareness (A)	18.447	1	18.447	9.28	.003	.067
Personalization (B)	1.119	1	1.119	0.56	.454	.004
A × B Interaction	0.021	1	0.021	0.01	.919	.000
Error	256.468	129	1.988			
Total	1055.556	133				
Corrected Total	275.977	132				
Model Summary	R^2	Adjusted R^2				
	.071	.049				

Note. *Dependent variable: Purchase Intention.*

b. Hypothesis 3: Moderation analysis with awareness of falsity as moderator

The last hypothesis we have defined was the assumption of a moderation effect between the variable awareness of falsity on the relationship of personalization and purchase intention. Moderation analysis was conducted using PROCESS (Model 1; Hayes, 2017) with 5,000 bootstrap samples and HC3

heteroscedasticity-consistent standard errors. Since personalization is a categorical variable we used the dummy-coded variable (0 as no personalization and 1 for personalization). Following the recommendations by Hayes (2017), continuous variables were mean-centered before computing the interaction term to reduce multicollinearity. We also added our two control variables celebrity endorsement and brand attitude since they can have an impact on purchase intention.

The overall Model was significant ($F(5,127) = 10.10$, $p < 0.001$), which indicates that the model performs better than no model. The model explains 26% of the variance ($R^2 = .26$, adjusted $R^2 = .24$). Looking at the interaction effect between personalization and awareness of deepfake ($b = -.20$, $SE = .22$, $t(127) = -.92$, $p = .362$) we see that there is no significant interaction effect and the model did not contribute additional explained variance (only 0.64% additional variance) ($\Delta R^2 = .006$, $F(1,127) = .84$, $p = .361$).

The control variable celebrity endorsement is not significantly impacting the model ($b = -.01$, $SE = .09$, $t(127) = -.16$, $p = .872$). The second control variable brand attitude however has a significant impact on the dependent variable ($b = .29$, $SE = .14$, $t(127) = 2.17$, $p = .032$). However, just like in the two way Anova the moderation analysis shows that personalization doesn't have a significant impact on purchase intention ($b = 1.36$, $SE = 1.21$, $t(127) = 1.12$, $p = .264$), while awareness of deepfakes does ($b = -.42$, $SE = .17$, $t(127) = -2.47$, $p = .015$) . Figure 10 shows that the there is seemingly no interaction between awareness of deepfakes and personalization, however the two lines are connecting in the end which shows that there might be a light interaction which however is not significant.

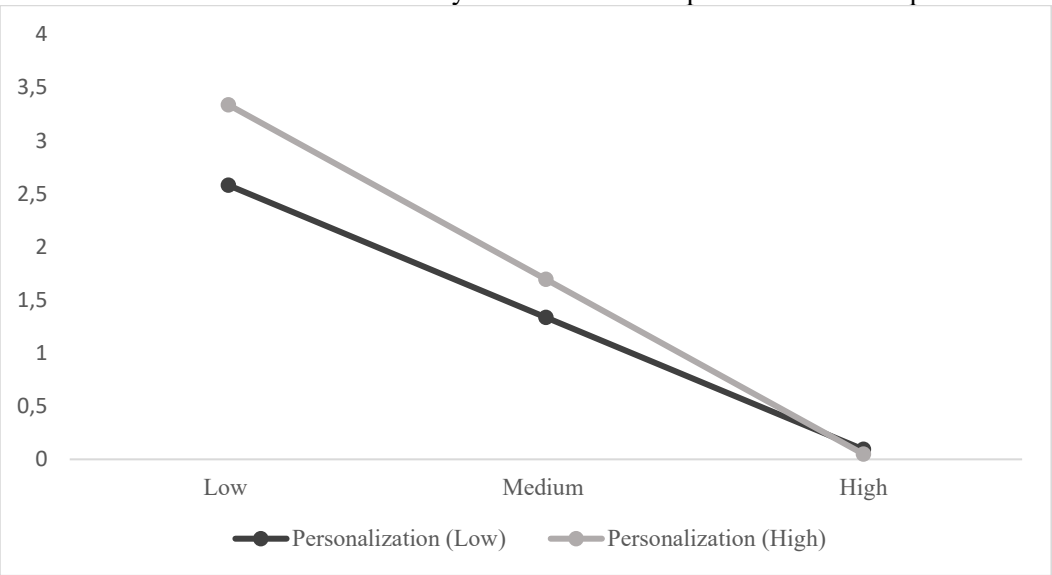
Conclusion: Hypothesis 3 has been rejected.

Table 7: Moderation analysis (PROCESS Model 1) predicting purchase intention from personalization, awareness of deepfakes, and their interaction

Predictor	B	SE	95% CI		t(127)	p
		HC3	LL	UL		
Constant	3.090	1.121	0.871	5.309	2.76	.007
Personalization (X)	1.357	1.208	- 1.034	3.748	1.12	.263
Awareness of Deepfakes (W)	- 0.422	0.171	- 0.760	- 0.084	- 2.47	.015
Personalization × Awareness (X×W)	- 0.200	0.219	- 0.635	0.233	- 0.92	.362
Brand Attitude (Covariate)	0.295	0.136	0.026	0.564	2.17	.032
Celebrity Endorsement (Covariate)	- 0.015	0.090	- 0.193	0.164	- 0.16	.872
Model Summary	R	R^2	Adjusted R^2	F	df	p
	.511	.262	.229	10.10	5,127	< .001
ΔR^2 for interaction term			R^2-change	F(HC3)	df	p
			.006	.838	1,127	.362

Note. *Dependent variable: Purchase Intention.*
Predictors: Personalization (X), Awareness of Deepfakes (W), and their interaction (X × W).
Covariates: Brand Attitude and Celebrity Endorsement.

Figure 10: Moderation of awareness of falsity on the interaction personalization on purchase intention



Note. *The Y-axis indicates the Purchase intention and on the X-Axis we have the awareness of falsity. The graph shows that the lines are almost parallel so there is no significant moderation of awareness of falsity on the interaction personalization and purchase intention.*

5.3.3 The effect of portrayal of realism and perceived likelihood of realism on awareness of falsity

Before testing Hypothesis 4 and 5 we did an explorative analysis to see whether the separation of perceived realism into two dimension is statistically justified. The exploratory factor analysis (principal component analysis with varimax rotation) using the five dimensions of perceived realism is visible in table a 32 in appendix 8. Following Kaiser's criterion[5] (Bortz & Schuster, 2011), we found that only two components had eigenvalues over 1, the third component had a variance of only 0.54 and was thus excluded. This indicates that perceived realism can indeed be grouped into two main dimensions instead of five.

Looking at the rotated component matrix in table a 32 in appendix 8, we found that plausibility, factuality, and typicality primarily load onto Factor 1, while perceptual quality and narrative consistency load onto Factor 2.

a. Hypothesis 4: Mediation analysis with portrayal of realism as mediator

In order to test Hypothesis 4 we performed a Mediation analysis through the PROCESS macro v5.0 (Hayes, 2022) using Model 4 and Bootstrapping 5000, represented in table 8. The total effect of awareness of deepfakes (X) and the dependent variable awareness of falsity (Y), represented in table 8, showed to be significant with $p = 0.032$ and explained 3.5% of the Variance.

[5] According to Kaiser's criterion the significant factors have a eigenvalue above 1 (Bortz & Schuster, 2011)

In Model 1, awareness of deepfakes significantly predicted portrayal of realism (M) (B = -0.49, p = .009).

In Model 2, where portrayal of realism was added as a mediator, the model as a whole was significant (p < .001) and explained 47.7% of the variance which is much higher than the R^2 from the first model. Portrayal of realism was a strong predictor of awareness of falsity (B = -0.74, p < .001). However, the direct effect of awareness of deepfakes on awareness of falsity was no longer significant (B = 0.08, p = .617). Since the direct effect is nonsignificant while the indirect effect is significant (B = 0.37, 95% CI [0.10, 0.65]), the confidence interval does not include zero, indicating a statistically significant mediation (Hayes, 2017) . We can thus conclude that there is a mediation of portrayal of realism. Awareness of deepfakes impacts portrayal of realism which in turns impacts awareness of falsity.

Conclusion: We therefore can support hypothesis 4.

Table 8: Regression analyses: testing the mediation pathways hypothesis 4 (PROCESS Model 4)

Model / Predictor	B	SE	95% CI		t(131)	p
			LL	UL		
Total Effect (X → Y)						
Awareness of deepfakes → Awareness of falsity	0.44	0.25	0.039	0.849	2.17	.032
Model summary: R = .19, R^2 = .035, F(1,131) = 4.70, p = .032						
Model 1 (X → M)						
Awareness of deepfakes → Portrayal of realism	- 0.493	0.186	- 0.861	- 0.126	- 2.66	.009
Model summary: R = .227, R^2 = .052, F(1,131) = 7.06, p = .009						
Model 2 (X + M → Y)						
Awareness of deepfakes → Awareness of falsity	0.079	0.157	- 0.232	0.389	0.50	.617
Portrayal of realism → Awareness of falsity	- 0.740	0.078	- 0.893	- 0.587	- 9.55	<.001
Model summary: R = .690, R^2 = .477, F(2,130) = 50.12, p < .001						

Note. *X = Awareness of Deepfakes (AD); M = Portrayal of Realism (PoR);Y = Awareness of Falsity (AF). All coefficients are unstandardized. Confidence intervals based on 5000 bootstrap samples. Indirect effect (a × b) =* 0.37*, 95% CI [*0.104, 0.655*].*

b. *Hypothesis 5: Mediation analysis with perceived likelihood of realism as mediator*

The total effect for the mediation analysis of hypothesis 5 is the same as for hypothesis 4, represented in table 08. The mediation specific results for hypothesis 5 are shown in table 09.

In Model 1 awareness of deepfakes doesn't not significantly impact perceived likelihood of realism (B = 0.17 p = .379).

However in model 2, both predictors, awareness of deepfakes (B = 0.52, p= .005) and perceived likelihood of realism (B=-0.47, p<.001) significantly predicted awareness of falsity. Thus both variables independently predict awareness of falsity.

The indirect effect is b = -0.08, of awareness of deepfakes on awareness of falsity through perceived likelihood of realism was not significant, since the confidence interval includes zero (95% CI [-0.28, 0.09], indicating that there could be a situation in which the effect of the mediator variable is null. Thus there is no mediation effect of perceived likelihood of realism.

Conclusion: Hypothesis 5 is rejected.

Table 9: Regression analyses: testing the mediation pathways hypothesis 5 (PROCESS Model 4)

Model / Predictor	B	SE	95% CI		t(131)	p
			LL	UL		
Model 1 (X → M)						
Awareness of deepfakes → Perceived likelihood of realism	0.172	0.195	-0.213	0.556	0.88	.379
Model summary: R = .078, R^2 = .006, F(1,131) = 0.78, p = .379						
Model 2 (X + M → Y)						
Awareness of deepfakes → Awareness of falsity	0.525	0.185	0.160	0.890	2.84	.005
Perceived likelihood of realism → Awareness of falsity	-0.471	0.094	-0.657	-0.284	-4.99	< .001
Model summary: R = .479, R^2 = .229, F(2,130) = 15.91, p < .001						

Note. *X = Awareness of Deepfakes (AD); M = Perceived Likelihood of Realism (PR); Y = Awareness of Falsity (AF). All coefficients are unstandardized. Confidence intervals based on 5000 bootstrap samples. Indirect effect (a × b) = -0.081, 95% CI [-0.281, 0.090].*

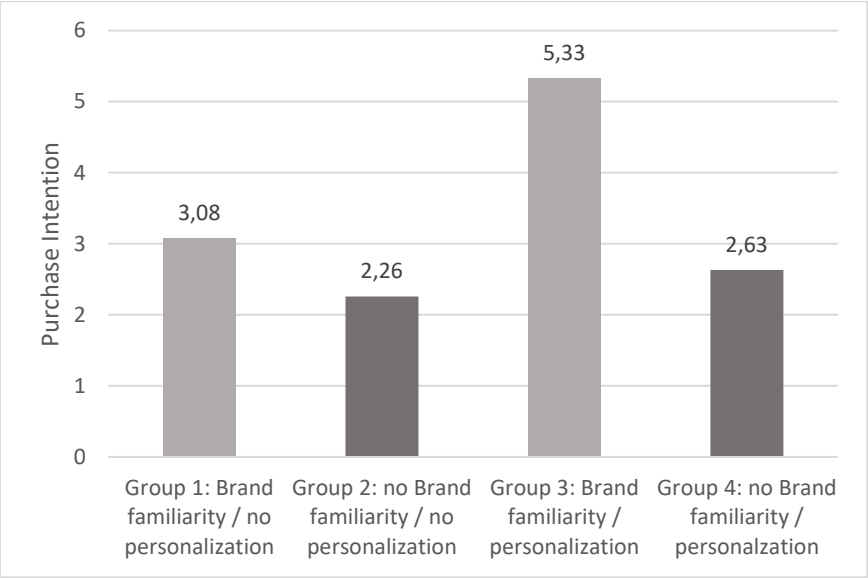
5.5 Explorative analysis

Based on the results of our main study we wanted to understand why personalization did not impact purchase intention.

In our experimental design, we chose to use an unknown brand (*Flower Knows*) to avoid any influence of brand familiarity on our construct (for example on perceived realism and awareness of falsity). However, as stated by Bleier and Eisenbeiss (2015), brand familiarity plays a crucial role in effectiveness in personalized ads: the authors found that personalized advertisements from known brands are more effective than those from unfamiliar ones. They explain that when respondents perceive an ad as personalized, their perceived intrusiveness and reactance towards the ad is reduced if the brand is known, since they recognize the advertiser's effort to tailor the message to their needs. When the marketer is unknown, this positive effect is less likely to occur. Viewers need to already have a certain level of trust in the brand to fully appreciate the personalization.

In the following we tried to take a graphical look at our data to see whether a trend can be identified.

Figure 11: Mean of the purchase intention between the groups having/having not brand familiarity and personalization

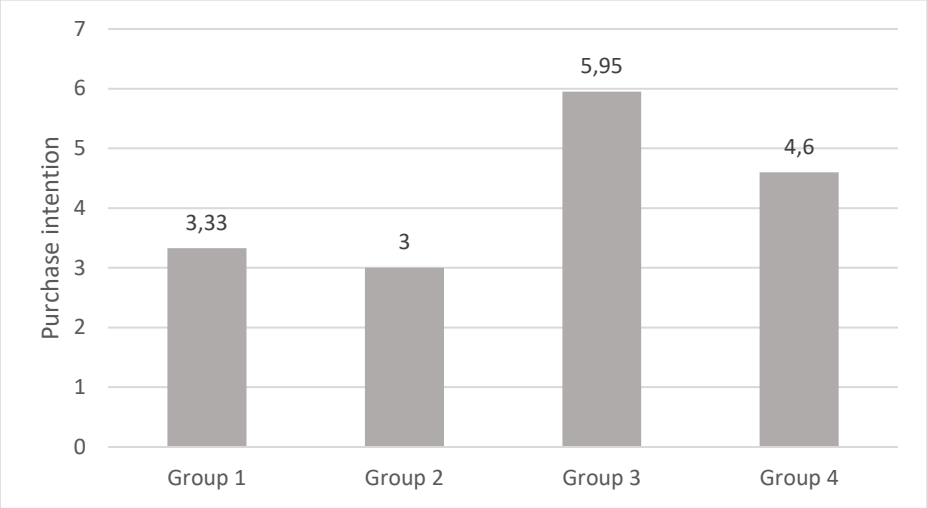


Note. *The figure shows the groups including Brand familiarity and personalization and their respective purchase intention. Group 1 N= 4; Group 2 N= 62; Group 3 N= 11; Group 4 N= 72*

We used the same sample however rebuilt our dataset: First we excluded respondents who failed the attention checks. Then we included those who indicated that they recognized the brand. An analyse of the standardized z-scores showed that all values were inside the acceptable range of ±3 (Osborne & Overbay, 2004) and thus that there are no extreme outliers.
Thus we had a total of 149 respondents, unfortunately only 15 of them were familiar with the brand. Due to this extremely uneven distribution (134 vs 15), it was not feasible to conduct a reliable statistical test of our hypotheses. Therefore, we decided to take a purely descriptive approach to the data analysis

In Figure 11, we observe that when brand familiarity is present, the personalization of advertisements appears to lead to a very high level of purchase intention (Group 3 > Group 1). When looking at the group without brand familiarity (Group 4), then the effect of personalization is very low. Group 4, including personalization is comparable to that of no personalization (Group 2) or brand familiarity alone (Group 1). This aligns with the findings of (Bleier & Eisenbeiss, 2015), who found that brand familiarity is a key factor for successful advertising personalization. However, due to the extremely unequal group sizes in our dataset, we cannot draw any definitive conclusions from these results.

Figure 12: Mean of the purchase intentions between the 4 experimental groups



Note. *This figure shows the means of purchase intentions between the 4 experimental groups only including the 15 respondents knowing the brand.*

Group 1: *no Awareness of Deepfakes / no Personalization N= 1;*

Group 2: *With awareness of Deepfakes / no Personalization N= 3;*

Group 3: *no Awareness of deepfakes / with Personalization N= 6*

Group 4: *with Awareness of Deepfakes / with personalization N= 5*

The Figure 12 shows that the highest purchase intention (M = 5.94) occurs when deepfakes are combined with personalization, group 3. However, as soon as participants are aware that deepfakes may have been used, the purchase intention drops significantly to M = 4.60, Group 4. In contrast, when no personalization is applied, the difference in purchase intention between respondents with and without deepfake awareness appears to be minimal (Group 1 and Group 2 of Figure 12).

This may suggest a potential interaction effect. Unfortunately we had very small number of respondents and in Group 1 only one respondent. Thus we are unable to conduct any meaningful statistical tests. As a result, these observations cannot provide clear or robust insights. We can only tentatively assume that brand familiarity may lead to an interaction effect between awareness and personalization.

VI. Discussion

In this work we wanted to understand how consumers' knowledge of deepfakes impacts purchase intention when marketers use deepfakes for hyper personalization. As stated previously research has until now focused on consumer reactions when confronted with disclosure of deepfakes but not on increased AI literacy, that is, increased awareness about the use and existence of deepfakes. Further the impact of personalization has rather focused on the use of generative AI but not on deepfakes, only a few studies have examined their impact; but did not sufficiently examine the interaction between awareness of deepfakes and personalization. With our experimental setting we tried to fill this gap and revealed some very interesting insights as that awareness of deepfakes negatively affected portrayal of realism and that personalization did not interact with awareness of deepfakes. In this part we will further discuss the found results.

6.1 Major results and theoretical implications

6.1.1 Extending the Campbell et al. (2021) framework to deepfake advertising

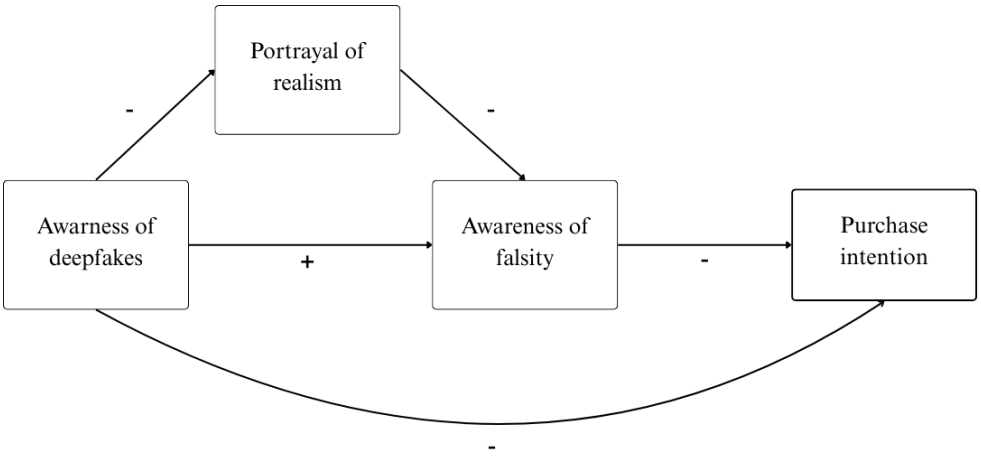
We were able to develop further the framework proposed by Campbell et al. (2021). The authors proposed the framework for different manipulation techniques in advertising. They proposed that the higher the manipulation techniques used the higher the perceived realism of the ad, which in turn reduces awareness of falsity and increases purchase intention. This work could confirm that with deepfake ads this framework is applicable, since we were able to support all the three assumptions (AA, AB, AC).

As shown in figure 13, we further added the concept of AI literacy through awareness of deepfakes to the framework. Manipulation sophistication being an external variable, that the advertiser can control himself, we decided to add awareness of deepfakes as an internal variable that consumers already bring with them. Adding this variable brings interesting insights, since people use previously acquired knowledge in order to evaluate the persuasion attempt. Not just the higher manipulation will impact perceived realism but also the media literacy and its further developed AI literacy is partially impacting this variable and thus needed to be added in this concept. We showed that awareness of deepfakes can be a part in this model and impacts its structure:

We were able to find that awareness of deepfakes is decreasing purchase intention as expected (H2b). These findings are in line with previous literature, in the sense that recognizing the ad manipulation with deepfake technology leads to lower persuasion outcome, as has been suggested by Campbell et al. (2021), empirically tested by Sivathanu and Pillai (2022) with the increase of perceived deception, or by Karpinska-Krakowiak and Eisend (2024) or Powers et al. (2023) with the use of disclosure. But we have brought more detailed understanding on how awareness of deepfakes impacted purchase intention and what is the mechanism behind. While in the model from Campbell et al. (2021) perceived realism has been showed as one full variable, we were able to show that it is more relevant to separate this variable into two dimensions.

Figure 13 shows the final model that has been confirmed through our hypothesis testing: it shows that portrayal of realism is a mediator between awareness of deepfakes and awareness of falsity. It also shows the direct effect we found between awareness of falsity on purchase intention and the direct effect between awareness of deepfakes on purchase intention.

Figure 13: Final model for deepfake advertisement with an unknown brand.



Note. *This figure shows the final model found in this work. It shows the mediation of portrayal of realism between awareness of deepfakes and awareness of falsity.*

6.1.2 Influence of awareness of deepfakes on purchase intention

Awareness of deepfakes is reducing portrayal of realism, which in turn increases awareness of falsity and decreases purchase intention. These results were somewhat surprising since we proposed that awareness of deepfakes was able to reduce the whole concept of perceived realism (H1) but found that awareness was only able to reduce one aspect of the concept: portrayal of realism (H4).

Perceived realism has 5 dimensions as proposed by Hall (2003): factuality, typicality, plausibility, narrative consistency and perceptual quality. In the context of traditional advertisement, it has been shown that media literacy can decrease perceived realism (Austin & Johnson,1997), or certain aspects just like typicality (Scull et al., 2014)[6]. While Austin and Johnson (1997) tested the full concept of perceived realism, our work on the other hand did not confirm this; hypothesis H1, stating that awareness of deepfakes will decrease perceived realism; could not be supported. Austin and Johnson (1997) did use classical TV-advertisement for alcohol, while our study used Deepfake technology. Also the work from these authors did include an extensive media literacy training, while we only addressed

[6] Scull et al. (2014) claims to have tested *perceived realism* but a deeper investigation of his questionnaire shows that he tested only *typicality*.

simple awareness about the technology. Therefore either the use of higher manipulation techniques like AI is impacting differently perceived realism or awareness of deepfakes are not enough and extensive media literacy intervention is needed to reduce perceived realism of the ad.

We later found that awareness of deepfakes only impacted significantly two of the five dimensions of perceived realism: perceptual quality and narrative consistency (H1d and H1c). Based on Karpinska-Krakowiak and Eisend (2024), who did separate the concept of perceived realism into two dimensions: portrayal of realism and factuality, and after performing a factor analysis (table a 35 appendix 8) we suggested that perceived realism can rather be separated in two different dimensions: perceived likelihood of realism and portrayal of realism. Thus we performed the mediation analysis of the two new variables and found that only portrayal of realism is a mediator in this concept (H4). Past research about the use of disclosure in the context of deepfakes ad, Karpinska-Krakowiak and Eisend (2024) showed that only one dimension of perceived realism is decreased: the factuality; but not significantly portrayal of realism. Our work identified that with increased awareness of deepfake the results are completely different than with disclosure. The increased awareness of deepfakes only significantly reduced the consumers portrayal of realism including narrative consistency and perceived quality, instead of perceived likelihood of realism, which included the dimension factuality, typicality and plausibility. But with awareness of deepfakes neither the concept of factuality nor the concept of typicality is reduced.

6.1.3 Disclosure vs. Awareness: Two distinct pathways through perceived realism

It therefore seems that disclosure and awareness of deepfakes have different impacts on perceived realism. Factuality is knowing whether the ad is real or not and disclosure is a clear statement, thus factuality is reduced with disclosed AI ads. Therefore the fact that factuality has not been reduced compared to the experiments done by Karpinska-Krakowiak and Eisend (2024) makes sense, since without disclosure the respondents cannot be sure that the ad is or is not AI. Thus factuality might not been impacted in our experimental setting. However, Karpinska-Krakowiak and Eisend (2024) said that well done AI deepfakes cannot be recognized as such and therefore portrayal is not reduced even when the deepfake is disclosed. But in this work, portrayal of realism has been reduced. We therefore assume that awareness of deepfakes acts differently than disclosure. People don't know if the video is fake or not, thus following the persuasion knowledge model (Friestad & Wright, 1994) people will recall on knowledge from the past when confronted with AI ads, that could contribute to the development of persuasion knowledge about deepfake manipulation. Consumers then might learn to associate such cues with potential persuasion attempts and learn how to recognize deepfakes they may look at details like mouth movements. AI-generated media can contain visual inconsistencies, such as unrealistic mouth movements, and previous research has shown that viewers react sensitively to such inconsistencies (Kaate et al., 2023). When later confronted with new deepfake advertisements, this knowledge may be

activated, leading participants to evaluate perceptual quality and narrative consistency more critically in order to identify possible manipulation. Once the persuasion attempt is recognized, consumers then employ coping strategies, which may ultimately reduce purchase intention, explaining that hypothesis H2b has been supported. One other reason for these results could be that participants in our study are already quite aware of AI in advertising,[7] and in consequence already have more experience in analyzing AI content. Another reason could be that the visual aspect in this work could have been too low, people recognizing the fake easier.

6.1.4 *Personalization* has no influence on *purchase intention*

The second finding of this study is that *personalization* did not influence *purchase intention*, which is in contradiction to our initial expectations; since hypothesis H2a, the negative relationship between personalization increase and purchase intention could not be supported.

We assumed that, in the case of a well-known brand, our test subjects might answer the questionnaire with a bias, thereby distorting our measurements. Therefore, in our work we chose to use an unknown Chinese brand. However, Bleier and Eisenbeiss (2015) showed in their work that trust in the brand is important for successful personalization. Other studies, including those using generative AI, have shown that personalization increases persuasiveness (Guo & Jiang, 2023; Kumar & Kapoor, 2024). But all these studies used well-known brands such as BMW and Mazda (Guo & Jiang, 2023) and a well-known Indian retailer for Kumar and Kapoor (2024). Therefore the lack of correlation between personalization and purchase intention in our work can be explained by the low brand familiarity among our test group, where none of our participants knew the brand.

Nevertheless, in our exploratory study, we found an indication, that in cases where respondents were familiar with the brand, personalization increased purchase intention and that there might be an interaction effect with awareness of deepfakes. The explorative analysis showed that the positive effects of personalization on purchase intention might get mitigated with higher awareness of deepfakes, since people will use coping strategies after recognizing the manipulation strategies of hyper personalization from the advertiser (Friestad & Wright, 1994). It would be important to further investigate this issue also because our dataset included only 15 respondents who are familiar with the brand, thus we could not draw any clear conclusions.

6.2 Research questions answers

This work stated an important research questions: *How does consumer awareness of deepfake-powered hyper-personalization influence purchase intention?* This question did address two issues, first how AI

[7] In our manipulation check, the measured AI literacy of both groups was at a higher level as was visible in table a 28 of the appendix 5.

literacy, awareness of deepfakes, impacted purchase intention, then how it did interact with hyper personalization.

We were able to give a clear answer on the first aspect of this question: *Consumer knowledge of deepfake-powered advertisement will make viewers take closer attention on the advertisement visual quality and its narrative consistency. This visual analysis of the viewer will increase his awareness of falsity and in turn reduce his purchase intention.* We were able to not just show that awareness of deepfakes has a negative impact on purchase intention but also show details behind the mechanism. This experiment was able to give researchers a more detailed framework that can be used to further develop the understanding of how awareness of deepfakes, is impacting purchase intention. Surprisingly we were able to show that awareness of deepfakes will be impacting the visual aspect of an ad and not the other aspects of perceived realism that have been impacted by media literacy when confronted with traditional media sophistications. Deepfakes seem to have also a different impact than traditional medias as the TV advertisements used by Austin and Johnson (1997), that impacted the full concept of perceived realism or looking at the experiment from where Scull et al. (2014) media literacy intervention impacted typicality. Unfortunately this work only focused on awareness of deepfakes and not on the whole concept of AI literacy, therefore this answer cannot be given for AI literacy itself but only for the case of awareness about the existence and use of deepfakes.

Looking at the second aspect we were not able to give a clear answer. Personalisation did not impact purchase intention, because in the context of our experiment we chose to use an unknown brand. This work therefore shows that deepfake powered hyper personalization is connected to a boundary condition which is the brand knowledge which is in line with the results from Bleier and Eisenbeiss (2015). Therefore we can say: *Deepfake powered hyper personalization is not effective when used for unknown brands.* Our explorative analysis however hinted towards an interaction effect between awareness of deepfakes and personalization when the brand is known.

6.3 Managerial implications

This work showed that the effectiveness of persuasion from deepfake ads can be influenced by the viewers awareness and by extend their AI literacy. More and more people will probably understand what deepfakes are, how to recognize them, how they are used and this will impact their purchase intention. Consumers will be more careful and looking for hints in the ad indicating the use of Deepfakes. The use of deepfakes may decrease purchase intention. Advertisers therefore need to be aware that as consumers become more familiar with the existence and use of AI, its application in advertising may increasingly risk reducing purchase intention.

However, as awareness of deepfakes only seems to impact portrayal of realism and simple disclosure as stated by Karpinska-Krakowiak and Eisend (2024) is not enough to decrease factuality. We might want

to suggest that advertiser to only use very high quality deepfakes. Based on these insights, very high-quality deepfakes may remain effective, increase purchase intention, even when simple disclosure is applied. However, future research needs to identify what happens if a consumer with higher awareness of deepfakes is confronted with simple disclosure. For advertisers it could be important to know whether in this case purchase intention is also decreased or if as long the portrayal of deepfake is high the simple disclosure is still not impacting the two dimensions of perceived realism.

Further besides the visual aspect advertisers should not forget about the narrative consistency, the story line of the ad needs to be consistent if using deepfakes, since it is also one of the main aspects that is reduced with higher awareness of deepfakes.

Further concerning hyper personalization, this work shows that using deepfakes to achieve hyper personalized advertisements for an unknown brand is unnecessary effort, it has the same effect than a non-personalized advertisement. Also the awareness of deepfakes will decrease purchase intention, therefore we cannot recommend advertisers to use deepfakes if advertiser only use it in order to achieve hyper personalization. When the brand is rather known we would therefore suggest that hyper personalization with deepfake might be a positive use case, also since it can reduce production costs. However further research is definitely needed in order to confirm this suggestion.

6.4 Limitations

Firstly the convenience sampling methods is problematic, it does not enable us to make general assumption about the population since the participants are mostly young students from the business environment. In reaching respondents through personal contacts and using websites like surveyswap.com which is targeting mainly student that might be more invested in AI and may have higher AI literacy. We therefore cannot make any clear assumptions about the whole population. Another study, that has access to probabilistic sampling methods, would be necessary to confirm if the findings here can be applied on the whole population.

Further, this study was not able to test how personalization and awareness of deepfakes would have interacted if the brand was known. The respondents used for the explorative analysis that knew the brand are too few to get any valuable insights in this setup. Therefore the dataset wasn't good to give insights on what would have happened if the brand was known. A new dataset would have been needed. Also, we were not able to recruit respondents that do enjoy these kind of product styles (fairy, pink makeup). In the manipulation check we could see that the perceived personalization did work, but in reality the people were not much attracted to the product itself (see appendix 5).

Also the deepfake quality was relatively low, the use of better deepfakes would have been closer to real advertising. Since marketers have more technological abilities to make extremely realistic deepfakes than us. The low quality of the deepfake makes it easier for the respondents to find signs indicating the

use of deepfakes. Then we need to mention that the experimental design might push the respondents to attentively watch the ad. Leading into respondents taking their time to analyse the video more thoroughly. In a real life context respondents might not take as much attention watching the ad than in the experimental setting.

Lastly we only manipulated the overall awareness of deepfakes because a concrete AI literacy intervention was not feasible with the available resources. Even though the manipulation check showed that we did successfully manipulated AI literacy, but it also showed that some respondents found the awareness text to be close to a disclosure. Thus a literacy intervention as it has been done in the work of Austin et al. (2007), Austin and Johnson (1997) and Scull et al. (2014), would have allowed us to test AI literacy as a broader construct and might also have avoided unintentionally directing respondents' attention to the AI-related aspects of the videos.

6.5 Future research

Firstly future research should test how personalization interacts with awareness of deepfakes in a context of a known brand. We believe that our observations in the explorative study might be confirmed, meaning that there is an interaction effect between awareness of deepfakes and personalization when a known brand is used.

The next step would be to see if AI literacy has a similar impact than awareness of deepfakes, through a real literacy intervention with several teachings and classes. Also interesting would be to further investigate whether there is an actual difference between media literacy in the traditional sense and AI literacy when considering perceived realism. Awareness of deepfakes being a part of AI literacy: thus our work suggests that AI literacy impacts rather portrayal of realism while it seems that media literacy as past research suggests impacts the whole concept of perceived realism. This needs to be further investigated.

Third, it would be interesting to know if respondents with high awareness will still try to find inconsistencies in deepfake in ads of extremely high quality. With almost undetectable AI it would have made sense that people take closer attention on plausibility or typicality (on the actors way of dressing or acting) than portrayal of realism. The knowledge people have on how to recognize AI would probably not be working with extremely realistic deepfakes. Therefore the question that arises is what if the ad would have better quality, unrecognizable AI? Then, when consumers can assume that deepfakes are mostly used in advertisement how will they react in the future? Will people then confuse real content with AI content? Currently purchase intention is reduced when consumers believe the ad is false, but will this keep on in the future when consumers need to assume that almost all advertisers use AI advertisements?

Lastly we want to look at the question that arises when looking at the interaction between AI literacy and disclosure: As Karpinska-Krakowiak and Eisend (2024) has shown only extended disclosure has significant impact on factuality, but what if the respondents have higher AI literacy? In Europe the AI Act requires advertisers to disclose ads that have used AI generated or AI manipulated content (Tagesschau, 2023)**.** But as shown in their study, simple disclosure of AI is not enough to decrease factuality (Karpinska-Krakowiak & Eisend, 2024). Simple disclosure that AI has been used, doesn't indicate what type of AI has been used, Deepfake, full generative AI, was the star really in the setup or only the background AI generated? In the next step it would be very interesting to see how AI literacy and simple disclosure interacts. Could simple AI disclosure be enough when the awareness of AI is high?

VII. Conclusion

This work aimed at understanding how awareness of deepfakes impacts purchase intention when advertisers use deepfake powered hyper personalized ads.

Firstly this work brought new relevant theoretical implications by further developing the framework from Campbell et al. (2021). On one side we successfully tested the model for the manipulation sophistication of deepfake and we also added a consumer side variable awareness of deepfakes, which is part of AI literacy. This work therefore showed that persuasion outcomes are not only impacted by the manipulation sophistication itself but also by the consumer knowledge about this manipulation, in this case their knowledge about deepfakes.

We were able to show that even if viewers have already a high level of awareness about deepfakes a little more increase of awareness of deepfakes could still impact negatively purchase intention. Central results of our work further included that awareness of deepfakes and disclosure impact perceived realism differently. An increase in awareness of deepfakes does not influence perceived realism, but it does influence the portrayal of realism, the critical examination of small errors in the presentation of the product and thus leads to a reduction in the purchase decision. From this it follows that successful advertising can certainly be based on deepfakes, but it requires perfect presentation and well structured narration.

This work enabled us to better understand how deepfake knowledge impacts purchase intention and we thus brought more details to the research results from Selent (2022), Sivathanu and Pillai (2022) and further developed the framework proposed by Campbell et al. (2021). We identified that the concept perceived realism can be separated into two main dimensions portrayal of realism and perceived likelihood of realism. While in this work awareness of deepfakes has shown to impact portrayal of realism, past literature has shown that disclosure rather impacts certain dimensions from perceived likelihood of realism (factuality) (Karpinska-Krakowiak & Eisend, 2024).

These findings show that growing knowledge about deepfakes in the population can result in lower persuasion outcomes when marketers use deepfakes. Questions arise how marketers could avoid this? What if consumers tend to believe deepfakes are used even though the video is real, since the technology becomes better and better and it is nearly impossible to identify deepfake ? Would the respondents then associate negative response to non-deepfake videos too?

On the other side we tried to answer the question of hyper personalization, however we could not find any significant impact of personalized advertisement on the final purchase intention. We found that this is due to the use of an unknown brand. Therefore this work found that hyper personalized deepfakes is bounded to the consumers connection with the brand. Brands that are unknown to the consumer are less trust worthy and thus consumers react negatively to personalization as stated by Bleier and Eisenbeiss

(2015). This work confirmed that for unknown brands the use of deepfakes for hyper personalization results in no better purchase intentions and thus it seems an unnecessary effort for marketers. However our exploratory analysis showed that for known brands hyper personalization is successful, but there might also be an interaction effect between awareness of deepfakes and personalization that will mitigate the positive effect. Nevertheless the exploratory analysis also hinted that even though the effect is mitigated it might still be better than no personalization. The next step would be to perform an experiment with a known brand and enough respondents to see what happens if the brand is known and if the proposed model is applicable in a situation with high brand familiarity.

In the end using deepfakes will still be interesting for marketers, since it reduces production costs and enable hyper-personalization. Advertisers need to keep in mind that awareness of deepfakes will increase in the next years and therefore the persuasion outcome of deepfake ads will be reduced. High quality visual deepfake ads will be necessary and hyper personalization deepfake ads can be beneficial even though their effect might be mitigated with awareness. This work shows how consumers think now but with AI, the fast technological evolution, the increasing awareness about this technologies, reactions towards AI advertisement might change over time. Today, consumers may react strongly negatively to deepfake advertisements, but as such ads become more common, consumers may gradually become accustomed to them, which may diminish these negative reactions. In any case better understanding the framework behind consumer reaction, as this work was able to bring forward, will be beneficial for future research.

Bibliography

Aljarah, A., Ibrahim, B., & López, M. (2024). In AI, we do not trust! The nexus between awareness of falsity in AI-generated CSR ads and online brand engagement. *Internet Research*, *35*(3), 1406–1426. https://doi.org/10.1108/INTR-12-2023-1156

Arachchi, H. A. D. M., & Samarasinghe, G. D. (2024). Impact of Deepfake Advertising Attributes on Consumers' Hedonic & Utilitarian Values and Brand Credibility. *2024 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, 1–6. https://doi.org/10.1109/SCSE61872.2024.10550666

Arnett, J. (2007). Message Interpretation Process Model. In *Encyclopedia of Children, Adolescents, and the Media* (Vols. 1–2, pp. 536–536). SAGE Publications, Inc. https://doi.org/10.4135/9781412952606.n277

Ashley, S., Maksl, A., & Craft, S. (2013). Developing a News Media Literacy Scale. *Journalism & Mass Communication Educator*, *68*(1), 7–21. https://doi.org/10.1177/1077695812469802

Aufderheide, P. (1993). *Media literacy: A report of the National Leadership Conference on Media Literacy, The Aspen Institute Wye Center, Queenstown, Maryland, December 7 - 9, 1992*. Aspen Inst.

Austin, E. W., & Johnson, K. K. (1997). Immediate and Delayed Effects of Media Literacy Training on Third Grader's Decision Making for Alcohol. *Health Communication*, *9*(4), 323–349. https://doi.org/10.1207/s15327027hc0904_3

Austin, E. W., & Meili, H. K. (1994). Effects of interpretations of televised alcohol portrayals on children's alcohol beliefs. *Journal of Broadcasting & Electronic Media*, *38*(4), 417–435. https://doi.org/10.1080/08838159409364276

Austin, E. W., Pinkleton, B. E., & Funabiki, R. P. (2007). The Desirability Paradox in the Effects of Media Literacy Training. *Communication Research*, *34*(5), 483–506. https://doi.org/10.1177/0093650207305233

Baek, T. H., & Morimoto, M. (2012). Stay Away From Me. *Journal of Advertising*, *41*(1), 59–76. https://doi.org/10.2753/JOA0091-3367410105

Bang, H., & Wojdynski, B. W. (2016). Tracking users' visual attention and responses to personalized advertising based on task cognitive demand. *Computers in Human Behavior*, *55*, 867–876. https://doi.org/10.1016/j.chb.2015.10.025

Belch, E. G., & Belch, A. M. (2003). *Advertising and Promotion: An Integrated Marketing Communications Perspective*. McGraw-Hill/Irwin. https://books.google.fr/books?id=nLTOAQAACAAJ

Bleier, A., & Eisenbeiss, M. (2015). The Importance of Trust for Personalized Online Advertising. *Journal of Retailing*, *91*(3), 390–409. https://doi.org/10.1016/j.jretai.2015.04.001

Bortz, J., & Schuster, C. (2011). *Statistik für Human- und Sozialwissenschaftler: Limitierte Sonderausgabe* (7. Aufl. 2010). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12770-0

Brehm, J. W. (1989). *Psychological Reactance: Theory and Applications*. *6*.

Bruner II, G. C. (2019). *Marketing Scales Handbook: Multi-Item Measures for Consumer Insight Research, Volume 10*. GCBII Productions, LLC. http://ebookcentral.proquest.com/lib/univie/detail.action?docID=5731640

BuzzFeed Video (Director). (2018, April 17). *You won't believe what Obama says in this video!* [Video recording]. https://www.youtube.com/watch?v=cQ54GDm1eL0

Cadbury Celebrations (Director). (2021, October 23). *Supporting Local Retailers This Diwali | Not Just A Cadbury Ad Campaign Video* [Video recording]. https://www.youtube.com/watch?v=5WECsbqAQSk&t=14s

Campbell, C., Plangger, K., Sands, S., & Kietzmann, J. (2021). Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising. *Journal of Advertising*, *51*(1), 22–38. https://doi.org/10.1080/00913367.2021.1909515

Campbell, C., Plangger, K., Sands, S., Kietzmann, J., & Bates, K. (2022). How Deepfakes and Artificial Intelligence Could Reshape the Advertising Industry: The Coming Reality of AI Fakes and Their Potential Impact on Consumer Behavior. *Journal of Advertising Research*, *62*(3), 241–251. https://doi.org/10.2501/JAR-2022-017

Chitrakorn, K. (2022). Chapter 9.2: How Deepfake could change fashion advertisiing. In M. Kirsten (Ed.), *Ethics of Data and Analytics: Concepts and Cases* (1st ed., pp. 372–374). Auerbach Publications. https://doi.org/10.1201/9781003278290

Cho, H., Shen, L., & Wilson, K. (2012). Perceived Realism: Dimensions and Roles in Narrative Persuasion. *Communication Research*, *41*(6), 828–851. https://doi.org/10.1177/0093650212450585

Cochran, J. D., & Napshin, S. A. (2021). Deepfakes: Awareness, Concerns, and Platform Accountability. *Cyberpsychology, Behavior, and Social Networking*, *24*(3), 164–172. https://doi.org/10.1089/cyber.2020.0100

Davenport, T. H. (2023). Hyper-Personalization for Customer Engagement with Artificial Intelligence. *Management and Business Review*, *3*(1–2), 29–36. https://doi.org/10.1177/2694105820230301006

De Keyzer, F., Dens, N., & De Pelsmacker, P. (2022). Let's get personal: Which elements elicit perceived personalization in social media advertising? *Electronic Commerce Research and Applications*, *55*, 101183. https://doi.org/10.1016/j.elerap.2022.101183

DeVellis, R. F., & Thorpe, C. T. (2022). *Scale development: Theory and applications* (Fifth edition). SAGE.

EDEKA. (2025, May 25). *It's big brain energy* [Social Media]. Instagram. https://www.instagram.com/reel/DKEa1TMsBhR/

Eisend, M., Plagemann, J., & Sollwedel, J. (2014). Gender Roles and Humor in Advertising: The Occurrence of Stereotyping in Humorous and Nonhumorous Advertising and Its Consequences for Advertising Effectiveness. *Journal of Advertising*, *43*(3), 256–273. https://doi.org/10.1080/00913367.2013.857621

Fomenko, Y. (2024, March 11). *It's my birthday!* [Social media video]. https://www.instagram.com/reel/DB6SEp_xzOw/?igsh=aG16YnI1czFzcGJm

Fransen, M. L., Smit, E. G., & Verlegh, P. W. J. (2015). Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01201

Friestad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research*, *21*(1), 1–31. https://doi.org/10.1086/209380

Guo, B., & Jiang, Z. (2023). Influence of personalised advertising copy on consumer engagement: A field experiment approach. *Electronic Commerce Research*, *25*(2), 1281–1310. https://doi.org/10.1007/s10660-023-09721-5

Hall, A. (2003). Reading Realism: Audiences' Evaluations of the Reality of Media Texts. *Journal of Communication*, *53*(4), 624–641. https://doi.org/10.1111/j.1460-2466.2003.tb02914.x

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. Guilford Publications. http://ebookcentral.proquest.com/lib/univie/detail.action?docID=5109647

Holdsworth, J., & Scapicchio, M. (2024, June 17). *What is deep learning?* IBM Think. https://www.ibm.com/think/topics/deep-learning

Kaate, I., Salminen, J., Santos, J., Jung, S.-G., Olkkonen, R., & Jansen, B. (2023). The realness of fakes: Primary evidence of the effect of deepfake personas on user perceptions in a design task. *International Journal of Human-Computer Studies*, *178*, 103096. https://doi.org/10.1016/j.ijhcs.2023.103096

Karagiannis, D., & Telesko, R. (2001). *Wissensmanagement: Konzepte der Künstlichen Intelligenz und des Softcomputing*. Oldenbourg.

Karpinska-Krakowiak, M. (Mags), & Eisend, M. (2024). Realistic Portrayals of Untrue Information: The Effects of Deepfaked Ads and Different Types of Disclosures. *Journal of Advertising*, *54*(3), 432–442. https://doi.org/10.1080/00913367.2024.2306415

Kharvi, P. L. (2024). Understanding the Impact of AI-Generated Deepfakes on Public Opinion, Political Discourse, and Personal Security in Social Media. *IEEE Security & Privacy*, *22*(4), 115–122. https://doi.org/10.1109/MSEC.2024.3405963

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, *63*(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Kietzmann, J., Mills, A. J., & Plangger, K. (2021). Deepfakes: Perspectives on the future "reality" of advertising and branding. *International Journal of Advertising*, *40*(3), 473–485. https://doi.org/10.1080/02650487.2020.1834211

Kim, B. K., Choi, J., & Wakslak, C. J. (2019). The Image Realism Effect: The Effect of Unrealistic Product Images in Advertising. *Journal of Advertising*, *48*(3), 251–270. https://doi.org/10.1080/00913367.2019.1597787

Kleine, F. (2022). *Perception of Deepfake Technology—The Influence of the Recipients' Affinity for Technology on the Perception of Deepfakes* [Masterthesis, Darmstadt University of Applied Sciences]. https://www.researchgate.net/publication/364254498_Perception_of_Deepfake_Technology_-_The_Influence_of_the_Recipients'_Affinity_for_Technology_on_the_Perception_of_Deepfakes

Kumar, M., & Kapoor, A. (2024). Generative AI and Personalized Video Advertisements. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4614118

Kwok, A. O. J., & Koh, S. G. M. (2021). Deepfake: A social construction of technology perspective. *Current Issues in Tourism*, *24*(13), 1798–1802. https://doi.org/10.1080/13683500.2020.1738357

Kyoorius. (2021). *Best Of Show- Black Elefant*. Kyooriuscreative2025. https://kyooriuscreative2025.awardsengine.com/?action=ows:entries.details&e=64203&project_year=2021

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727

MacKenzie, S. B., & Lutz, R. J. (1989). An Empirical Examination of the Structural Antecedents of Attitude toward the Ad in an Advertising Pretesting Context. *Journal of Marketing*, *53*(2), 48–65. https://doi.org/10.1177/002224298905300204

Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. *Scientific Reports*, *14*(1), 4692. https://doi.org/10.1038/s41598-024-53755-0

Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K. (2023). Deepfakes: Deceptions, mitigations, and opportunities. *Journal of Business Research*, *154*, 113368. https://doi.org/10.1016/j.jbusres.2022.113368

Nan, L. (2025, September 7). *Is C-beauty having a K-beauty moment?* JingDaily. https://jingdaily.com/posts/flower-knows-judydoll-is-c-beauty-having-a-k-beauty-moment

Nguyen, T. T., Nguyen, Q. V. H., Nguyen, D. T., Nguyen, D. T., Huynh-The, T., Nahavandi, S., Nguyen, T. T., Pham, Q.-V., & Nguyen, C. M. (2022). Deep learning for deepfakes creation and detection: A survey. *Computer Vision and Image Understanding*, *223*, 103525. https://doi.org/10.1016/j.cviu.2022.103525

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Practical Assessment, Research, and Evaluation*, *9*(1). https://doi.org/10.7275/qf69-7k43

Peter, J. P., & Olson, J. C. (2010). *Consumer behavior & marketing strategy* (9. ed). McGraw-Hill Irwin.

Pfiffelmann, J. (n.d.). *Repertory of Marketing Scales*. Marketing Scales. Retrieved September 23, 2025, from https://www.jean-pfiffelmann.com/marketing-scales/

Plohl, N., Mlakar, I., Aquilino, L., Bisconti, P., & Smrke, U. (2024). Development and Validation of the Perceived Deepfake Trustworthiness Questionnaire (PDTQ) in Three Languages. *International Journal of Human–Computer Interaction*, *41*(11), 6786–6803. https://doi.org/10.1080/10447318.2024.2384821

Powers, G., Johnson, J. P., & Killian, G. (2023). To Tell or Not to Tell: The Effects of Disclosing Deepfake Video on US and Indian Consumers' Purchase Intention. *Journal of Interactive Advertising*, *23*(4), 339–355. https://doi.org/10.1080/15252019.2023.2260399

Ramos-Clemente, M. (2022, February 17). *Did You Know? Anna Delvey's Court Looks in Netflix's "Inventing Anna" Are Unbelievably Accurate*. Preview Fashion.

https://www.preview.ph/fashion/netflix-inventing-anna-delvey-sorokin-court-looks-a00318-20220217?s=32lt2t2o5cn5ulksusdmt17u8g

Scull, T., Malik, C., & Kupersmidt, J. (2014). A Media Literacy Education Approach to Teaching Adolescents Comprehensive Sexual Health Education. *Journal of Media Literacy Education*. https://doi.org/10.23860/jmle-6-1-1

Selent, P. (2022). *Deepfakes im Marketing: Eine Analyse von Deepfake-Marketing mit exemplarischer Ausarbeitung eines eigenen Deepfake-Modells zur Überprüfung von Effekten bei Rezipienten* [Application/pdf]. 14122 KB, 109 pages. https://doi.org/10.25933/OPUS4-2839

Semeradova, T., & Weinlich, P. (2019). Computer Estimation of Customer Similarity With Facebook Lookalikes: Advantages and Disadvantages of Hyper-Targeting. *IEEE Access*, *7*, 153365–153377. https://doi.org/10.1109/ACCESS.2019.2948401

Seow, J. W., Lim, M. K., Phan, R. C. W., & Liu, J. K. (2022). A comprehensive overview of Deepfake: Generation, detection, datasets, and opportunities. *Neurocomputing*, *513*, 351–371. https://doi.org/10.1016/j.neucom.2022.09.135

Shanahan, T., Tran, T. P., & Taylor, E. C. (2019). Getting to know you: Social media personalization as a means of enhancing brand loyalty and perceived quality. *Journal of Retailing and Consumer Services*, *47*, 57–65. https://doi.org/10.1016/j.jretconser.2018.10.007

Similarweb. (2025, May). *Flowerknows.co Website Analysis for May 2025*. Similarweb. https://www.similarweb.com/website/flowerknows.co/#geography

Sivathanu, B., & Pillai, R. (2022). The effect of deepfake video advertisements on the hotel booking intention of tourists. *Journal of Hospitality and Tourism Insights*, *6*(5), 1669–1687. https://doi.org/10.1108/JHTI-03-2022-0094

Solomon, M. R. (2017). *Consumer behavior: Buying, having, and being* (Twelfth Edition). Pearson.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th edition, Pearson new international edition). Pearson.

Tagesschau. (2023, September 12). *Einigung zum "AI Act": EU will KI-Anwendungen streng regulieren*. tagesschau.de. https://www.tagesschau.de/ausland/europa/eu-kuenstliche-intelligenz-regulierung-100.html

Vasist, P. N., & Krishnan, S. (2022). Engaging with deepfakes: A meta-synthesis from the perspective of social shaping of technology theory. *Internet Research*, *33*(5), 1670–1726. https://doi.org/10.1108/INTR-06-2022-0465

White, T. B., Zahay, D. L., Thorbjørnsen, H., & Shavitt, S. (2007). Getting too personal: Reactance to highly personalized email solicitations. *Marketing Letters*, *19*(1), 39–50. https://doi.org/10.1007/s11002-007-9027-9

Whittaker, L., Letheren, K., & Mulcahy, R. (2021). The Rise of Deepfakes: A Conceptual Framework and Research Agenda for Marketing. *Australasian Marketing Journal*, *29*(3), 204–214. https://doi.org/10.1177/1839334921999479

Whittaker, L., Mulcahy, R., Letheren, K., Kietzmann, J., & Russell-Bennett, R. (2023). Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda. *Technovation*, *125*, 102784. https://doi.org/10.1016/j.technovation.2023.102784

Wilson, J. (2022, October 13). Deepfake: Post The Bruce Willis Controversy What Disruption To Entertainment Could Be Caused. *Forbes*. https://www.forbes.com/sites/joshwilson/2022/10/13/deepfake-post-the-bruce-willis-controversy-what-disruption-to-entertainment-could-be-caused/

Wonderlandmovies (Director). (2018, February 19). *Zalando #whereveryouare campaign with Cara Delevingne* [Video recording]. https://www.youtube.com/watch?v=OIGmaHaTu58

Zero Malaria Britain (Director). (2019, September 5). *David Beckham speaks nine languages to launch Malaria Must Die Voice Petition* [Video recording]. https://www.youtube.com/watch?v=QiiSAvKJIHo

I have made every effort to locate all copyright holders and obtain their permission to use the images in this work. Should any copyright infringement nevertheless come to light, please contact me.

Acknowledgements

At this point, I would like to thank Univ.-Prof. Dr. Martin Eisend, M.A. for his patience, his guidance and valuable feedback.

I also want to thank all participants who contributed to this study as well as my father and my mother for helping me read through the thesis. I want to especially thank my uncle Roland for giving me advice based on his experience as a journal reviewer. Although marketing is not his field, the insights provided were extremely helpful.

I want to thank my brother Julien and my friends Sarah, Laura, Claudia, Deni, Kim, André and Denise for being by my side and believing in me during this difficult past year.

Appendices

Appendix 1

Appendix 1: Data description and cleaning

a. Data cleaning:

In the following we will describe how we decided to clean the data issued from our experiment.

We had a total of 293 respondents. However, only 218 completed the whole questionnaire.

We had added two attention check questions. The first one was related to the deepfake awareness text. This was shown to Group 2 and Group 4. The second one was an attention question checking if the respondents watched the videos attentively: it was asking which brand was shown in the video. The second question was showed to all respondents. We only kept those respondents that passed these attention questions. Then 186 respondents remained.

Next, we excluded participants who did not recognize the celebrity they had selected, which was crucial to ensure the manipulation worked as intended. 27 respondents that did not recognize the celebrity were therefore removed.

We also excluded respondents who knew Neha Sharma, this celebrity was shown only in the group that did not receive personalized advertising. In order to avoid that these respondents think the advertisement was tailored to their ethnicity or preferences we removed these 6 respondents.

We then excluded one child (14 years old), primarily due to data protection concerns and because minors generally cannot make purchase decisions independently. Also we decided to excluded respondents over 70 years old, including one respondent over 90.

Furthermore, we removed respondents who were familiar with the brand Flower Knows. According to Aljarah et al. (2024), brand familiarity is a significant moderator for awareness of falsity. To avoid bias in our results, we excluded respondents who knew the brand. The brand was chosen specifically because it is still relatively unknown to the European public, allowing us to use its products in the video without strong prior associations. Familiarity, either positive or negative, could have distorted the results. This step reduced the sample size to 135.

Finally, after reviewing box plots of our key variables, we excluded three outliers with z-scores of ±3.0 (Osborne & Overbay, 2004): one for awareness of falsity ($z = 3.14$) and two for perceived realism ($z = 3.26$; $z = 3.81$). These responses were deemed too extreme and were therefore excluded.

The final sample consisted of $N = 133$.

b. *Demographics and their characteristics distribution between the 4 experimental groups:*

It was important for us to make sure that the characteristics gender and age were similar in the 4 groups.

We saw that the group sizes were similar (G1 = 33, G2 = 30, G3 = 35, G4 = 35).

Then we looked at the ages and performed an ANOVA, see table a 6 that showed that there are no significant age differences ($F(3,129) = 1.55$, $p = 0.205$). Gender was similarly distributed across groups, since the chi-square test, see table a 6, we conducted (excluding the three participants who selected "Other" or "Non-binary") was not significant ($\chi^2(3, N = 130) = 3.35$, $p = 0.341$). The country of residence was also evenly distributed, as did the educational background, further details can be seen on table a 2 and a 3.

c. *Data description:*

Table a 1: Demographics age of the respondents overall and between the four experimental groups

	Age overall	Group 1	Group 2	Group 3	Group 4
N	133	33	30	35	35
Mean	27.16	25.64	28.90	26.74	27.51
Median	26.00	25.11	27.50	25.00	26.00
Std. Deviation	6.22	5.50	4.40	7.80	6.27
Min.	18	18	23	20	19
Max.	65	47	40	65	55
Variance	38.74	30.24	19.33	60.90	39.32
Skewness	2.86	2.07	.92	3.93	2.54
Kurtosis	12.81	6.42	-.03	17.89	10.20

Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Table a 2: Educational background

	Group 1	Group 2	Group 3	Group 4	Total
High school Diploma or equivalent	6	2	5	5	18
Vocational training	2	0	1	2	5
Bachelor degree	18	14	21	19	72
Master Degree	7	12	7	8	34
Doctorate or Higher	0	2	1	1	4

Note. *n=133*

Table a 3: Country of residence

	Group 1	Group 2	Group 3	Group 4	Total
German speaking countries	28	27	30	31	116
Other European countries	2	3	3	2	10
Outside of the Europe	3	0	2	2	7

Note. *n=133*

Table a 4: Gender distribution

	Group 1	Group 2	Group 3	Group 4	Total
Female	23	20	29	27	99
Male	10	9	5	7	31
Other	0	1	1	1	3

Note. *n=133*

Table a 5: Test of normality for the age distribution

	Statistic	df	Sig.
Group 1	.81	33	<.001*
Group 2	.90	30	.011
Group 3	.55	35	<.001*
Group 4	.78	35	<.001*

Note. *A Shapiro Wilk test has been performed*

Table a 6: Distribution of gender and age between the experimental groups

Aged distribution ANOVA

	Group	M	SD	df	F	Sig.	η^2
Descriptives	Group 1	25.64	5.40	-	-	-	
	Group 2	28.90	19.33	-	-	-	
	Group 3	26.74	7.80	-	-	-	
	Group 4	27.51	6.27	-	-	-	
ANOVA: Age distribution between the groups				3, 129	1.55	.205	.035

Gender distribution Chi-Square test

	df 1	$\chi 2$	Sig.	V
Gender distribution between the groups	3	3.35	.341	.16

Note. *For the age of the respondents: Normality was assumed and the levene test showed homogeneous variances (F(3,129) =.116, p=.951).For the gender we only looked at male and female since one of the assumptions for the Chi-Square test is to have at least 5 answers per categorical group.*

Appendix 2

Appendix 2: Scale validity and reliability

Table a 7: Reliability: Internal consistency of the variables tested

	Cronbach Alpha
Awareness of falsity	
AF02 (product claims, 3 items)	.92
AF03 (product unrelated claims, 3 items)	.93
AF04 (presented reality, 2 items)	.83
Awareness of falsity overall construct	0.87
Perceived Realism	
PV01 (plausibility, 5 items)	.88
PV02 (typicality, 3 items)	0.75
PV03 (factuality, 3 items)	.84
PV04 (narrative consistency, 5 items)	.85
PV05(perceptual quality, 5 items)	.84
Perceived realism overall construct	.90
Control variables	
BP05 (brand attitude, 5 items)	.93
CE01 (brand endorsement, 4 items)	.90
DQ01(deepfake quality, 6 items)	.81

Note. *AF0X: Awareness of Falsity Scale number X; PV0X: perceived realism scale number X; BP0X: brand attitude number X; CE0X: celebrity endorsement scale number X and DQ0X: Deepfake quality number X. All scales cane be found in the Excel code sheet.*

Table a 8: Construct Validity : Factor analysis awareness of falsity, rotated component matrix

PCAT item	Factor loading		
	1	2	3
Factor 1: Product Unrelated Claims			
AF02_01 Unconvincing / Convincing	.909	-	-
AF02_02 Unbelievable / Believable	.919	-	-
AF02_03 Dishonest / Honest	.848	-	-
Factor 2: Product Related Claims			
AF03_01 Generally truthful	-	.898	-
AF03_02 Accurately informed	-	.882	-
AF03_03 Believable	-	.870	-
Factor 3: Presented Realism			
AF04_01 Not realistic / Realistic	-	-	.864
AF04_02 Could exist in real life / Likely in real life	-	-	.936

Note. *Extraction Methode: Principal Component Analysis/ Rotation Method: Varimax with Kaiser normalization. Loadings <0.40 have been left out. The factor Analysis showed that the construct validity is given. Just as theoretical explained by Campbel et al. (2021), the construct is build up on 3 factors. All factor loads are mainly on their respective factor and we could not identify substantial cross-loadings.*

Appendix 3:

Appendix 3: Control variables

a. Control variable 1: Celebrity endorsement

Table a 9 shows the descriptive data of each celebrity of the variable celebrity endorsement. The video including the celebrity Bella Poarch had to be excluded since only one respondent selected her. To test whether the presented celebrities led to different brand endorsement perceptions, we conducted a Welch ANOVA, visible in table a 10. The Welch ANOVA showed no significant differences between the celebrities ($N = 132$, $F(5, 19.60) = 1.54$, $p = 0.224$). The same result was found when comparing the four study groups see table a 11. As the assumptions for the ANOVA was also not met, the levene test in table a 11 showed no homogeneous variances, we thus performed a one-Welch ANOVA, which was not significant ($F(3, 71.34) = 0.28$, $p = 0.840$). Thus there were no differences between the groups and the participants felt that the celebrities shown in the groups were a similar good fit to the brand.

We then used simple linear regression, to check whether brand endorsement impacts purchase intention. As showed in table a 12 the regression was significant ($F(1, 131) = 5.26$, $p = 0.023$). The model explained 3.9% of the variance. Thus celebrity endorsement predicts purchase intention ($\beta = -0.20$, $t(131) = -2.29$, $p = 0.023$) and we need to control the variable "celebrity endorsement" in further analyses involving purchase intention.

Table a 9: Celebrity endorsement between the different celebrities shown

Celebrity (Group)	N	M	Median	Min	Max	SD	Variance	Skewness	Kurtosis
Lady Gaga	12	2.79	2.50	1.00	6.00	1.45	2.10	0.97	0.90
Anya Taylor-Joy	4	3.69	3.38	1.00	7.00	2.58	6.64	0.60	-0.30
Anne Hathaway	23	4.41	4.75	1.00	7.00	1.83	3.37	-0.34	-0.71
Ariana Grande	9	3.78	4.00	2.25	6.50	1.60	2.55	0.78	-0.65
Margot Robbie	21	3.44	3.00	1.25	5.25	1.23	1.51	0.07	-1.18
Neha Sharma	63	3.54	3.75	1.00	6.25	1.14	1.29	0.04	0.17

Note. *Celebrity Bella Poarch has been excluded since only one person has chosen this celebrity.*

Table a 10: Welch ANOVA : Difference of the celebrity endorsement between the celebrities shown

Test	Group	Mean	SD	Statistic	df_1	df_2	p	η^2
Shapiro–Wilk (Normality)	Lady Gaga	-	-	.913	-	12	.231	-
	Anya Taylor-Joy	-	-	.978	-	4	.892	-
	Anne Hathaway	-	-	.949	-	23	.283	-
	Ariana Grande	-	-	.846	-	9	.068	-
	Margot Robbie	-	-	.937	-	21	.187	-
	Neha Sharma	-	-	.959	-	63	.033*	-
Levene's Test (Homogeneity)	Based on mean	-	-	2.95	5	126	.015	-
Descriptive Statistics	Lady Gaga	2.80	1.45	-	-	-	-	-
	Anya Taylor-Joy	3.69	2.58	-	-	-	-	-
	Anne Hathaway	4.41	1.83	-	-	-	-	-
	Ariana Grande	3.78	1.60	-	-	-	-	-
	Margot Robbie	3.44	1.23	-	-	-	-	-
	Neha Sharma	3.54	1.14	-	-	-	-	-
Welch's ANOVA (Equality of Means)	-	-	-	1.538	5	19.60	.224	.09

Note. *We chose a Welch ANOVA since the distribution of the respondents in the groups were unequal and the homogeneity of the variances was not given.*

Table a 11: Welch ANOVA : Difference between the celebrity endorsement between the 4 experimental groups

Test	Group	M	SD	Statistic	df_1	df_2	p	η^2
Shapiro–Wilk (Normality)	Group 1	-	-	.958	-	33	.203	-
	Group 2	-	-	.954	-	29	.197	-
	Group 3	-	-	.956	-	34	.213	-
	Group 4	-	-	.915	-	34	.013*	-
Levene's Test (Homogeneity of Variances)	Based on mean	-	-	6.325	3	129	< .001*	-
Descriptive Statistics	Group 1	3.60	1.16	-	-	-	-	-
	Group 2	3.45	1.13	-	-	-	-	-
	Group 3	3.75	1.82	-	-	-	-	-
	Group 4	3.67	1.53	-	-	-	-	-
Welch's ANOVA	Between Groups			.279	3	71.34	.840	.006

Note. *We chose a Welch ANOVA, since the homogeneity of the variances was not given.*

Table a 12: Simple linear regression prediction purchase intention from the variable celebrity endorsement

Effect	Estimate (B)	SE	95% CI		p
			LL	UL	
Fixed effects					
Intercept	3.139	.336	2.48	3.80	<.001
Celebrity Endorsement	-.197	.086	-.366	-.028	.023
Model summary	R^2	Adjusted R	df1	df2	F
	.039	.031	1	131	5.26

Note. *Dependent variable: Purchase Intention.*

b. *Control variable 2: Brand attitude*

To test whether the variable brand attitude impacted the four groups, we conducted a one-way ANOVA. Table a 13 shows the result, it can be seen that the difference between the means of purchase intention of the experimental groups was not significant ($F(3,129) = 0.522$, $p = 0.836$), meaning that brand attitude did not significantly vary between the groups. As we did with the other variables we performed a simple regression analysis in table a 14, which was statistically significant ($F(1,131) = 6.91$, $p = 0.010$, also only explaining 5.0% of the variance. Brand attitude significantly predicted purchase intention ($\beta = -0.36$, $t(131) = 2.63$, $p = 0.010$) and thus we need to control brand attitude when analysing purchase intention.

Table a 13: One way ANOVA : Difference between the brand attitude between the 4 experimental groups

Test	Group	M	SD	Statistic	df_1	df_2	p	η^2
Shapiro–Wilk (Normality)	Group 1	-	-	.905	-	33	.007*	-
	Group 2	-	-	.910	-	30	.015*	-
	Group 3	-	-	.876	-	35	<.001*	-
	Group 4	-	-	.916	-	34	.013*	-
Levene's Test (Homogeneity of Variances)	Based on mean	-	-	.285	3	129	.836	-
Descriptive Statistics	Group 1	4.81	.97	-	-	-	-	-
	Group 2	4.59	.88	-	-	-	-	-
	Group 3	4.62	.83	-	-	-	-	-
	Group 4	4.56	.89	-	-	-	-	-
One-Way ANOVA	Between Groups			.522	3	129	.668	.012

Note. *We chose a One Way ANOVA, since the requirements for the ANOVA have been met (normal distribution, the Shapiro Wilk test did not indicate normal distribution, however the QQ plots indicate a sufficient normal like distribution to perform the ANOVA and homogeneous variances)*

Table a 14: Simple linear regression prediction purchase intention from the variable brand attitude

Effect	Estimate (B)	SE	95% CI		p
			LL	UL	
Fixed effects					
Intercept	.754	.646	-.512	2.02	.245
Brand attitude	.359	.137	.090	.627	.010
Model summary	R^2	Adjusted R	df1	df2	F
	.050	.043	1	131	6.92

Note. *Dependent variable: Purchase Intention.*

c. *Control variable 3: Deepfake quality*

Maintaining consistent deepfake quality across all videos was crucial, so as not to distort the results. As shown in (Kaate et al., 2023), inconsistencies in deepfake quality can trigger the uncanny valley effect. In order to evaluate the quality, we used a scale developed by Plohl et al. (2024)**.** The Likert scale included questions like "The face of the person in the video (or parts of it) was distorted." or " The mouth of the person in the video was moving strangely." (Plohl et al., 2024, p. 7). We specifically focused on face-related items, as this was the area altered by the deepfake technology.

It was important that the overall video quality remained consistent across the celebrities. To test whether deepfake quality had an impact across the different videos, we conducted a Welch ANOVA. Although the normality plots looked acceptable and homogeneity of variances was given, visible in table a 16, the Welch ANOVA was chosen as it is more robust in cases where group sizes differ significantly as was the case here. As visible in table a 16 the Welch ANOVA revealed no significant difference in deepfake quality between the different videos.

Additionally, we performed a One-Way ANOVA, visible in table a 15, since the group distribution was similar and variance homogeneity was met, to test for significant differences across the four defined research groups. Again, no significant difference was found.

Table a 15: One way ANOVA : Difference between the deepfake quality between the 4 experimental groups

Test	Group	M	SD	Statistic	df_1	df_2	p	η^2
Shapiro–Wilk (Normality)	Group 1	-	-	.973	-	33	.578	-
	Group 2	-	-	.976	-	30	.720	-
	Group 3	-	-	.915	-	35	.010*	-
	Group 4	-	-	.878	-	35	.001*	-
Levene's Test (Homogeneity of Variances)	Based on mean	-	-	1.02	3	128	.385	-
Descriptive Statistics	Group 1	4.74	1.18	-	-	-	-	-
	Group 2	4.60	1.07	-	-	-	-	-
	Group 3	5.00	1.07	-	-	-	-	-
	Group 4	4.97	1.43	-	-	-	-	-
One-Way ANOVA	Between Groups			.878	3	128	.454	.02

Note. *We chose a One Way ANOVA, since the requirements for the ANOVA have been met (normal distribution, the Shapiro Wilk test did not indicate normal distribution for group 3 and 4, however the QQ plots indicate a sufficient normal like distribution to perform the ANOVA and homogeneous variances)*

Table a 16: One way ANOVA : Difference of the deepfake quality between the celebrities shown

Test	Group	Mean	SD	Statistic	df_1	df_2	p	η^2
Shapiro–Wilk (Normality)	Lady Gaga	-		.891	-	12	.121	-
	Anya Taylor-Joy	-		.827	-	4	.161	-
	Anne Hathaway	-		.879	-	23	.010*	-
	Ariana Grande	-		.877	-	9	.145	-
	Margot Robbie	-		.865	-	21	.008*	-
	Neha Sharma	-		.991	-	63	.918	-
Levene's Test (Homogeneity)	Based on mean	-		.710	5	126	.617	-
Descriptive Statistics	Lady Gaga	4.97	1.19	-	-	-	-	-
	Anya Taylor-Joy	4.87	1.98	-	-	-	-	-
	Anne Hathaway	5.10	1.25	-	-	-	-	-
	Ariana Grande	4.65	1.55	-	-	-	-	-
	Margot Robbie	5.09	1.11	-	-	-	-	-
	Neha Sharma	4.67	1.12	-	-	-	-	-
One Way ANOVA (Equality of Means)	-	-	-	.691	5	126	.631	.03
Welch ANOVA				0.643	5	19.9	.670	

Note. *Deepfake perceived quality of each deepfake video used in our experiment, N is the amount of respondents on a 6 Likert scale questions from* Plohl et al.(2024, p. 7). *The participants answer on a scale from 1 to 7, 7 meaning lowest quality. M is the mean of the value about all questions and SD is the standard deviation. The overall quality of the videos is similar which was important for this study. We chose a Welch ANOVA even though assumptions for he test have been met because of the different sample sizes(The Shapiro Wilk test indicated that only the respondents from Anne Hathaway and Margot Robbie are not normally distributed; a closer look at the QQ plots however showed that normal distribution can be assumed). Bella Poarch has been excluded since the celebrity only had one respondent.*

Appendix 4:

Appendix 4: Data analysis overview and descriptives

a. Data analysis overview

Table a 17: Data analysis overview

H	IV	DV	Moderator / Mediator / Covariate	Statistical test and Assumptions tested
AA	**PR**	**PI**	**CV: CE, BA**	**Multiple Linear Regression** Normality: P-P plot (assumption met) figure a4 Homoscedasticity: Residual scatterplot (no violations detected) figure a5 Linearity: Residual plots (no systematic pattern) figure a5 Multicollinearity: VIF = 1.04–1.12 (< 10)
AB	**PR**	**AF**	**–**	**Linear Regression** Normality: P-P plot (assumption met) figure a6 Homoscedasticity: Residual scatterplot (no violations detected) figure a7 Linearity: Residual plots (no systematic pattern observed) figure a7
AC	**AF**	**PI**	**CV: CE, BA**	**Multiple Linear Regression** Normality: P-P plot (assumption met) figure a8 Homoscedasticity: Residual scatterplot (no violations detected) figure a9 Linearity: Residual plots (no systematic pattern observed) figure a9 Multicollinearity: VIF = 1.27 (< 10)

H	IV	DV	Moderator / Mediator / Covariate	Statistical test and Assumptions tested
H1	**AD**	**PR**	**–**	**Independent Sample t-Test** Normality: Shapiro-Wilk test not significant (p_no awareness = .540; p_awareness = .755) thus assumption is met Homogeneity of variances: Levene's test not significant (F(1,131) = 0.66, p = .416), thus the variances are homogeneous. **MANOVA** <u>*Normality: Given*</u> *Plausibility:* Shapiro Wilk test significant (p_no awareness = <.001; p_awareness = .005) but the Q-Q plots in figure a10 indicate acceptable normality *Factuality:* Shapiro Wilk test significant but the Q-Q plots in figure a11 indicate acceptable normality *Typicality:* Shapiro Wilk test significant but the Q-Q plots in figure a12 indicate acceptable normality *Perceptual Quality:* Shapiro Wilk test significant for the first group (p_group1 no awareness = .045;p_group2 awareness = .150) but the Q-Q plot in figure a13 indicate acceptable normality *Narrative Consistency:* Shapiro Wilk test not significant (p_no awareness = .255; p_awareness = .146) <u>*Homogeneity of variances: assumption is confirmed*</u> Plausibility: Levene's test not significant (F(1,131) =1.24, p =.268) Factuality: Levene's test not significant (F(1,131) =2.77, p =.098) Typicality: Levene's test not significant (F(1,131) =2.98, p =.086)

H	IV	DV	Moderator / Mediator / Covariate	Statistical test and Assumptions tested
				Perceptual Quality: Levene's test not significant (F(1,131) =.73, p =.384) Narrative Consistency: : Levene's test not significant (F(1,131) = .002, p =.968) <u>*Multicollinearity:*</u> A collinearity analysis using Pearson correlations among the dependent variables showed that correlations ranged from r = .11 to r = .58. As none of the correlations exceeded the recommended threshold of r = .90 (Tabachnick & Fidell, 2013) multicollinearity was not considered a concern.
H2	**AD, P**	**PI**	**–**	**Two-Way ANOVA** Normality: Shapiro-Wilk significant but the Q-Q plots in figures a 14, a15, a16 and a17 indicate acceptable normality Homogeneity of variances: Levene's test not significant (p = .192)
H3	**P**	**PI**	**Moderator: AF**	**Moderation (PROCESS Model 1, Hayes 2017)** Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .04 >.05 But use of HC3 robust p = .17 Linearity: Residual plots of AF and PI (no systematic pattern observed) see figure a9 Multicollinearity: VIF range = 1.02–2.71 Important outliers: Bonferroni Outlier Test p = .032 >0.05
H4	**AD**	**AF**	**Mediator: PoR**	**Mediation (PROCESS Model 4, Hayes 2017)** <u>AD→PoR</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .94 >.05 Multicollinearity: VIF = 1.00 (< 10)

H	IV	DV	Moderator / Mediator / Covariate	Statistical test and Assumptions tested
				No important outliers: Bonferroni Outlier Test p = 1.00 >0.05 <u>AD+PoR→AF</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .074 >.05 Linearity: Scatterplot of AF and PoR in figure a18 indicate linearity Multicollinearity: VIF = 1.05 (< 10) No important outliers: Bonferroni Outlier Test p = .45 >0.05 <u>AD→AF</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .21 >.05 Multicollinearity: VIF = 1.00 (< 10) No important outliers: Bonferroni Outlier Test p = .13 >0.05
H5	**AD**	**AF**	**Mediator: PLR**	**Mediation (PROCESS Model 4, Hayes 2017)** <u>AD→PLR</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .45 >.05 Multicollinearity: VIF = 1.00 (< 10) No important outliers: Bonferroni Outlier Test p = .39>0.05 <u>AD+PLR→AF</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .14 >.05 Linearity: Scatterplot of AF and PLR in figure a19 indicate linearity Multicollinearity: VIF = 1.01 (< 10)

H	IV	DV	Moderator / Mediator / Covariate	Statistical test and Assumptions tested
				No important outliers: Bonferroni Outlier Test p = .26 >0.05 <u>AD→AF</u> Normality: use of Bootstrapping Homoscedasticity: Breusch-Pagan Test p = .21 >.05 Linearity: Residual plots (no systematic pattern observed) Multicollinearity: VIF = 1.00 (< 10) No important outliers: Bonferroni Outlier Test p = .13 >0.05

Note. *P:personalization, PR. perceived realism, AD: awareness of deepfakes, AF: awareness of falsity, PoR: perceived portrayal of realism, PLR: perceived likelihood of realism, CV: control variables, CE: celebrity endorsement, BA: Brand attitude.*

b. Descriptive main variables overviews

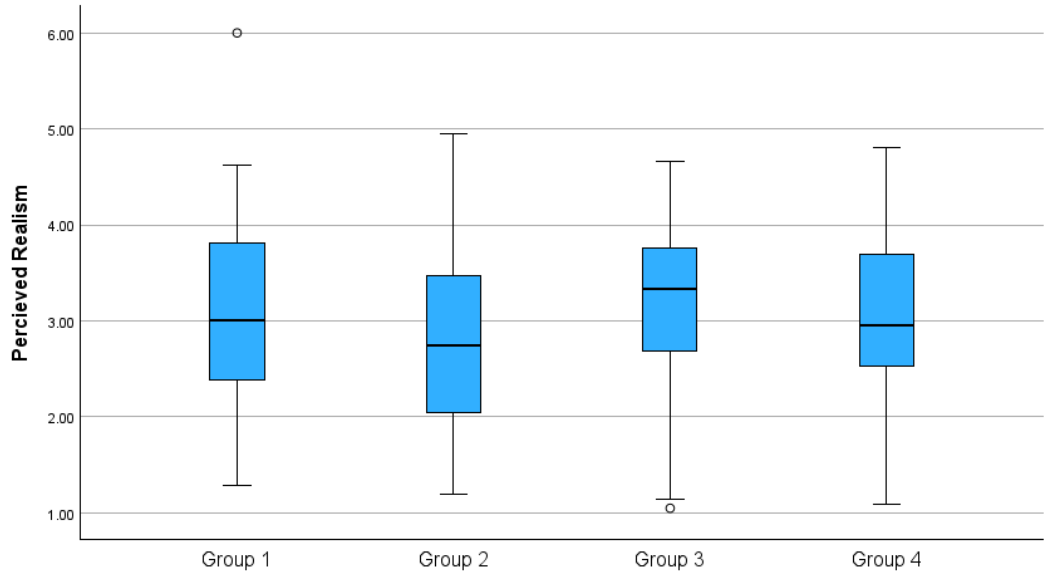
PERCEIVED REALISM

Table a 18: Descriptive statistic perceived realism between the 4 experimental Groups

	Perceived Realism	Group 1	Group 2	Group 3	Group 4
N	133	33	30	35	35
Mean	3.02	3.03	2.87	3.15	3.01
Median	3	3	2.74	3.33	2.95
Std. Deviation	.93	1.04	.95	.91	.83
Min.	1.05	1.29	1.19	1.05	1.10
Max.	6	6	4.95	4.67	4.81
Variance	.87	1.09	.913	.837	.69
Skewness	.15	.57	.44	-.57	.02
Kurtosis	-.18	.49	-.44	-.17	-.22

Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Figure a 1: Boxplot of perceived realism across experimental conditions



Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Table a 19: Descriptive statistic perceived realism separated between the manipulated variables awareness of deepfakes and personalization

	Awareness of deepfakes		Personalization	
	No	Yes	No	Yes
N	68	65	30	35
Mean	3.09	2.95	2.96	3.08
Median	3.19	2.86	2.86	3.12
Std. Deviation	.97	.89	.99	.87
Min.	1.05	1.10	1.19	1.05
Max.	6	4.95	6	4.81
Variance	.95	.79	.99	.76
Skewness	.07	.22	.52	-.29
Kurtosis	.04	-.45	.09	-.34

Note. *n=133*

Table a 20: Descriptive statistic perceived realism and the 5 subdimensions.

	Perceived Realism		Typicality		Factuality		Plausibility		Narrative-consistency		Perceptual Quality	
Mean	3.02		2.31		2.37		2.83		3.85		3.2	
Median	3		2		2		2.6		4		3	
SD	.93		1.19		1.20		1.43		1.30		1.25	
Min.	1.05		1.		1		1		1		1	
Max.	6		6.33		6		6.80		6.60		6	
Variance	.87		1.42		1.44		2.05		1.69		1.56	
Skewness	.15		.97		.67		.50		-.08		.35	
Kurtosis	-.18		.67		-.23		-.68		-.75		-.35	
Awareness Groups	G1	G2	G1	G2	G1	G2	G1	G2	G1	G2	G1	G2
Mean	3.09	2.95	2.17	2.45	2.57	2.17	2.61	3.06	4.08	3.61	3.46	2.93
SD	.97	.89	1.08	1.29	1.27	1.09	1.33	1.51	1.31	1.25	1.30	1.15

Note. *n=133 Group 1: no awareness of deepfakes; Group 2: awareness of deepfakes; SSD: Standard Deviation*

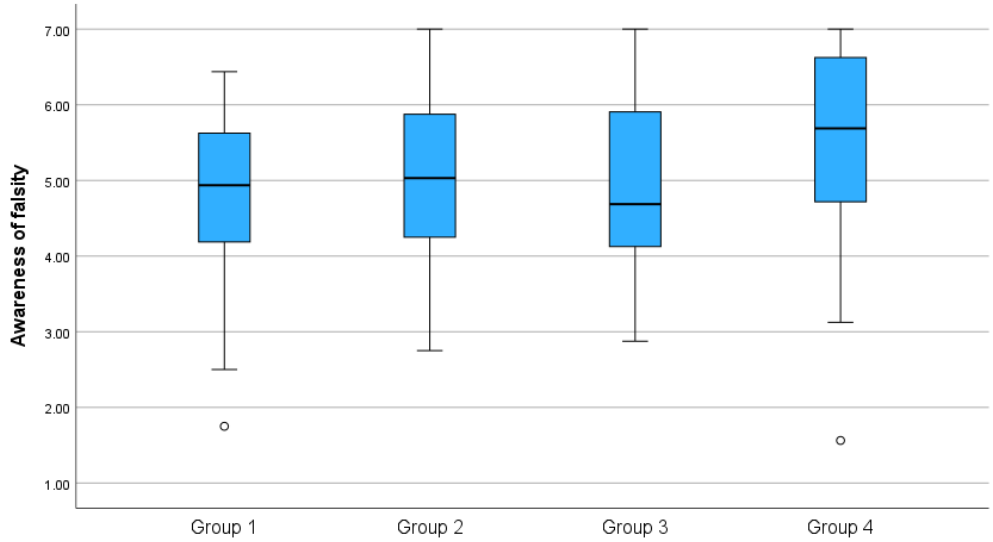
AWARENESS OF FALSITY

Table a 21: Descriptive statistic awareness of falsity between the 4 experimental Groups

	Awareness of Falsity	Group 1	Group 2	Group 3	Group 4
N	133	33	30	35	35
Mean	5.09	4.80	5.13	4.94	5.47
Median	5.12	4.94	5.03	4.69	5.69
Std. Deviation	1.18	1.03	1.13	1.12	1.36
Min.	1.56	1.75	2.75	2.88	1.56
Max.	7	6.44	7	7	7
Variance	1.40	1.07	1.27	1.26	1.84
Skewness	-.41	-.941	-.19	.02	-.99
Kurtosis	-.22	1.28	-.56	-1.07	.55

Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Figure a 2: Boxplot of awareness of falsity across experimental conditions



Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Table a 22: Descriptive statistic awareness of falsity separated between the manipulated variables awareness of deepfakes and personalization

	Awareness of deepfakes		Personalization	
	No	Yes	No	Yes
N	68	65	30	35
Mean	4.87	5.32	4.96	5.21
Median	4.78	5.56	4.94	5.44
Std. Deviation	1.08	1.26	1.08	1.26
Min.	1.75	1.56	1.75	1.56
Max.	7	7	7	7
Variance	1.16	1.58	1.18	1.60
Skewness	-.36	-.63	-.47	-.46
Kurtosis	-.07	-.07	.35	-.47

Note. *n=133*

Table a 23: Descriptive statistics for the subdimensions of the construct

Statistic	Presented Reality	Product-related Claims	Product-unrelated Claims
N	133	133	133
Mean	5.75	4.84	4.89
Median	6.25	4.67	5.00
Std. Deviation	1.44	1.37	1.70
Min	1.75	1.67	1.00
Max	7.00	7.00	7.00
Variance	2.07	1.88	2.88
Skewness	−1.03	−0.14	−0.53
Kurtosis	0.08	−0.76	−0.55

Note. *N = 133. Descriptive statistics for the four subdimensions of the construct: Awareness of Falsity, Presented Reality, Product-related Claims, and Product-unrelated Claims.*

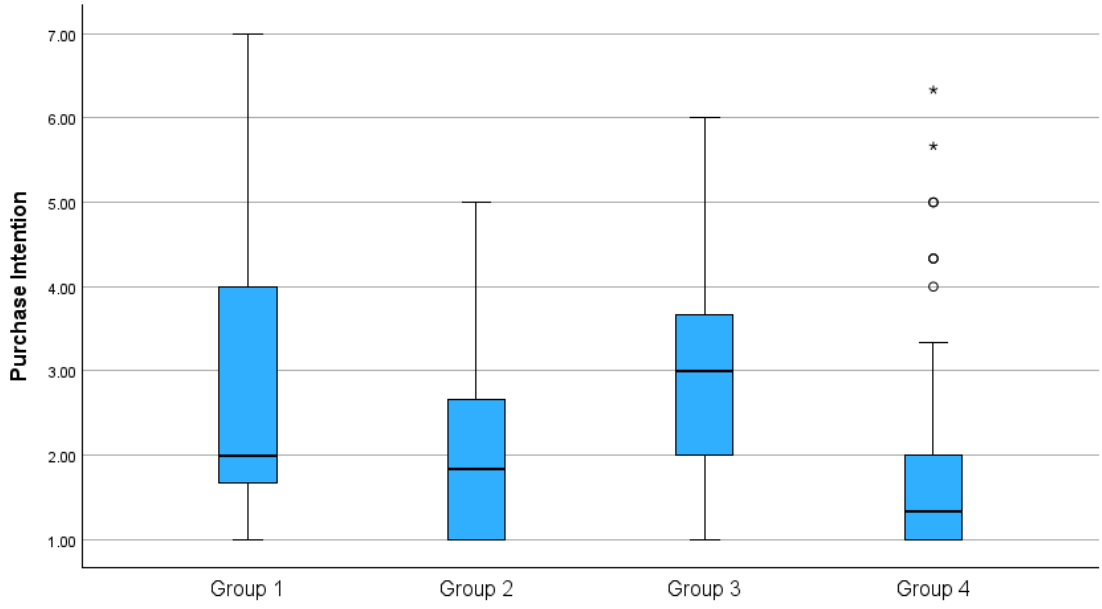
PURCHASE INTENTION

Table a 24: Descriptive statistics for purchase intention between the 4 experimental groups

Statistic	Purchase Intention	Group 1	Group 2	Group 3	Group 4
N	133	33	30	35	35
Mean	2.42	2.68	1.96	2.89	2.11
Median	2.00	2.00	1.83	3.00	1.33
Std. Deviation	1.45	1.58	1.01	1.36	1.57
Min	1.00	1.00	1.00	1.00	1.00
Max	7.00	7.00	5.00	6.00	6.33
Variance	2.09	2.50	1.02	1.86	2.46
Skewness	.97	.98	.94	.40	1.41
Kurtosis	.20	.32	.97	−.50	.79

Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Figure a 3: Boxplot of purchase intention across experimental conditions



Note. *n=133, Group 1: no Awareness of Deepfakes / no Personalization; Group 2: With awareness of Deepfakes / no Personalization; Group 3: no Awareness of deepfakes / with Personalization; Group 4: with Awareness of Deepfakes / with personalization*

Table a 25: Descriptive statistics for purchase intention separated by the manipulated variables awareness of deepfakes and personalization

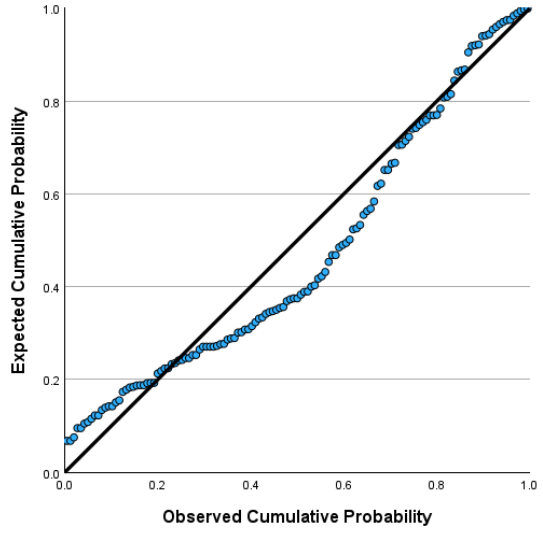
Statistic	Awareness of Deepfakes		Personalization	
	No	Yes	No	Yes
N	68	65	63	70
Mean	2.78	2.04	2.33	2.50
Median	2.33	1.67	2.00	2.00
Std. Deviation	1.47	1.33	1.38	1.52
Min	1.00	1.00	1.00	1.00
Max	7.00	6.33	7.00	6.33
Variance	2.15	1.78	1.90	2.28
Skewness	.69	1.44	1.21	.80
Kurtosis	−.14	1.51	1.33	−.43

Note. *N = 133.*

c. Assumption tests

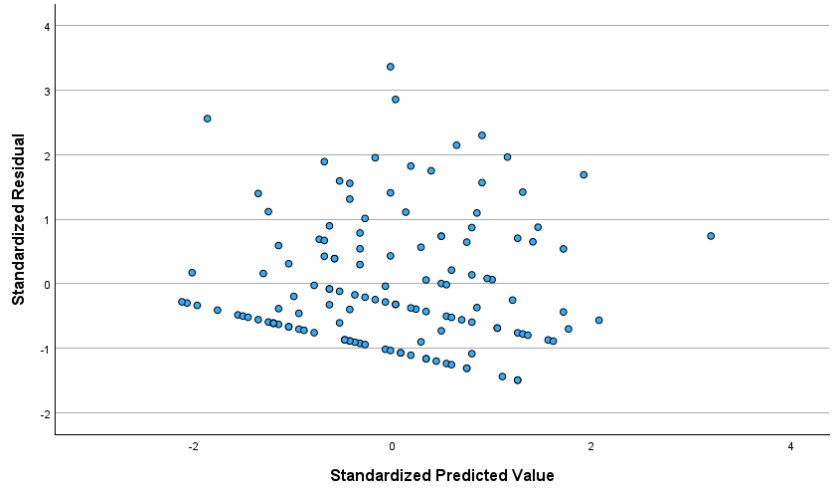
Assumption A:

Figure a 4: Normality P-P Plot of regression standardized residual (DV: purchase intention)



Note. *The plotted points lie close to the diagonal line, indicating that the residuals are approximately normally distributed. Thus, the assumption of normality of residuals can be considered given.*

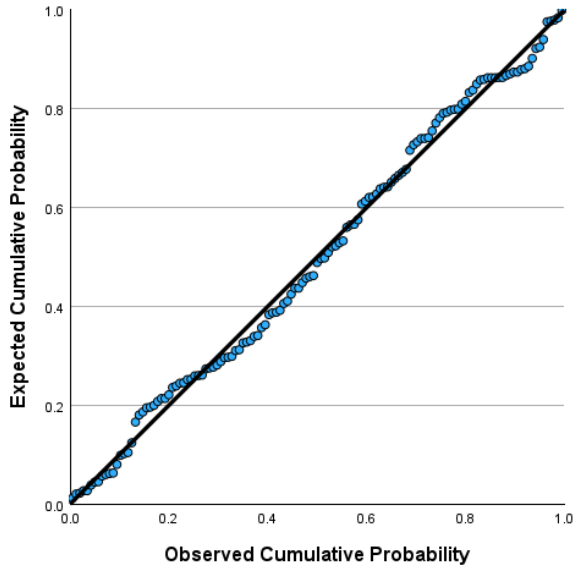
Figure a 5: Scatterplot of the residuals from the regression model performed in assumption A (DV: purchase intention; IV: perceived realism)



Note. *The scatterplot of residuals is used to identify patterns in the residuals and thus to assess the assumptions of linearity and homoscedasticity. There seems to be no major trend recognizable in the scatterplot, thus we can assume that the assumptions are given.*

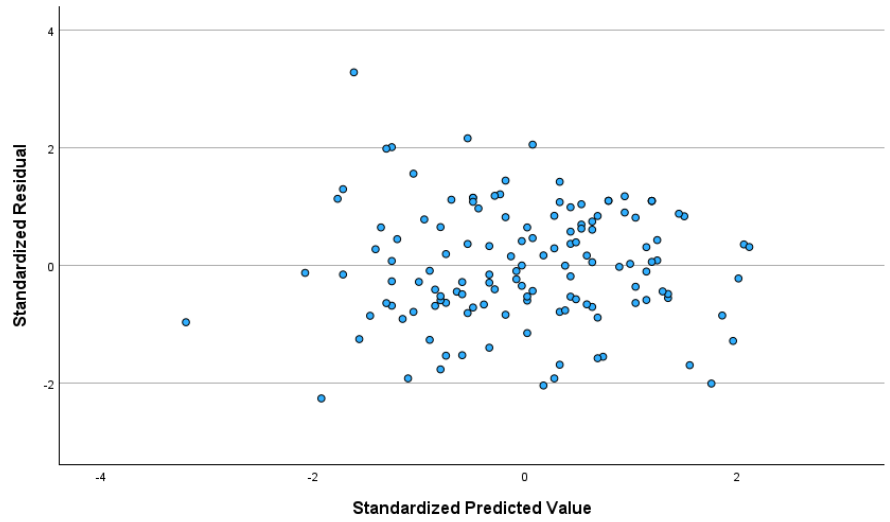
Assumption B:

Figure a 6: Normality P-P Plot of regression standardized residual (DV: awareness of falsity)



Note. *The plotted points are forming a well distinguishable diagonal line indicating that the normality assumption of the residuals of the regression model is given.*

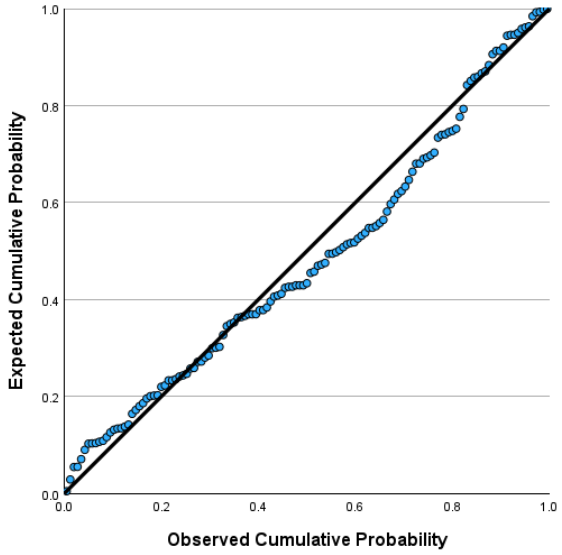
Figure a 7: Scatterplot of the residuals from the regression model performed in assumption B (DV: awareness of falsity; IV: perceived realism)



Note. *The scatterplot shows a cloud of points with no distinguishable funnel shape that could indicate heteroscedasticity or other systematic patterns. Thus, the assumptions of linearity and homogeneity can be considered satisfied*.

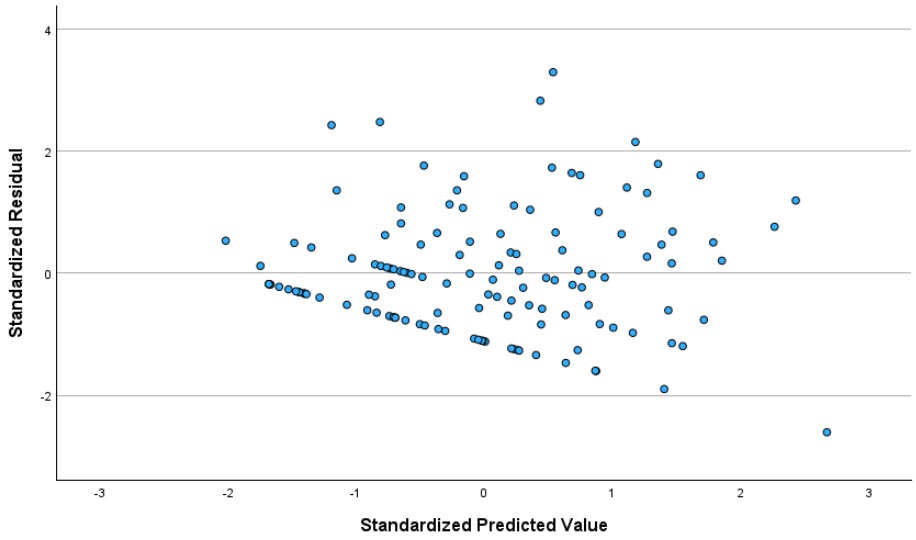
Assumption C:

Figure a 8: Normality P-P Plot of regression standardized residual (DV: purchase intention)



Note. *The plotted points are following the diagonal line indicating that the normality assumption of the residuals of the regression model is can be accepted.*

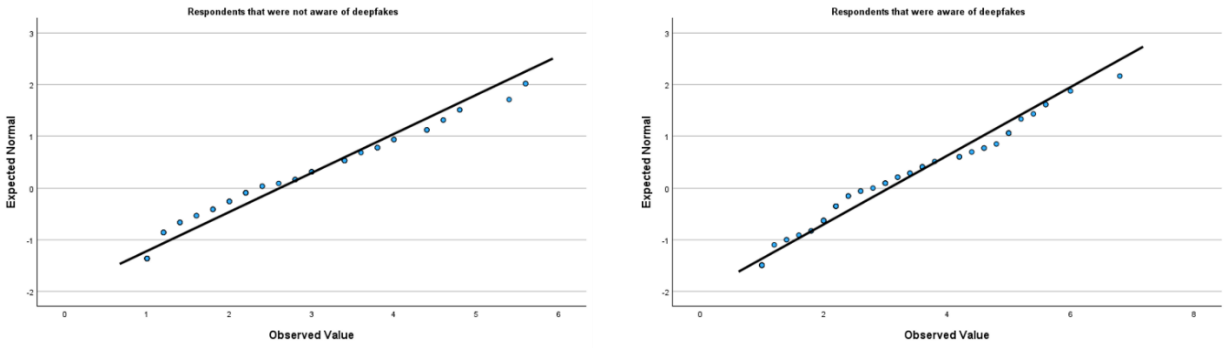
Figure a 9: Scatterplot of the residuals from the regression model performed in assumption C (DV: purchase intention; IV: awareness of falsity)



Note. *The scatterplot shows no major trend recognizable in the scatterplot, thus we can assume that the assumptions are given.*

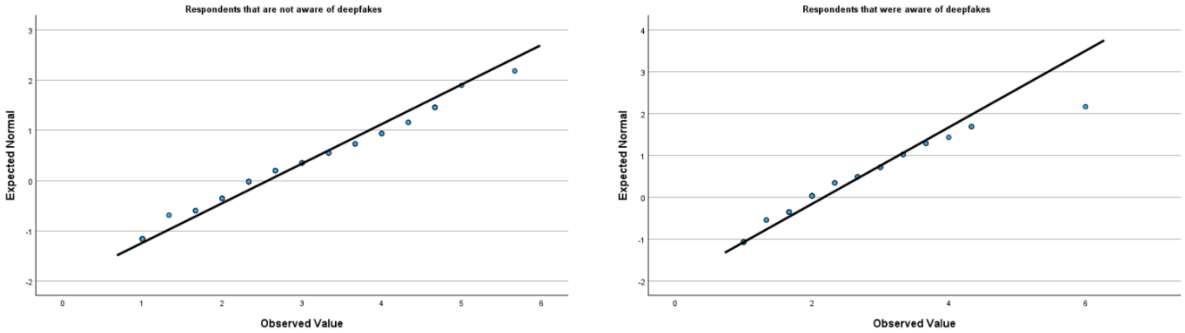
Hypothesis 1:

Figure a 10: Q-Q plots of plausibility across the awareness-of-deepfakes conditions



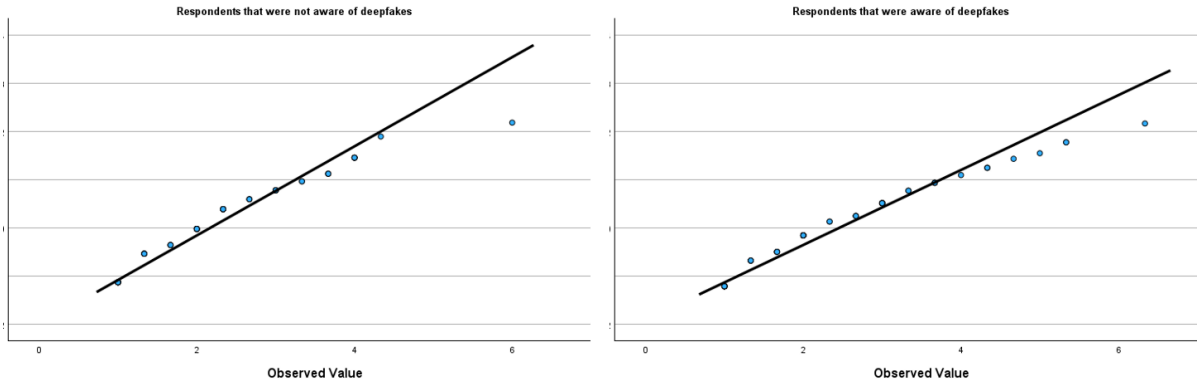
Note. *The plotted points in both QQ plots follow well the diagonal line. We can therefore assume normality.*

Figure a 11: Q-Q plots of factuality across the awareness-of-deepfakes conditions



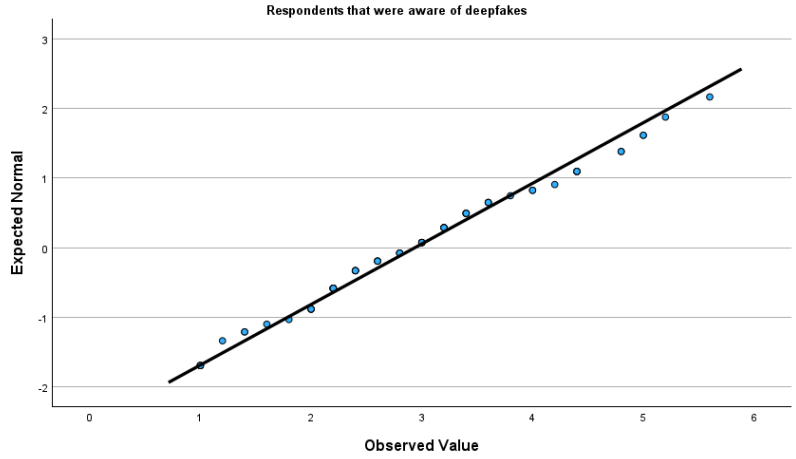
Note. *The plotted points in both QQ plots follow well the diagonal line, therefore we assume normality.*

Figure a 12: Q-Q plots of typicality across the awareness-of-deepfakes conditions



Note. *The plotted points in both QQ plots follow well the diagonal line, so we can assume normality.*

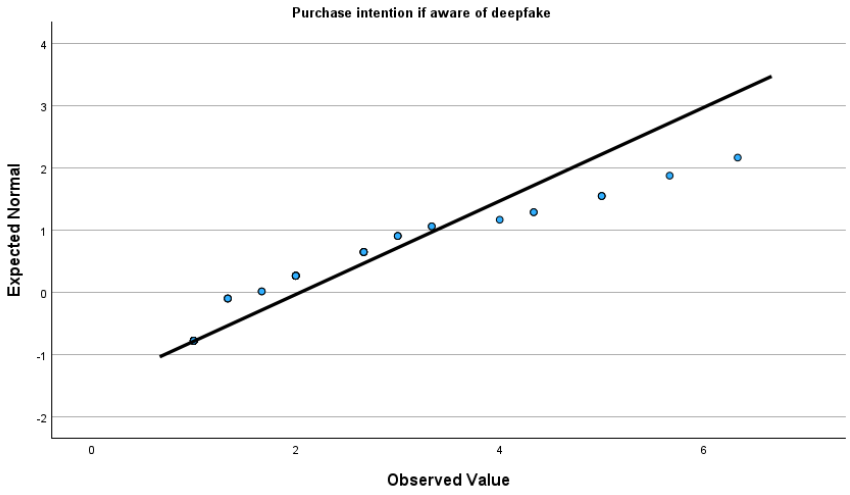
Figure a 13: QQ plot of the variable perceptual quality when respondents were aware of deepfakes



Note. *Since the Shapiro Wilk test was not significant for the none aware group, only the QQ plot for the aware group was analysed. The plotted points align very well to the diagonal line, we thus assume normality.*

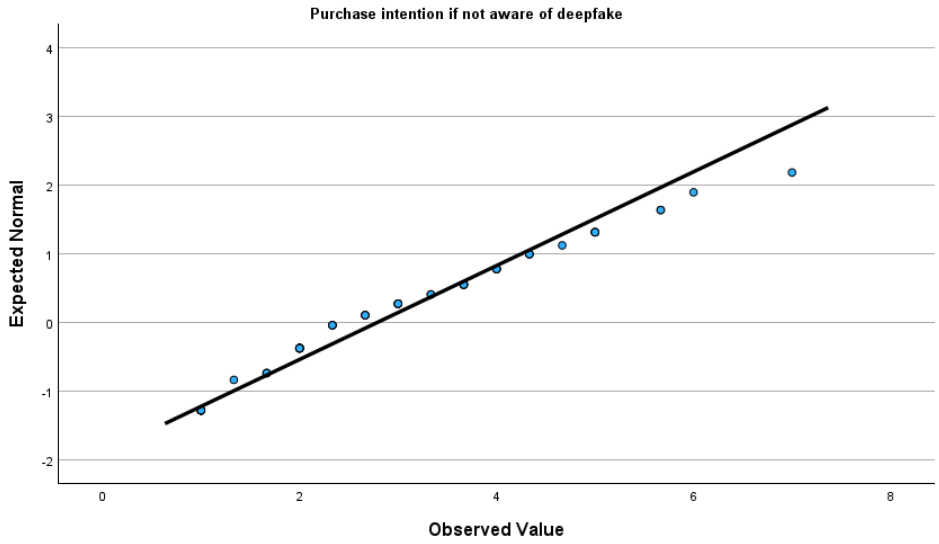
Hypothesis 2:

Figure a 14: QQ plot of the variable purchase intention when respondents were aware of deepfakes



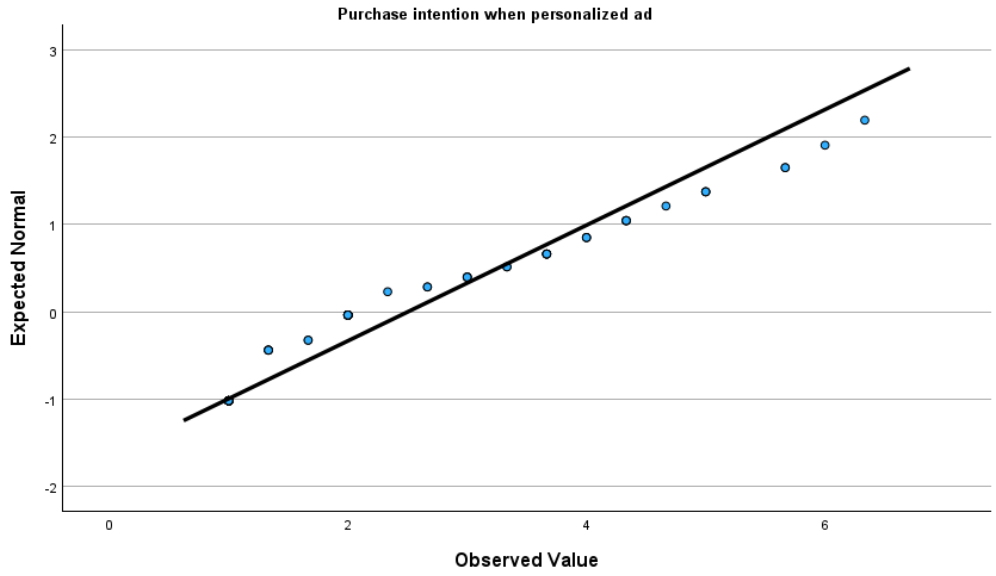
Note. *The plotted points follow the diagonal reference line reasonably well. In the upper range, the points deviate slightly from the line, indicating minor deviations from normality. However, the Q-Q plot overall appears acceptable; therefore, the assumption of normality can be considered satisfied.*

Figure a 15: QQ plot of the variable purchase intention when respondents were not aware of deepfakes



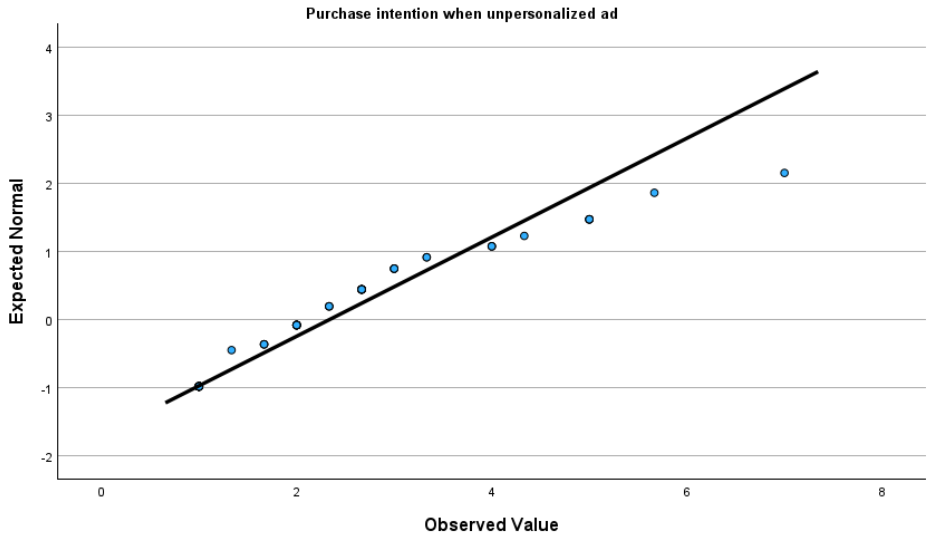
Note. *The plotted points follow well the diagonal line, thus we can assume normality.*

Figure a 16: QQ plot of the variable purchase intention when respondents were confronted with personalized ads



Note. *The plotted points follow well the diagonal line, thus we can assume normality.*

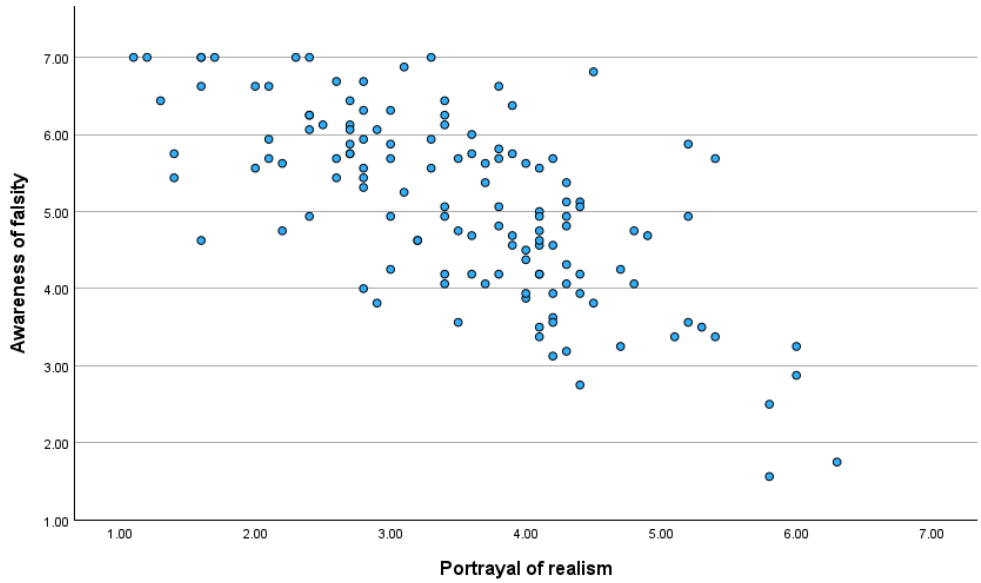
Figure a 17: QQ plot of the variable purchase intention when respondents were confronted with non-personalized ads



Note. *The plotted points follow well the diagonal line, thus we can assume normality.*

Hypotheses 4:

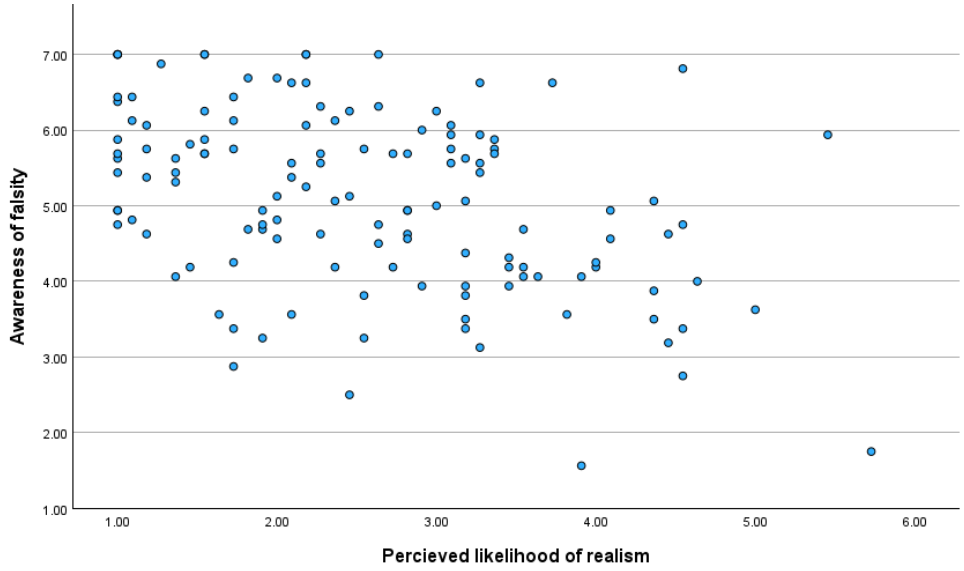
Figure a 18: Scatterplot of portrayal of realism (predictor) and awareness of falsity



Note. *The scatterplot shows a negative linear relationship between awareness of falsity and portrayal of realism. Thus the assumption for linearity is confirmed.*

Hypotheses 5:

Figure a 19: Scatterplot of perceived likelihood of realism (predictor) and awareness of falsity



Note. *The scatterplot shows a negative linear relationship between awareness of falsity and perceived likelihood of realism. Thus the assumption for linearity is confirmed.*

Appendix 5:
Appendix 5: Manipulation check

a. Manipulation check in main study

In our main study, we conducted manipulation checks using simple assessments:

The manipulation of awareness was tested via an attention check presented directly after the awareness text. While an attention question does not constitute a manipulation check per se, it served as a prerequisite to ensure participants had actually read the text, which was essential for the manipulation to take effect.

We also tested whether the personalization manipulation was successful by asking participants to indicate which celebrity they had seen in the video. Recognition was used as a proxy for familiarity, based on the assumption that participants are unlikely to have a preference for a celebrity they do not recognize. Consequently, we excluded participants who failed to recognize the celebrity shown to them.

Additional scales were not included in the main questionnaire, as the survey was already quite long and we wanted to prevent cognitive fatigue from overly complex items.

b. Manipulation check design follow-up study

After completing the main study, we decided to conduct a follow-up study. This one was only focusing on the manipulation checks for awareness and personalization. This decision was based on insights from the initial results, after which question raised whether the manipulation has been perceived as expected.

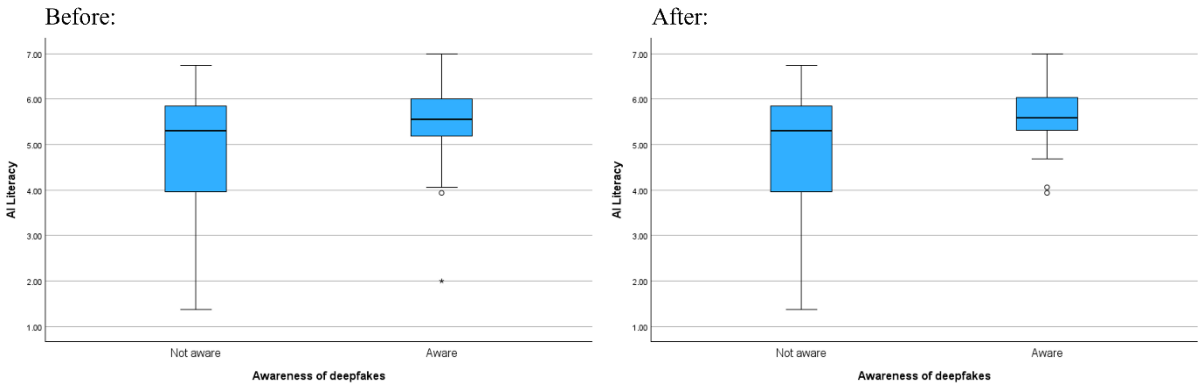
Thus to avoid premature or incorrect conclusions, we conducted this post-study to evaluate the experimental design.

The manipulation check initially included 108 participants, of which 26 did not complete the full questionnaire. Participants were mainly recruited via the platform SurveySwap, as well as through new contacts reached directly via Instagram and WhatsApp who had not participated in the main study. Nevertheless, the collection method was the same as in the main study.

The remaining 82 datasets were cleaned using the same procedure as in the main study. This included attention checks, celebrity recognition, brand knowledge assessments, and the exclusion of participants who recognized Neha Sharma, the celebrity used in the non-personalized groups.

We also identified and excluded one outlier based on a box plot analysis of awareness of deepfakes between the "aware" and "not aware" groups. The boxplots can be seen in figure a20, it shows how the removal of the outlier changed the distribution of the data.

Figure a 20: Boxplots of AI literacy by awareness of deepfakes before and after outlier removal



Note. T*he left boxplot shows the distribution including an extreme outlier in the aware group and the right boxplot presents the distribution after removing this outlier.*

The final sample had 48 participants, distributed as follows: Group 1: N = 16, Group 2: N = 12, Group 3: N = 12, Group 4: N = 8

Looking at the demographics, represented in table a 30 and table a 31, this sample had a similar distribution to the main study. Similar to the main study: the majority were female (87.5%), most respondents came from German-speaking countries (56.2%), most participants, 89.7% (compared to 90.0% in the main study) , were between 18 and 34 years old M = 24.79, SD = 6.22, Min = 18, Max = 44) and participants had completed higher education (62.5%).

The manipulation check followed the same design as the main study, but tested the variables AI literacy and perceived personalization. We also decided to include an item to assess whether participants perceived the text as a disclosure for deepfakes ("*I perceived the text at the beginning as a disclosure about the ad.*").

We used the 8-item Likert scale for perceived personalization from De Keyzer et al. (2022) and added 3 items adapted to our study design (e.g., "*The ad felt personal because it featured a celebrity I like*", "*The ad felt personal to me because it featured a celebrity I had indicated as my preference*", "*The ad felt like it was based on data collected about me*"). The resulting 11-item scale showed high internal reliability (Cronbach's Alpha = 0.951).

For the variable AI literacy, we did not find a suitable existing scale and therefore developed our own. We based it on the three domains used by Ashley et al. (2013) in their media literacy scale for newspapers. The authors named those three domains:

First *Authors and Audiences*, which Ashley et al. (2013) explained is the understanding that authors (in our case, brands) aim for profit, use persuasion, and target specific audiences. Then there is *Messages and Meaning,* which they explain to be the awareness that messages are constructed using specific production techniques to influence audience attitudes and behaviour. And finally the authors defined *Representation and Reality*, the knowledge of how media can shape perceptions of reality.

We used these domains to guide the development of our own items. However, the original developed scale from Ashley et al. (2013) was primarily focused on news articles. To address this, we also incorporated AI literacy competencies from Long and Magerko (2020), who classified these competencies into five aspects: "*What is AI?, What can AI do?, How does AI work?, How should AI be used? How do people perceive AI?*" (Long & Magerko, 2020, p. 3). We then developed items for each of these aspects (see appendix 6), by following the criteria proposed by DeVellis and Thorpe (2022, pp. 99–102) (see appendix 7).

An explorative factor analysis (principal Component with Varimax rotation), has shown that the items all relate to 3 main factors, just as proposed from Ashley et al. (2013), 3 items however did show significant cross loadings and were thus excluded. In table a 27 identifies the 3 items and shows the cross loadings of the 3 items that were excluded. The resulting 13 item scale has shown to have a high internal consistency (Cronbach's Alpha: .944). Table a 26 further shows the Cronbach's alpha for the 3 factors of the variable AI Literacy

Table a 26: Reliability: Internal consistency of the AI literacy subscales and manipulation check scales

Construct	Cronbach's α	Number of Items
AI Literacy – Factor 1: Messages and Meanings	.948	7
AI Literacy – Factor 2: Authors & Audiences	.922	4
AI Literacy – Factor 3: Representation & Reality	.772	2
AI Literacy (overall scale)	.944	13
Personalization (overall scale)	.951	11

Table a 27: Construct validity: Factor analysis AI literacy, rotated component matrix

PCAT item	Factor loading		
	1	2	3
Factor 1: Messages and Meanings			
AD01_01 I consider myself knowledgeable about Deepfakes.	.756	-	-
AD01_02 I know what Deepfakes are.	.836	-	-
AD01_03 I understand that Deepfakes use AI to mimic real people.	.830	-	-
AD01_04 I know what Deepfakes can do, like face swaps or voice imitation.	.842	-	-
AD01_13 I am aware that Deepfakes could be misused to deceive or harm people.	.746	-	-
AD01_14 I understand ethical concerns around consent and misinformation in Deepfake usage.	.865	-	-
AD01_15 I have seen public debates or news stories discussing Deepfakes.	**.576**	**-**	**.487**
AD01_16 I understand that someone decides exactly how a Deepfake should move, speak, or look.	.644	-	-
Factor 2: Authors and Audiences			
AD01_05 I know that companies use Deepfakes in advertising.	-	.702	.489
AD01_06 I am aware that advertisers use Deepfakes to change consumer behaviour.	-	.875	-
AD01_07 I know that Deepfakes can be used to tailor ads to individual people.	.420	.803	-
AD01_08 I know that advertisers can use Deepfakes to change a model's face based on my preferences.	-	.833	-
Factor 3: Representation and Reality			
AD01_09 I am not confident in my understanding of how Deepfakes work.	**-**	**-**	**-**
AD01_10 I understand the technical process behind Deepfakes.	**-**	**.637**	**.601**
AD01_11 I know that Deepfakes are generated using training data and neural networks.	-	-	.745
AD01_12 I understand that humans decide what data is used to train Deepfakes.	-	-	.761

Note. *Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser normalization. AD01_X is the items name in the excel codebook.*
Loadings below .40 were omitted. The factor analysis identified a clear three-factor structure, we decided to remove item AD01_15, AD01_09 and AD01_10, since we identified strong cross loadings over the 3 factors.

c. *Results of the manipulation check*

We tested whether the manipulations for awareness of deepfakes and personalization worked as intended. A t-test was conducted to compare the awareness and no-awareness groups, and a one-way ANOVA was used to test differences between the personalization and no-personalization groups.

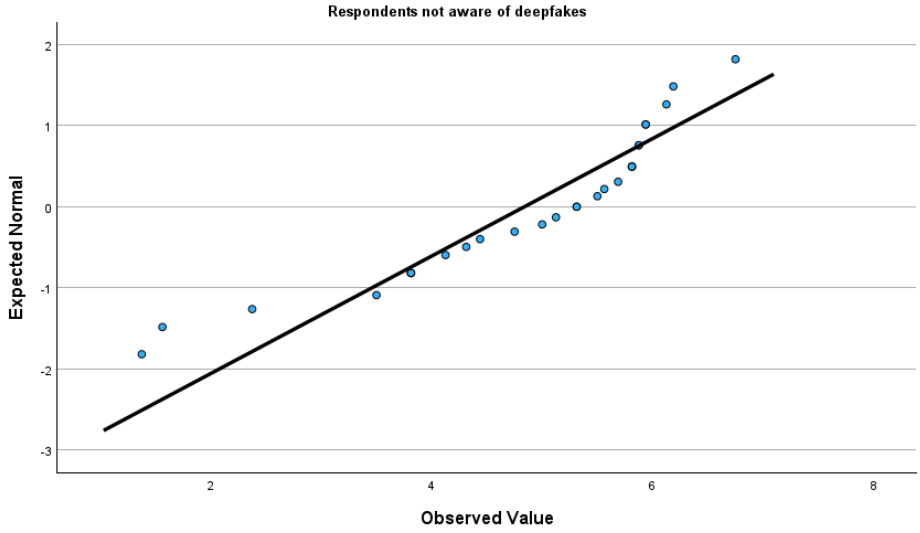
Results: both "manipulations for awareness" and "personalization" were significantly positively increased.

Awareness of deepfakes

For the variable awareness of deepfakes, we tested whether there is a measurable difference in AI literacy between Group 2 ($N = 20$, $M = 5.73$, $SD = .72$), who read the deepfake introduction text, and Group 1 ($N = 28$, $M = 5.73$, $SD = 1.48$), who did not. A t-test was conducted to determine whether this difference was significant.

The assumptions for the t-test were checked: the Shapiro–Wilk test showed that Group 2 was normally distributed ($W(20) = .95$, $p = .475$). However, for Group 1, the Shapiro–Wilk test was significant ($W(28) = .86$, $p = .001$), indicating a violation of the normality assumption. Nevertheless, the Q-Q plot in figure a21 generated via SPSS's Explore function figured an observable linear distribution, so we assumed that the normality assumption is confirmed.

Figure a 21: QQ plot of the variable AI literacy when respondents were not confronted with the included awareness text.



Note. *The plotted points follow well the diagonal line, thus we can assume normality.*

Since the Levene's test showed that the variances were heterogeneous ($F(1, 46) = 9.01$, $p = .004$), we chose to use the Welch t-test (equal variances not assumed). This test yielded a significant result ($t(41.45) = -2.20$, $p = .034$). The effect size was on a medium level ($d = -.58$), indicating a higher level of AI literacy in Group 2. We can therefore conclude that the deepfake awareness manipulation was effective. Table a 28 regroups the presented data of the Shapiro Wilk test, levene test and performed Welch t-test.

Table a 28: Independent samples t-test: Differences in AI literacy between awareness groups

Test	Group	n	M	SD	Statistic	df	p	95% CI [LL, UL]		Cohen's d
								LL	UL	
Shapiro–Wilk (Normality)	Group 1	28	5.02	1.48	.858	28	.001	-	-	-
	Group 2	20	5.73	0.72	.956	20	.475	-	-	-
Levene's Test (Homogeneity)	-	-	-	-	9.01	46	.004	-	-	-
Independent t-test	Equal variances not assumed	-	-	-	-2.20	41.45	.034	-1.36	-.058	-.579

Note. *Group 1 = low awareness; Group 2 = high awareness. The Shapiro Wilk test was significant which shows that the data in group1 is not normally distributed however the QQ plot shows a relatively well aligned scatterplot so we assume that normality is given.*

Personalisation

We performed the same analysis for the variable personalisation, an overview of the analysis is given in table a 29 also showing the descriptive data of the two groups. First, we performed a T-test between the two groups (Group 1: without personalisation/Group 2: with personalisation). The normality distribution was given (Shapiro-Wilk Test: Group 1: W(28) = .93, p = .069; Group 2: W(20) = .969, p = .739), but the variances were not homogeneously distributed (F(1, 46) = 5.99, p = .018). We therefore performed a Welch T-test, which turned out to be significant (t(29.33) = -3.52, p = .001). The effect size was large (d = 1.11), showing that the group that was shown the preferred celebrity (Group 2) had a higher perceived personalisation than Group 1.

Table a 29: Independent samples t-test: Differences in personalization between binary groups

Test	Group	n	M	SD	Statistic	df	p	95% CI		Cohen's d
								LL	UL	
Shapiro–Wilk (Normality)	Group 1	28	2.30	0.99	.932	28	.069	-	-	-
	Group 2	20	3.73	1.61	.969	20	.739	-	-	-
Levene's Test (Homogeneity)	-	-	-	-	5.99	46	.018	-	-	-
Independent t-test	Equal variances not assumed	-	-	-	-3.52	29.33	.001	-2.25	-0.60	-1.11

Note. *Group 1 = no personalization; Group 2 = personalization.*
A significant difference indicates that the manipulation successfully altered the perceived personalization level.

Personalization open question

Lastly we had also added an open-ended question at the end of the questionnaire. This was meant to get more detailed insight on how the respondents felt about the personalisation. The two groups (personalised vs. non-personalised) showed significantly different perceived personalisation. But thanks to the open-ended answers, we were able to see that the users did not really feel personally targeted by the ad. From the 20 respondents in the personalised group, 6 said they felt the ad was adapted to them:

"*I love make-up and these aesthetics*", "*Yes, because it featured a celebrity that I had indicated that I liked*".

Three respondents were unsure, but 11 said they did not feel targeted:

"*I felt it was explicitly made for this study and each participant gets the celebrity they voted*", "*No, because even though I like Lady Gaga, I do not really identify with the lifestyle: pink, luxury etc*.", "*No, it was not my style at all. Just because I like the celebrity I do not specially feel targeted*".

d. *Manipulation check data demographics*

Table a 30: Descriptive statistics for age (N = 48)

Statistic	Value
Mean (M)	24.79
Median	23.50
Standard Deviation (SD)	6.22
Variance	38.72
Minimum	18
Maximum	44
Skewness	1.59
Kurtosis	2.41

Table a 31: Demographic characteristics of the sample (N = 48)

Variable	Category	n	%
Gender	Male	4	8.3
	Female	42	87.5
	Non-binary	1	2.1
	Other	1	2.1
Country of residence	EU (German-speaking)	27	56.3
	EU (non-German-speaking)	5	10.4
	Non-EU countries	16	33.3
Education	Primary school	1	2.1
	High school diploma or equivalent	16	33.3
	Vocational training	1	2.1
	Bachelor's degree	16	33.3
	Master's degree	13	27.1
	Doctorate or higher	1	2.1

Appendix 6

Appendix 6: Scale building

a. Awareness of falsity scale

In order to build the scale awareness of falsity we used the three aspects proposed by Campbell et al. (2021) as a theoretical core: (1) product-related claims, (2) product-unrelated claims, and (3) presented reality.

For **product-related claims**, we chose to use the ad credibility scale from Kim, Ratneshwar, and Thorson (2017, as cited in Bruner II, 2019). Ad credibility refers to the extent to which consumers believe the claims in the ad about the brand or product are truthful (MacKenzie & Lutz, 1989). The concept of awareness of falsity, according to Campbell et al. (2021), seems to go beyond ad credibility, incorporating not just product-related claims but also unrelated claims and the presented reality. In our understanding the concept of ad credibility is similar to the first dimension of awareness of falsity: product related claims. Therefore, while ad credibility aligns with the first dimension, we decided to use the scale of ad credibility but to modify it to better reflect falsity: Unlike ad credibility, awareness of falsity excludes trust in the advertiser (MacKenzie & Lutz, 1989), focusing instead on the content of the ad itself. Hence, we used a modified version of the ad credibility scale for the first dimension, including items from Kim, Ratneshwar, and Thorson (2017, as cited in (Bruner II, 2019): *"This advertisement's product-related claim is generally truthful."; "This advertisement's product-related claim leaves one feeling accurately informed."; "This advertisement's product-related claim is believable."*. These items indicate how well and how correctly the product-related claims were perceived by the consumer. The three-item scale showed high reliability (Cronbach's $\alpha = 0.92$).

For **product-unrelated claims**, we focused on celebrity endorsement and adapted the semantic differential scale by Eisend et al. (2014), modifying it to assess the product-unrelated aspects of the celebrity in the ad (see appendix 9 for the full scale used). This scale had a Cronbach's α of 0.93.

For **presented reality**, we adapted a scale from Pfiffelmann et al. (2023), which also showed good reliability (Cronbach's $\alpha = .83$), although it contained only two items.

We conducted an exploratory factor analysis (principal component analysis with varimax rotation), as shown in table a 8 from appendix 2. The results showed that the items loaded clearly onto three factors. According to Kaiser's criterion, meaning that the significant factors have a eigenvalue above 1 (Bortz & Schuster, 2011), three components were retained, as the fourth had a variance of only .29 (<1), providing little additional information. This fits with the theoretical framework by Campbell et al., (2021). The factor loadings for the three dimensions were all above .848, with low cross-loadings (max = .307). The construct validity of the scale is given and we thus used this scale in our main study.

b. *AI literacy scale*

The following you can find the developed AI literacy scale, leaning towards Long and Magerko (2020, p. 2) definition. To ensure a clear, grammatical correct formulation of the items, an AI tool (ChatGPT) was used :

"What is AI?":

- I know what Deepfakes are
- I consider myself knowledgeable about Deepfakes

"What can AI do?":

- I understand that deepfakes use AI to mimic real people.
- I know what deepfakes can do, like face swaps or voice imitation.
- I know that companies use Deepfakes in advertising.
- I am aware that advertisers use Deepfakes to change consumer behaviour.
- I know that deepfakes can be used to tailor ads to individual people.
- I know that advertisers can use deepfakes to change a model's face based on my preferences.

"How does AI work?":

- I am not confident in my understanding of how Deepfakes work.
- I understand the technical process behind Deepfakes.
- I know that deepfakes are generated using training data and neural networks.
- I understand that humans decide what data is used to train deepfakes.

"How should AI be used?":

- I understand that someone decides exactly how a deepfake should move, speak, or look.
- I am aware that deepfakes could be misused to deceive or harm people.
- I understand ethical concerns around consent and misinformation in deepfake usage.

"How do people perceive AI?":

- I have seen public debates or news stories discussing deepfakes.

Appendix 7:

Appendix 7: Quality criteria used for the scale development

The following you can find the quality criteria's we have used, according to DeVellis and Thorpe (2022)

- Avoid double-barrelled items (include only one idea at a time)
- Avoid ambiguous wording (e.g., pronoun references, misplaced modifiers)
- Avoid multiple negatives
- Avoid biased or leading formulations (e.g., acquiescence bias)
- Avoid exceptionally lengthy items

Appendix 8:

Appendix 8: Construct validity of analysis 2

Table a 32: Construct validity: Factor analysis of perceived realism (rotated component matrix)

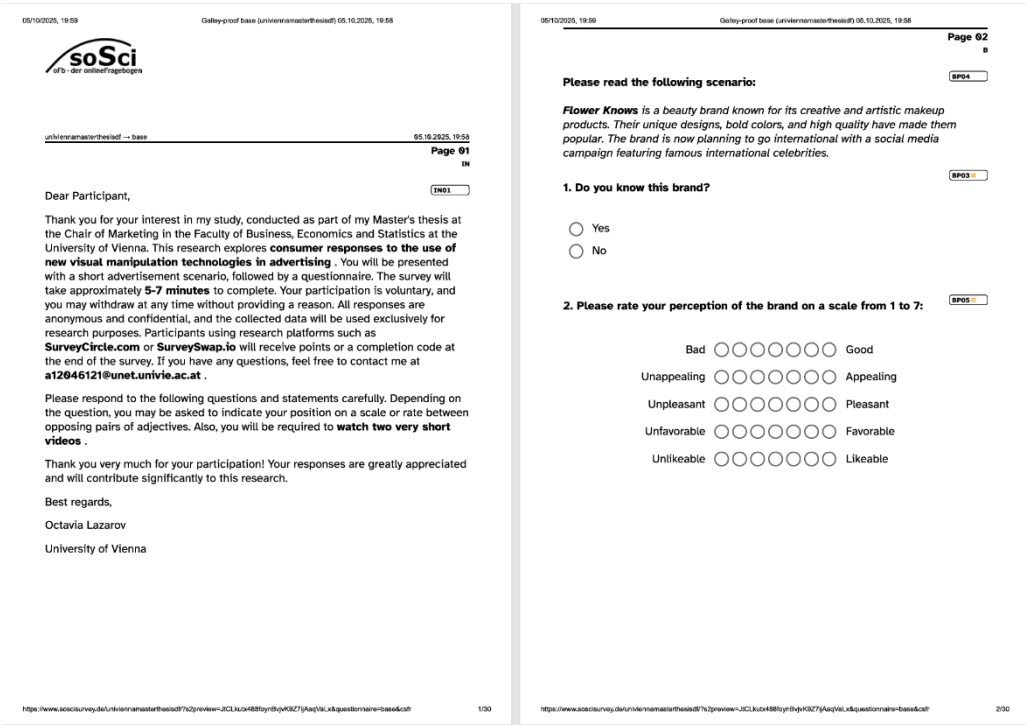
PCAT item	Factor loading	
	1	2
Factor 1: Perceived Likelihood of realism		
Plausibility	.808	-
Factuality	.719	.455
Typicality	.894	-
Factor 2: Portrayal of realism		
Perceptual Quality	-	.784
Narrative Consistency	-	.874

Note. *Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser normalization. Factor loadings < .40 were suppressed.*

Appendix 9:

Appendix 9: Full questionnaire of the experiment

The following screenshots are taken from the SoSci Survey PDF report and document the full questionnaire used in the main experiment. The layout shown here corresponds to the report export and not to the exact interface seen by participants.



05/10/2025, 19:59 Galley-proof base (univiennamasterthesisdf) 05.10.2025, 19:58

soSci
oFb - der onlineFragebogen

univiennamasterthesisdf → base 05.10.2025, 19:58

Page 01
IN

IN01

Dear Participant,

Thank you for your interest in my study, conducted as part of my Master's thesis at the Chair of Marketing in the Faculty of Business, Economics and Statistics at the University of Vienna. This research explores **consumer responses to the use of new visual manipulation technologies in advertising** . You will be presented with a short advertisement scenario, followed by a questionnaire. The survey will take approximately **5-7 minutes** to complete. Your participation is voluntary, and you may withdraw at any time without providing a reason. All responses are anonymous and confidential, and the collected data will be used exclusively for research purposes. Participants using research platforms such as **SurveyCircle.com** or **SurveySwap.io** will receive points or a completion code at the end of the survey. If you have any questions, feel free to contact me at **a12046121@unet.univie.ac.at** .

Please respond to the following questions and statements carefully. Depending on the question, you may be asked to indicate your position on a scale or rate between opposing pairs of adjectives. Also, you will be required to **watch two very short videos** .

Thank you very much for your participation! Your responses are greatly appreciated and will contribute significantly to this research.

Best regards,

Octavia Lazarov

University of Vienna

https://www.soscisurvey.de/univiennamasterthesisdf/?s2preview=JtCLkutx488foynBvjvK8Z7jAaqVaLx&questionnaire=base&cstr 1/30

05/10/2025, 19:59 Galley-proof base (univiennamasterthesisdf) 05.10.2025, 19:58

Page 02
B

BP04

Please read the following scenario:

Flower Knows *is a beauty brand known for its creative and artistic makeup products. Their unique designs, bold colors, and high quality have made them popular. The brand is now planning to go international with a social media campaign featuring famous international celebrities.*

BP03

1. Do you know this brand?

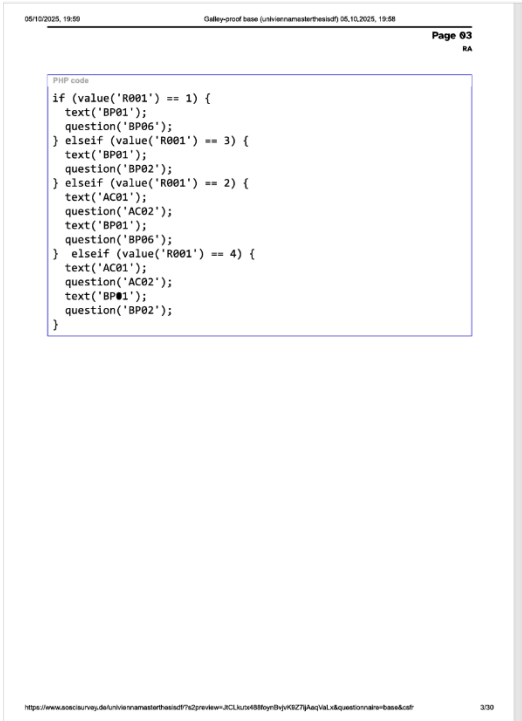
- ○ Yes
- ○ No

BP05

2. Please rate your perception of the brand on a scale from 1 to 7:

Bad	○○○○○○○	Good
Unappealing	○○○○○○○	Appealing
Unpleasant	○○○○○○○	Pleasant
Unfavorable	○○○○○○○	Favorable
Unlikeable	○○○○○○○	Likeable

https://www.soscisurvey.de/univiennamasterthesisdf/?s2preview=JtCLkutx488foynBvjvK8Z7jAaqVaLx&questionnaire=base&cstr 2/30



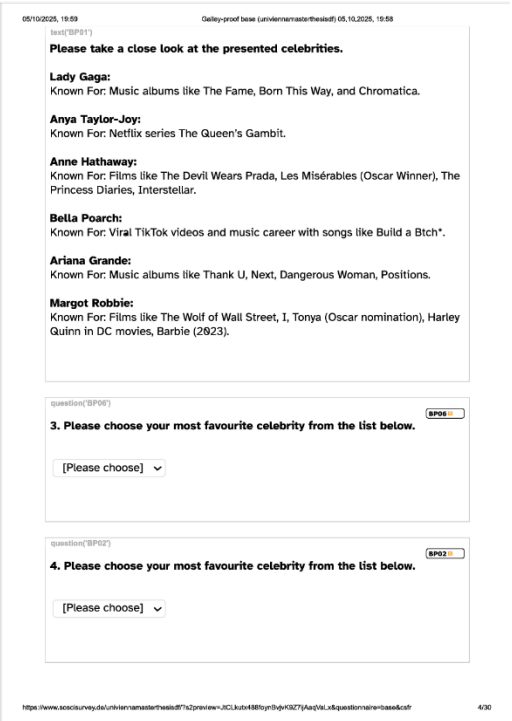
05/10/2025, 19:59 Galley-proof base (univiennamasterthesisdf) 05.10.2025, 19:58

Page 03
RA

PHP code

```
if (value('R001') == 1) {
  text('BP01');
  question('BP06');
} elseif (value('R001') == 3) {
  text('BP01');
  question('BP02');
} elseif (value('R001') == 2) {
  text('AC01');
  question('AC02');
  text('BP01');
  question('BP06');
}  elseif (value('R001') == 4) {
  text('AC01');
  question('AC02');
  text('BP01');
  question('BP02');
}
```

https://www.soscisurvey.de/univiennamasterthesisdf/?s2preview=JtCLkutx488foynBvjvK8Z7jAaqVaLx&questionnaire=base&cstr 3/30



05/10/2025, 19:59 Galley-proof base (univiennamasterthesisdf) 05.10.2025, 19:58

text('BP01')

Please take a close look at the presented celebrities.

Lady Gaga:
Known For: Music albums like The Fame, Born This Way, and Chromatica.

Anya Taylor-Joy:
Known For: Netflix series The Queen's Gambit.

Anne Hathaway:
Known For: Films like The Devil Wears Prada, Les Misérables (Oscar Winner), The Princess Diaries, Interstellar.

Bella Poarch:
Known For: Viral TikTok videos and music career with songs like Build a Btch*.

Ariana Grande:
Known For: Music albums like Thank U, Next, Dangerous Woman, Positions.

Margot Robbie:
Known For: Films like The Wolf of Wall Street, I, Tonya (Oscar nomination), Harley Quinn in DC movies, Barbie (2023).

question('BP06')

BP06

3. Please choose your most favourite celebrity from the list below.

[Please choose]

question('BP02')

BP02

4. Please choose your most favourite celebrity from the list below.

[Please choose]

https://www.soscisurvey.de/univiennamasterthesisdf/?s2preview=JtCLkutx488foynBvjvK8Z7jAaqVaLx&questionnaire=base&cstr 4/30

text('AC01')

AC01

5. Deepfake Technology: Redefining Advertising

The following scenario presents an extract from an article on new technologies in advertising. Please read it carefully.

*Advertisers today are increasingly using **deepfakes** in their campaigns to create hyper-personalized ads on social media platforms. A deepfake is a type of artificial intelligence (AI) technology that manipulates video or audio to make it look like someone is saying or doing something they never actually did. For example, a deepfake ad could feature your favourite celebrity promoting a product, even though it is not actually the same person. Someone else might see the same advertisment film, but featuring the face of a his prefered celebrity. Brands ensure that all celebrities involved have consented to the use of their faces in order to avoid legal complications.*

question('AC02')

AC02

6. In this text, what technology do advertisers use nowadays to achieve hyper-personatization?

- ◯ Deepfake
- ◯ Photoshop
- ◯ CGI

Page 04
LG

VA01

Please watch the following **first** video attentively. It is an advertisement for Flower Knows products, made for social media.

Flower Knows - Advertisment

VA12

Please watch the **second** video attentively.

Flower Knows - Advertisment

Page 05
MR

VA06

Please watch the following **first** video attentively. It is an advertisement for Flower Knows products, made for social media.

Flower Knows - Advertisment

VA14

Please watch the **second** video attentively.

Flower Knows - Advertisment

Page 11
CE

VA10

The two videos you watched were **identical**. This was done to ensure that you watched the video twice for attention purposes, so everyone gives it the same level of focus.

VA08

7. Please select the correct brand name that appeared in the video:

- ◯ Flower Knows
- ◯ Florette
- ◯ Dior

VA09

8. Please select the celebrity's name presented in the video.

[Please choose]

CE01 CE

9. On a scale from 1 to 7, please evaluate the celebrity as an endorser (i.e. representative or advocate) of this brand on the following attributes:

The celebrity as an endorser of this brand can be described as...

Appropriate	◯◯◯◯◯◯◯	Inappropriate
Fits	◯◯◯◯◯◯◯	Does not fit
Effective	◯◯◯◯◯◯◯	Not effective
Poor endorser	◯◯◯◯◯◯◯	Strong endorser

Page 12
PV

PV01

10. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The ad showed something that could possibly happen in real life.	◯	◯	◯	◯	◯	◯	◯
The event in the ad portrayed possible real-life situations.	◯	◯	◯	◯	◯	◯	◯
The story in the ad could actually happen in real life.	◯	◯	◯	◯	◯	◯	◯
Never in real life would what was shown in the ad happen.	◯	◯	◯	◯	◯	◯	◯
Real people would do the things shown in the ad.	◯	◯	◯	◯	◯	◯	◯

PV02

11. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Not many people are likely to experience the event portrayed in the ad.	◯	◯	◯	◯	◯	◯	◯
The ad portrayed an event that happens to a lot of people.	◯	◯	◯	◯	◯	◯	◯
What happened in the ad is what happens to people in the real world.	◯	◯	◯	◯	◯	◯	◯

PV03

12. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The ad was based on facts.	◯	◯	◯	◯	◯	◯	◯
The ad showed something that had really happened.	◯	◯	◯	◯	◯	◯	◯
What was shown in the ad had actually happened.	◯	◯	◯	◯	◯	◯	◯

PV04

13. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The ad showed a coherent story.	○	○	○	○	○	○	○
The story portrayed in the ad was consistent.	○	○	○	○	○	○	○
Parts of the ad were contradicting each other.	○	○	○	○	○	○	○
The story portrayed in the ad made sense.	○	○	○	○	○	○	○
The event in the ad had a logical flow.	○	○	○	○	○	○	○

PV05

14. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The visual elements of the ad were realistic.	○	○	○	○	○	○	○
The audio elements of the ad were realistic.	○	○	○	○	○	○	○
The acting in the ad was realistic.	○	○	○	○	○	○	○
The scenes in the ad were realistic.	○	○	○	○	○	○	○
I felt that the overall production elements of the ad were realistic.	○	○	○	○	○	○	○

Page 13
AF

15. Please indicate how much you agree with the following statements: AF01

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I can rely on this advertisement to provide the truth.	○	○	○	○	○	○	○
The aim of this advertisement is to inform the consumer.	○	○	○	○	○	○	○
I believe this advertisement is informative.	○	○	○	○	○	○	○
This advertisement is truthful.	○	○	○	○	○	○	○
This advertisement provides reliable information about the quality and performance of the product.	○	○	○	○	○	○	○
This advertisement communicates the truth well.	○	○	○	○	○	○	○
This advertisement presents a true picture of the product being advertised.	○	○	○	○	○	○	○
I feel I've been accurately informed after viewing this advertisement.	○	○	○	○	○	○	○
This advertisement provides consumers with essential information.	○	○	○	○	○	○	○

16. Please recall the product claims presented in the advertisement and indicate the extent to which you agree with the following statements. AF02

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
This advertisement's product-related claim is generally truthful.	○	○	○	○	○	○	○
This advertisement's product-related claim leaves one feeling accurately informed.	○	○	○	○	○	○	○
This advertisement's product-related claim is believable.	○	○	○	○	○	○	○

17. On a scale from 1 to 7, please rate the celebrity's presence in the advertisement on the following dimensions: AF03

The presence of the celebrity is...

Unconvincing	○ ○ ○ ○ ○	Convincing
Unbelievable	○ ○ ○ ○ ○	Believable
Dishonest	○ ○ ○ ○ ○	Honest

18. On a scale from 1 to 7, please rate the advertisement: AF04

The scenario in this advertisement is...

Not realistic	○ ○ ○ ○ ○	Realistic
Could exist unlikely in real life	○ ○ ○ ○ ○	Likely in real life

Page 14
PI

19. Please indicate how much you agree with the following statements. PI01

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I would like to try this product.	○	○	○	○	○	○	○
I would buy this product if I happened to see it.	○	○	○	○	○	○	○
I would actively seek out this product in a store to purchase it.	○	○	○	○	○	○	○

Page 15
DQ

20. Please indicate how much you agree with the following statements. DQ01

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
The face of the person in the video (or parts of it) was distorted.	○	○	○	○	○	○	○
The facial features of the person in the video changed during the video.	○	○	○	○	○	○	○
The face of the person in the video (or parts of it) was blurry.	○	○	○	○	○	○	○
The person in the video behaved authentically.	○	○	○	○	○	○	○
The mouth of the person in the video was moving strangely.	○	○	○	○	○	○	○
The gestures displayed by the person in the video did not seem natural.	○	○	○	○	○	○	○

Page 16
D

21. What is your gender? DI01

- ○ Male
- ○ Female
- ○ Non-binary
- ○ Other

22. How old are you? DI02

23. What is your highest level of education? DI03

- ○ No formal education
- ○ Primary school
- ○ High school diploma or equivalent
- ○ Vocational training
- ○ Bachelor's degree
- ○ Master's degree
- ○ Doctorate or higher

24. Where do you currently live? (country) DI04

Last Page

Thank you for completing this questionnaire!

Thank you for your time and effort in completing this survey. Your responses are very helpful to this research.

I would like to inform you that all the videos shown during this survey are **deepfakes**. Special thanks to the artist **Yulia Fomenko (@fomenkojulli on Instagram)** for granting permission to use her original video as the basis for this study. The deepfake videos were created using DeepFaceLab and serve exclusively scientific purposes.

If you are using research platforms, you will receive your completion code on the next page.

Your answers were transmitted, you may close the browser window or tab now.

Löse den folgenden Survey Code unter https://www.surveycircle.com ein und erhalte durch SurveyCircle kostenlos Teilnehmer für deine Studie: F18L-XYY2-68RV-SL4V

Survey Code mit einem Klick einlösen: https://www.surveycircle.com/F18L-XYY2-68RV-SL4V/

B. Sc. Octavia Lazarov, Universität Wien – 2025

Appendix 10:

Appendix 10: Full questionnaire of the manipulation check

The following screenshots are taken from the SoSci Survey PDF report and document the full questionnaire used in the manipulation check. The layout shown here corresponds to the report export and not to the exact interface seen by participants.

05/10/2025, 20:00 Galley-proof base (univiennamasterthesisdf2) 05.10.2025, 20:00

soSci
oFb - der onlineFragebogen

univiennamasterthesisdf2 → base 05.10.2025, 20:00

Page 01
IN

IN01

Dear Participant,

Thank you for your interest in my study, conducted as part of my Master's thesis at the Chair of Marketing in the Faculty of Business, Economics and Statistics at the University of Vienna. This research explores **consumer responses to the use of new visual manipulation technologies in advertising** . You will be presented with a short advertisement scenario, followed by a questionnaire. The survey will take approximately **5-7 minutes** to complete. Your participation is voluntary, and you may withdraw at any time without providing a reason. All responses are anonymous and confidential, and the collected data will be used exclusively for research purposes. Participants using research platforms such as **SurveyCircle.com** or **SurveySwap.io** will receive points or a completion code at the end of the survey. If you have any questions, feel free to contact me at **a12046121@unet.univie.ac.at** .

Please respond to the following questions and statements carefully. Depending on the question, you may be asked to indicate your position on a scale or rate between opposing pairs of adjectives. Also, you will be required to **watch two very short videos** .

Thank you very much for your participation! Your responses are greatly appreciated and will contribute significantly to this research.

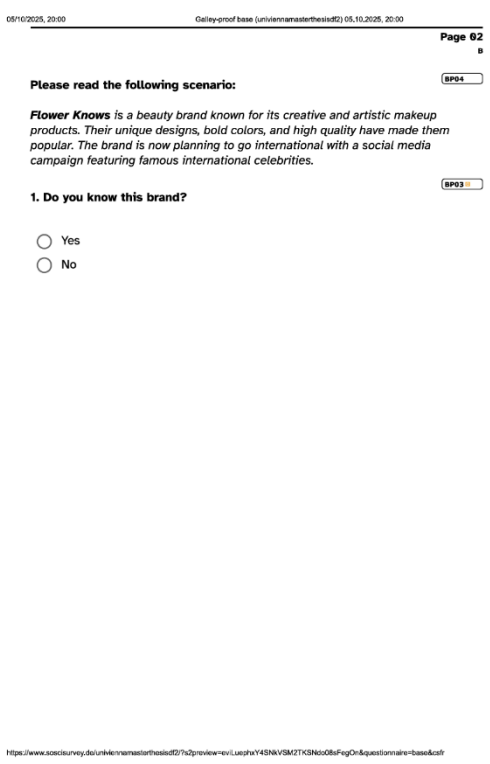
Best regards,

Octavia Lazarov

P.S: This survey contains Karma to get free survey responses at SurveySwap.io

University of Vienna

https://www.soscisurvey.de/univiennamasterthesisdf2/?s2preview=evILuephxY4SNkVSM2TKSNdo08sFegOn&questionnaire=base&csfr 1/36



05/10/2025, 20:00 Galley-proof base (univiennamasterthesisdf2) 05.10.2025, 20:00

Page 02
B

BP04

Please read the following scenario:

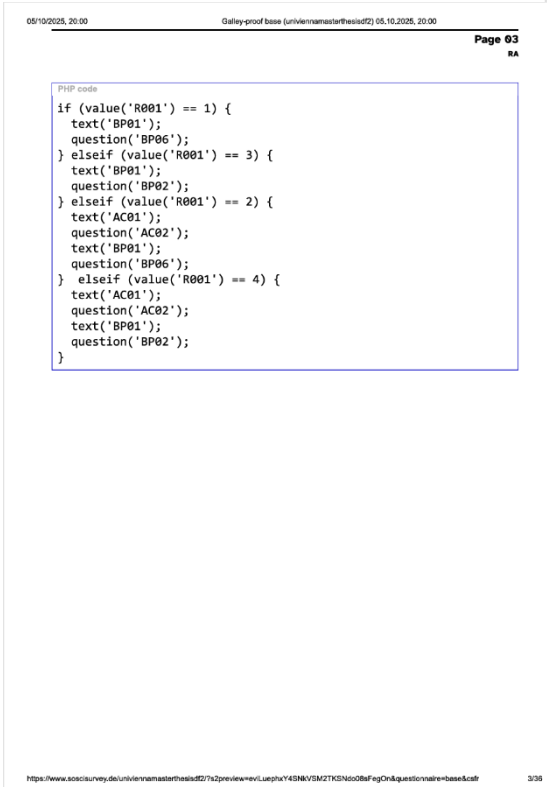
Flower Knows *is a beauty brand known for its creative and artistic makeup products. Their unique designs, bold colors, and high quality have made them popular. The brand is now planning to go international with a social media campaign featuring famous international celebrities.*

BP03

1. Do you know this brand?

- ◯ Yes
- ◯ No

https://www.soscisurvey.de/univiennamasterthesisdf2/?s2preview=evILuephxY4SNkVSM2TKSNdo08sFegOn&questionnaire=base&csfr 2/36



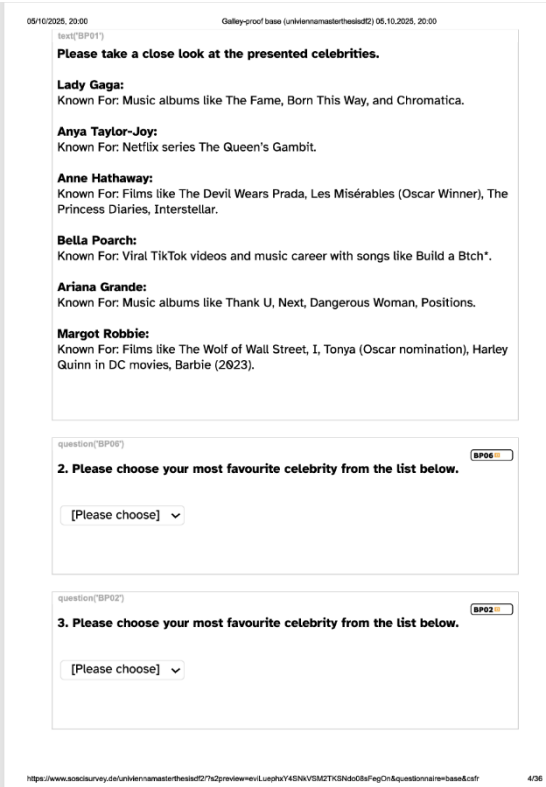
05/10/2025, 20:00 Galley-proof base (univiennamasterthesisdf2) 05.10.2025, 20:00

Page 03
RA

PHP code

```
if (value('R001') == 1) {
  text('BP01');
  question('BP06');
} elseif (value('R001') == 3) {
  text('BP01');
  question('BP02');
} elseif (value('R001') == 2) {
  text('AC01');
  question('AC02');
  text('BP01');
  question('BP06');
}  elseif (value('R001') == 4) {
  text('AC01');
  question('AC02');
  text('BP01');
  question('BP02');
}
```

https://www.soscisurvey.de/univiennamasterthesisdf2/?s2preview=evILuephxY4SNkVSM2TKSNdo08sFegOn&questionnaire=base&csfr 3/36



05/10/2025, 20:00 Galley-proof base (univiennamasterthesisdf2) 05.10.2025, 20:00

text('BP01')

Please take a close look at the presented celebrities.

Lady Gaga:
Known For: Music albums like The Fame, Born This Way, and Chromatica.

Anya Taylor-Joy:
Known For: Netflix series The Queen's Gambit.

Anne Hathaway:
Known For: Films like The Devil Wears Prada, Les Misérables (Oscar Winner), The Princess Diaries, Interstellar.

Bella Poarch:
Known For: Viral TikTok videos and music career with songs like Build a Btch*.

Ariana Grande:
Known For: Music albums like Thank U, Next, Dangerous Woman, Positions.

Margot Robbie:
Known For: Films like The Wolf of Wall Street, I, Tonya (Oscar nomination), Harley Quinn in DC movies, Barbie (2023).

question('BP06')

BP06

2. Please choose your most favourite celebrity from the list below.

[Please choose]

question('BP02')

BP02

3. Please choose your most favourite celebrity from the list below.

[Please choose]

https://www.soscisurvey.de/univiennamasterthesisdf2/?s2preview=evILuephxY4SNkVSM2TKSNdo08sFegOn&questionnaire=base&csfr 4/36

text('AC01')

AC01

4. Deepfake Technology: Redefining Advertising

The following scenario presents an extract from an article on new technologies in advertising. Please read it carefully.

*Advertisers today are increasingly using **deepfakes** in their campaigns to create hyper-personalized ads on social media platforms. A deepfake is a type of artificial intelligence (AI) technology that manipulates video or audio to make it look like someone is saying or doing something they never actually did. For example, a deepfake ad could feature your favourite celebrity promoting a product, even though it is not actually the same person. Someone else might see the same advertisment film, but featuring the face of a his prefered celebrity. Brands ensure that all celebrities involved have consented to the use of their faces in order to avoid legal complications.*

question('AC02')

AC02

5. In this text, what technology do advertisers use nowadays to achieve hyper-personalization?

- ◯ Deepfake
- ◯ Photoshop
- ◯ CGI

VA01

Please watch the following **first** video attentively. It is an advertisement for Flower Knows products, made for social media.

Flower Knows - Advertisment

VA12

Please watch the **second** video attentively.

Flower Knows - Advertisment

VA06

Please watch the following **first** video attentively. It is an advertisement for Flower Knows products, made for social media.

Flower Knows - Advertisment

VA14

Please watch the **second** video attentively.

Flower Knows - Advertisment

Page 11
CE

VA19

The two videos you watched were **identical**. This was done to ensure that you watched the video twice for attention purposes, so everyone gives it the same level of focus.

VA08

6. Please select the correct brand name that appeared in the video:

- ◯ Flower Knows
- ◯ Florette
- ◯ Dior

VA09

7. Please select the celebrity's name presented in the video.

[Please choose]

Page 12
P

P001

8. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
This ad is tailored to my preferences.	◯	◯	◯	◯	◯	◯	◯
I believe this ad is customized to my needs.	◯	◯	◯	◯	◯	◯	◯
This ad was targeted at me as a unique individual.	◯	◯	◯	◯	◯	◯	◯
I believe that this ad is customized to my characteristics.	◯	◯	◯	◯	◯	◯	◯
This ad	◯	◯	◯	◯	◯	◯	◯

was personalized according to my profile.	○	○	○	○	○	○	○
There was personal information in the ad.	○	○	○	○	○	○	○
The ad was targeted at me.	○	○	○	○	○	○	○
I could recognize myself in the group the ad was targeted at.	○	○	○	○	○	○	○
The ad felt personal because it featured a celebrity I like.	○	○	○	○	○	○	○

The ad felt personal to me because it featured a celebrity I had indicated as my preference.	○	○	○	○	○	○	○
The ad felt like it was based on data collected about me.	○	○	○	○	○	○	○

9. Did you feel like the advertisement was personally designed for you? Why or why not? P002

Page 13
AD

AD01

10. Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I know what Deepfakes are.	○	○	○	○	○	○	○
I consider myself knowledgeable about Deepfakes.	○	○	○	○	○	○	○
I understand that Deepfakes use AI to mimic real people.	○	○	○	○	○	○	○
I know what Deepfakes can do, like face swaps or voice imitation.	○	○	○	○	○	○	○
I know that companies use	○	○	○	○	○	○	○

Deepfakes in advertising.

I am aware that advertisers use Deepfakes to change consumer behaviour.	○	○	○	○	○	○	○
I know that Deepfakes can be used to tailor ads to individual people.	○	○	○	○	○	○	○
I know that advertisers can use Deepfakes to change a model's face based on my preferences.	○	○	○	○	○	○	○

I am not confident in my understanding of how Deepfakes work.	○	○	○	○	○	○	○
I understand the technical process behind Deepfakes.	○	○	○	○	○	○	○
I know that Deepfakes are generated using training data and neural networks.	○	○	○	○	○	○	○
I understand that humans decide what data is used to train Deepfakes.	○	○	○	○	○	○	○

I am aware that Deepfakes could be misused to deceive or harm people.	○	○	○	○	○	○	○
I understand ethical concerns around consent and misinformation in Deepfake usage.	○	○	○	○	○	○	○
I have seen public debates or news stories discussing Deepfakes.	○	○	○	○	○	○	○
I understand that someone decides exactly how a Deepfake should	○	○	○	○	○	○	○

move, speak, or look.

Page 14
DC

PHP code

```
if (value('R001') == 2) {
  question('D001');
} elseif (value('R001') == 4) {
  question('D001');
} elseif (value('R001') == 1) {
  goToPage('D');
} elseif (value('R001') == 3) {
  goToPage('D');
}
```


question('D001')

D001

11. Thinking back to the introductory text and the video you just saw.

Below is the text that was presented to you before the videos.

Deepfake Technology: Redefining Advertising

Advertisers today are increasingly using deepfakes in their campaigns to create hyper-personalized ads on social media platforms. A deepfake is a type of artificial intelligence (AI) technology that manipulates video or audio to make it look like someone is saying or doing something they never actually did. For example, a deepfake ad could feature your favourite celebrity promoting a product, even though it is not actually the same person. Someone else might see the same advertisment film, but featuring the face of a his prefered celebrity. Brands ensure that all celebrities involved have consented to the use of their faces in order to avoid legal complications.

In the following items, we refer to the term disclosure.

Disclosure Definition: it refers to a short message included inside an ad that explains how the ad was made or what kind of content to expect (e.g., if the video was created or altered using AI).

Please indicate how much you agree with the following statements:

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I perceived the text at the beginning as a disclosure about the ad.	○	○	○	○	○	○	○
The introductory text appeared to provide general information about the upcoming video.	○	○	○	○	○	○	○
After reading the text, I expected the video to contain deepfake content.	○	○	○	○	○	○	○

I interpreted the text as a factual explanation about AI and deepfakes.	○	○	○	○	○	○	○
I understood that the video could have been either a deepfake or a real video.	○	○	○	○	○	○	○

12. What is your gender? DI01

- Male
- Female
- Non-binary
- Other

13. How old are you? DI02

14. What is your highest level of education? DI03

- No formal education
- Primary school
- High school diploma or equivalent
- Vocational training
- Bachelor's degree
- Master's degree
- Doctorate or higher

15. Where do you currently live? (country) DI04

Thank you for completing this questionnaire!

Thank you for your time and effort in completing this survey. Your responses are very helpful to this research.

I would like to inform you that all the videos shown during this survey are **deepfakes**. Special thanks to the artist **Yulia Fomenko (@fomenkojulli on Instagram)** for granting permission to use her original video as the basis for this study. The deepfake videos were created using DeepFaceLab and serve exclusively scientific purposes.

Your answers were transmitted, you may close the browser window or tab now.

The following code gives you Karma that can be used to get free research participants at SurveySwap.io.

Go to: https://surveyswap.io/sr/YIH2-QT07-KULP

Or, alternatively, enter the code manually: YIH2-QT07-KULP

B. Sc. Octavia Lazarov, Universität Wien – 2025

Appendix 11:
Appendix 11: The Stimuli

The stimulus chosen was a video that had to meet several criteria: First, it needed to match the brand, *Flower Knows,* aesthetic (pink, floral, girlish, fairy-tale/princess-like). Second, the actor's face needed to be well visible, meaning no hair or objects in front of it. Third, we needed a video where the actor wore a clear wig. Because we believe that some celebrities have hairstyles tied to their identity. That's why we chose a video from a well-known fashion influencer called Yulia Fomenko (**@fomenkojulli)**. She was so kind to agree, that we use her video in the context of this master thesis. The video features her in a baroque-style royal outfit and a large baroque wig, showing how she celebrates her birthday. Link to the original video: **Video**. Her face remains centred in the frame, with expressive facial movements. We edited the video to make it look like she was gifting herself *Flower Knows* makeup for her birthday. We added product claims and the *Flower Knows* logo.

We then applied a deepfake, replacing her face with those of selected celebrities. The selection was based on specific criteria: First the celebrities had to be widely known in Western culture. Further, while different celebrities have different public images, we aimed to offer a diverse selection so that the ad could still be seen as a hyper-personalized experience. For the non-personalized deepfake, we chose a celebrity not widely known in the West: Indian actress Neha Sharma. The last selection criterion was skin tone. We had to choose celebrities whose skin tone closely matched that of the original influencer. This was due to technical limitations, as we were not able to alter the skin tone of the entire video without compromising the deepfake's visual quality. The last criterion we had, was the availability of training data for each celebrity. While it's possible to create custom training data from interviews or other videos, the dataset must be large and well-cleaned to produce high-quality deepfakes. We thus chose to use the datasets made available from the website **deepfakevfx.com**, which offers training sets for a wide range of celebrities. In order to have enough data to train the AI, we chose to use datasets with a minimum of 5,000 images.

In the following table a 33 can be seen the links to the respective stimuli:

Table a 33: Celebrity videos and their respective YouTube link access (only accessible via the link/do not share)

VIDEO	LINK
Video Lady Gaga:	https://youtube.com/shorts/_tHuzk5YgCA
Video Anya Taylor-Joy:	https://youtube.com/shorts/26SJmG_wE50
Video Anne Hathaway:	https://youtube.com/shorts/uKRIZvM9HUg
Video Bella :	https://youtube.com/shorts/20PQBI39olE
Video Ariana Grande:	https://youtube.com/shorts/He0l-cHOtPs
Video Margot Robbie:	https://youtube.com/shorts/lmesy3tqy34
Video Neha Sharma:	https://youtube.com/shorts/L0Uh0CL1KwI

The Figure a 22 is showing an example screenshot of one of the created deepfakes compared to a screenshot from the original Video from Yulia Fomenko (@fomenkojulli). You can find her video here: https://www.instagram.com/reel/DB6SEp_xzOw/?igsh=aG16YnI1czFzcGJm

Figure a 22: Example of the original video vs the deepfake version

Note. *On the left side is the screenshot of the original video by Fomenko (2024), who gave us permission to use this video in the context of this master thesis. On the right side is one of the deepfakes, featuring Anya Taylor-Joy.*

Appendix 12:
Appendix 12: Definitions

Neural Networks:

A simple network has one input layer, one output layer, and one or two hidden layers. These layers are linked by weighted connections. Consider an image of a white number on a black background. The input layer contains many nodes. Each node represents a pixel. A white pixel activates its node strongly, a grey pixel gives a medium value, and a black pixel gives a low value. In this case the value of each node in the input layer depends on the pixel's brightness. These values are multiplied by their weighted connections, which represent the importance associated to each node. Nodes in the first hidden layer connect to several input nodes. An activation function is applied to produce the value of the hidden node, which is the sum of the weighted values of each node from the input layer connected to the node of the hidden layer and a bias term. The same process continues until the output layer is reached. The output node with the highest value represents the network's prediction, for example, the digit it believes is shown in the image. During training, machine learning, the network adjusts the weights of its nodes after each example to improve predictions. Standard neural networks have one or two hidden layers. (Karagiannis & Telesko, 2001)

Generative AI:

Generative AI includes all AI models, generally trained on very larges amounts of data in order to generate new content in response to human input, like text to image (Kumar & Kapoor, 2024). These models generate the output based on probabilities that have been established in the training phase, the AI will then propose a continuation or composition that has the highest probability to fit the prompt inputted (Matz et al., 2024). Some generative AI examples could be Midjourney, producing images, ChatGPT, producing text or Sora, producing videos.

Source Credibility:

Source credibility, so Powers et al. (2023) is as how trustworthy and reliable the sender of a message seems to be. Then following Ohanian' findings from 1990, Powers et al. (2023) further explained that source credibility is a construct build up from 3 sub dimensions including perceived trustworthiness, attractiveness of the source and expertise.

Cognitive dissonance:

The theory says that people naturally prefer consistency, they experience discomfort when exposed to contradictions, in order to reduce this discomfort the person will change his attitude or behaviours to restore consistencies. For example when a person is smoking but knows that smoking is bad for his health, there will be a cognitive dissonance between the persons actual behaviour and their belief. In order to reduce the inconsistency the person can choose between 3 dissonance reducing strategies. First eliminating the behaviour that creates the inconstancy: he could simply stop smoking. The second one is adding, adding new explications to the belief as like looking for people that even though smoked lived a long life. Or lastly changing the belief and question the veracity of the statement that smoking is bad for the health. (Solomon, 2017)