

**Research Bank**

Journal article

**The rise of deepfakes : A conceptual framework and research agenda for marketing**

**Whittaker, Lucas, Letheren, Kate and Mulcahy, Rory**

This is an Accepted Manuscript of an article published as:

Whittaker, L., Letheren, K. and Mulcahy, R. (2021). The rise of deepfakes : A conceptual framework and research agenda for marketing. *Australasian Marketing Journal*, 29(3), pp. 204-214. <https://doi.org/10.1177/1839334921999479>

This work is licensed under [Creative Commons Attribution-NonCommercial 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/).

# **The Rise of Deepfakes: A conceptual framework and research agenda for marketing**

## **Abstract**

Deepfakes, digital content created via machine learning, a form of artificial intelligence technology, are generating interest amongst marketers and the general population alike and are often portrayed as a ‘phantom menace’ in the media. Despite relevance to marketing theory and practice, deepfakes – and the opportunities for benefit or deviance they provide – are little understood or discussed. This paper introduces deepfakes to the marketing literature and proposes a typology, conceptual framework, and associated research agenda, underpinned by theorizing based on balanced centrality, to guide the future investigation of deepfakes in marketing scholarship. The paper makes an argument for balance (i.e. situations where all stakeholders benefit) and it is hoped that this paper may provide a foundation for future research and application of deepfakes as ‘a new hope’ for marketing.

**Keywords:** Artificial intelligence, deepfakes, generative adversarial networks, machine learning, synthetic media, balanced centrality

## **1. Introduction**

In 2017, there was a considerable debate regarding the digital resurrection of Carrie Fisher (the actor who portrays Princess Leia) in the Star Wars: The Rise of Skywalker movie, due to her passing before shooting (Lowry, 2017). To retain authenticity, Carrie’s daughter, Billie Lourd, acted as a stand-in actor during shooting and Carrie Fisher’s facial and voice features were digitally overlaid. Such digital manipulation of visual and audio is not just limited to Hollywood blockbusters, and is becoming increasingly facilitated by artificial intelligence (AI) technology. This phenomenon, known as the “deepfake”, is a new form of AI-enabled content creation which is being implemented by marketers and customers. This

conceptual paper sets out to understand how AI has led to “The Rise of Deepfakes” and discusses the relevance and need for future marketing scholarship in this area. This paper also synthesizes literature from marketing, information science, and communication relating to deepfakes and AI to develop a conceptual model and research agenda.

AI is a growing industry which is expected to reach \$118 billion USD by 2025 (Tractica, 2019), and AI-enabled robots are already amongst us, in sectors such as retail, foodservice, tourism, and health (Mende, Scott, van Doorn, Grewal, & Shanks, 2019). Yet it is not only the *speed* of development that is unprecedented; it is also *how* the development is occurring. Given machine learning technologies and collaboration opportunities over the internet, development can sometimes be shared, or even contributed to by the AI itself if autonomous machine learning functionalities have been programmed in by trained information scientists. AI is a complex technology with broad applications across multiple fields. For instance, AI can be used to enhance algorithms (for example, in AI-enabled biometrics for recruiting; van Esch, Black, Franklin, & Harder, 2020), to provide the ‘mind’ inside embodied robots and virtual assistants, and to detect patterns and provide advice in daily life – such as in share market trading. Synthetic media is also an application of AI and describes automatically and artificially generated media that blur the lines between real and fake (Whittaker, Kietzmann, Kietzmann, & Dabirian, 2020). Such applications are not necessarily malicious. Indeed, customers already interact with AI such as chatbots and virtual assistants. However, what remains to be seen is whether customers are accepting of deepfake technologies, which may obscure the fact that seemingly ‘real’ content was created via AI.

There have been some initial discussions into the emergent AI-based deepfake technology. Deepfakes are the product of AI and the ‘deep learning’ machine learning technique which trains deep neural networks. These networks can be thought of as a brain containing many interconnected neurons, with these artificial neurons being referred to as

‘units’. Though each of these units complete a simple computation, the multitude of units within a deep neural network together can perform complex tasks (Kietzmann, McClure, & Kriegeskorte, 2019). These deep neural networks can be trained to artificially generate and manipulate video, images, and audio to create hyper-realistic fake content (Kietzmann, Lee, McCarthy, & Kietzmann, 2020). This example demonstrates how deepfakes can be viewed as a type of synthetic media as the realistic deepfake output is automatically and artificially generated (Whittaker et al., 2020). With deepfakes, it is difficult to tell if the person in the video is real or fake, or alternatively for customers to realize this is a question they should be asking. There are numerous marketing opportunities for the application of deepfakes bringing (a new) hope that deepfakes can be used for the benefit of both customers and business – offering stronger engagement across cultural boundaries, enhanced messaging personalization and effectiveness, and even opportunities for customer-led co-creation. However, the potential of deepfakes as a ‘phantom menace’ should also not be ignored. Indeed, consent and awareness are two central issues with deepfakes as the potential for deviance is evident – consider the hijacking of political or marketing campaigns, criminal activities like phishing or revenge pornography, or even the potential for psychological harm stemming from the fluidity deepfakes seemingly grant to mortality. Despite deepfakes possessing a light and dark side, marketing scholarship has not yet thoroughly theorized or discussed these perspectives. Thus, this paper seeks to introduce deepfakes and theorize directions for future research to explore the light and dark side of deepfakes.

The purpose of this paper is therefore threefold: first, to define and introduce deepfakes to the marketing literature, second, to differentiate deepfakes from existing marketing practices to highlight their unique contribution to the discipline; and third, to provide a conceptual framework and research agenda to guide marketing scholars’ investigation of this new phenomenon. The introduction of deepfakes, alongside a typology

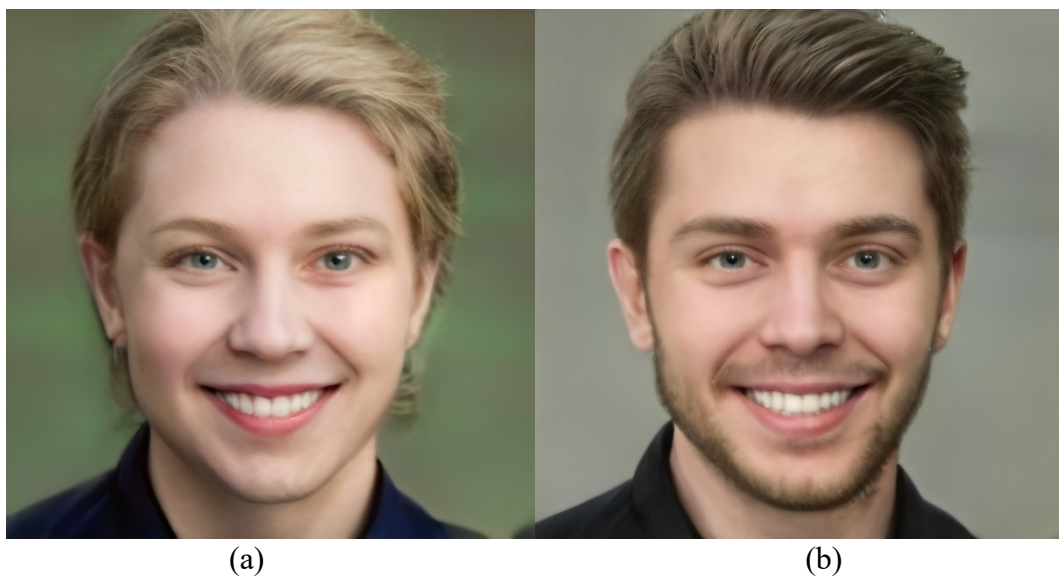
and conceptual framework to the marketing literature is important for two main reasons. First, the creation of authentic, persuasive content is a key aim of marketing – deepfakes offer a new way of providing such content, but as yet this technology is little understood outside of information science and has not yet been integrated with marketing theory and differentiated from existing marketing practice. Second, as with any technology that is not yet well understood, there is potential for unexpected – and perhaps immoral – uses. Hence, marketers need to know more about potential opportunities for deviance utilizing deepfakes, both on the business-side and the customer-side, as well as understanding how deepfakes may affect the overall marketing ecosystem. It is only through more research, shedding light on the application of deepfakes and their relation to marketing theory and practice, that we can hope to guide safe, sustainable practices for deepfakes in the marketing field.

This paper applies a Star Wars metaphor throughout to discuss both the light and the dark side of deepfakes and how we may yet find balance by embracing the ‘force’ of deepfakes for businesses and customers. The remainder of this paper is organized as follows. First, we define and introduce deepfakes, providing a typology that differentiates deepfakes from other marketing practices. Following this, we provide an overview of the benefits of AI – with a particular focus on the deepfake applications of AI – for businesses and customers through the theoretical lens of balanced centrality which sets the foundation for the proposed conceptual framework. Finally, we present our conceptual framework and its propositions which provide a research agenda for deepfakes.

## **2. What are Deepfakes? A New Hope or Phantom Menace for Marketing?**

A deepfake can be defined as a product of AI technology and machine learning which merges, combines, replaces or superimposes various types of content to produce a form of synthetic media that obscures the distinction of authenticity (Maras & Alexandrou, 2019). Deepfake media first rose to public awareness in 2017 and can take various forms within

digital media - whether audio, visual, or audio-visual – as summarized by Kietzmann and colleagues (2020). For example, photographic deepfakes enable faces and bodies to be swapped or blended into someone else’s seamlessly. As demonstrated in Figure 1, images of three individuals can be integrated in different ways to create deepfake content. These two images were generated using *Artbreeder* ([www.artbreeder.com](http://www.artbreeder.com)), an online tool which uses deepfake-based technology to allow users to merge existing images and generate new ones. The original three images (of two male co-authors and one female co-author) were uploaded and merged together as shown in Figure 1, with characteristics of the female co-author being more pronounced in the left image, and the male characteristics of the two co-authors being more pronounced in the right image.



**Figure 1.** Deepfakes generated by merging images of the authors – (a) female dominant version, (b) male dominant version.

Deepfakes are not restricted to visual content and can also permit voice-swapping, with audio being able to be synthesized via either changing or imitating the voice of someone else (Kietzmann et al., 2020). Text can also be converted into the voice of a chosen individual (Saito, Takamichi, & Saruwatari, 2018). Beyond photographic and audio deepfakes, deepfakes can be used within video, where faces can be swapped or morphed (Kietzmann et al., 2020). Beyond faces, deepfakes can even permit full-body puppetry, with the movement

of an individual within a video being manipulated via transposition from another person's movement (Chan, Ginosar, Zhou, & Efros, 2019). Perhaps the most sophisticated evolution of deepfakes, audio-visual deepfakes allow for the manipulation of both the facial movements and spoken words of an individual within video, convincingly making someone appear to say things which they have never said (Whittaker et al., 2020).

Deepfakes are rapidly evolving in sophistication and will eventually be undetectable to the untrained eye (Maras & Alexandrou, 2019). Indeed, the two main factors driving their proliferation through social media includes their increasing accessibility and believability, as deepfakes are becoming easier to produce with customer-grade apps such as *Zao* and *FakeApp*, and also harder to distinguish from authentic media due to their increasing sophistication (Kietzmann et al., 2020). In sum, deepfakes give both individuals and organizations the power to create highly realistic, yet synthetic representations of whoever they please.

While the deep neural networks which facilitate deepfakes can artificially generate and manipulate audio-visual content, it is also noteworthy to briefly discuss that these networks can generate entirely new, yet realistic content in the form of generative adversarial networks (GANs). GANs comprise of a generator network and discriminator network, whereby the generator network (acting the role of counterfeiter) generates content aiming to deceive the discriminator network (acting the role of a counterfeit detective) (Goodfellow et al., 2014). Over time, the generator network learns to improve its output to eventually deceive the discriminator network – which when facilitating deepfake creation can result in highly realistic synthesized content (Whittaker et al., 2020). That is, the AI that generates content becomes proficient enough at doing so that the AI that judges the authenticity of the content can no longer tell the difference. After this training process, the generator can create entirely new content which has high similarity to the original source input, such as a person's voice or

face (Whittaker et al., 2020). As GANs can generate novel output, they are slightly different to deepfakes. However, as GANs utilize AI-driven machine learning to create synthesized human content (like deepfakes), we refer to human-like content generated by GANs as deepfakes within our paper. Next, we conceptualize how deepfakes differ from other marketing practices which leverage human realism and/or technologically-facilitated intelligence.

### *2.1 A Typology of Human Realism and Technologically-facilitated Intelligence within Marketing Practice*

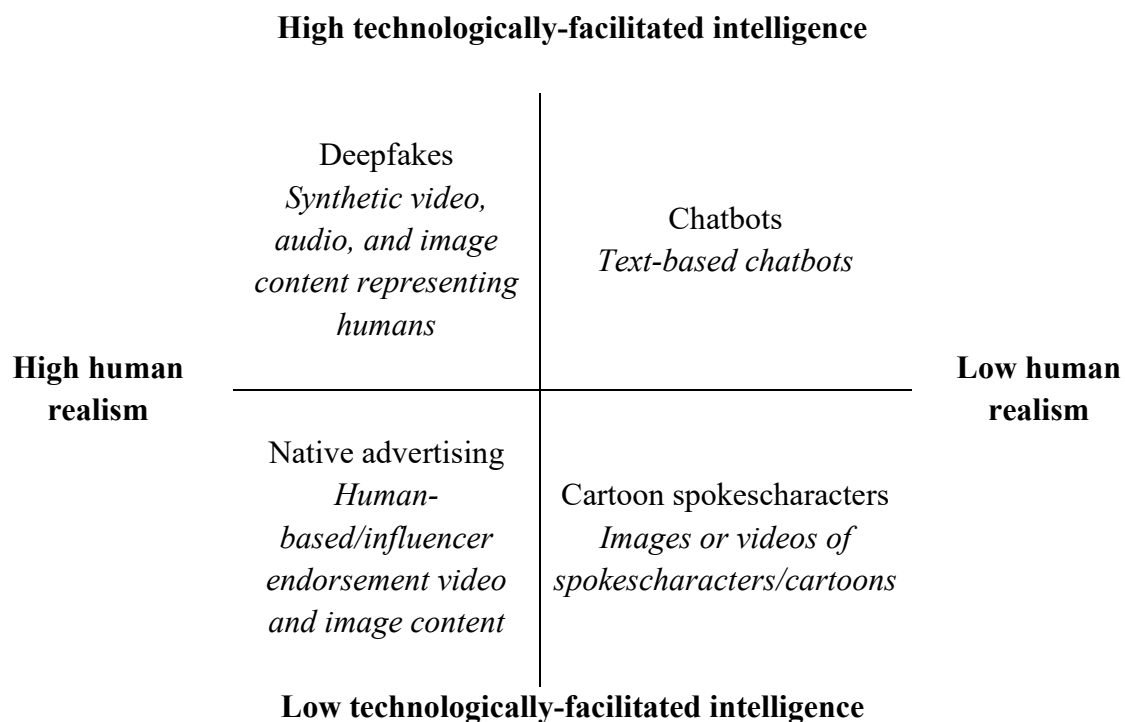
This section presents a typology that highlights the placement of deepfakes amongst other marketing practices, and hence, the unique contribution of deepfakes to the marketing discipline. Our typology is developed to provide a more precise and nuanced understanding of deepfake phenomena through differentiation from other marketing practices, hence highlighting a conceptual contribution, in accordance with MacInnis (2011). Differentiation creates scholarly contribution by distinguishing and classifying phenomena in accordance with underlying dimensions of comparison to articulate important differences to add clarity and reduce confusion between phenomena to create new insight (MacInnis, 2011). Our proposed typology differentiates between deepfakes and related marketing practices to provide new insight into the unique characteristics and advantages of deepfakes that can be leveraged by marketers.

Two dimensions of comparison constitute the typology's axes: human realism and technologically-facilitated intelligence. Firstly, human realism refers to the ability of digital content to be representative of humans in the real world, for example the use of humanistic physical features, faces, and voices. This dimension goes beyond anthropomorphism (the granting of human features or characteristics) and considers the *perception* of realism or



authenticity. Perceived realism is important because the sense of authenticity can lead to enhanced trust (Sundar, 2008).

The second axis deals with levels of technologically-facilitated intelligence, and describes the reliance upon, or utilization of, intelligent technologies to automate digital content creation. We conceptualize that high-intelligence content relies more heavily on AI, whereas low-intelligence content has reduced utilization, or absence of AI, relying more on human input to facilitate content creation. Such a conceptualization is made in accordance with the Passive Interactive Proactive (PIP) typology (Letheren, Russell-Bennett, Mulcahy, & McAndrew, 2019), which proposes that as humans place greater reliance on technology, the technology moves from adopting a passive role (requiring cognitive and behavioral human input to operate), to becoming interactive (providing personalized output and making some actions on behalf of the human), and finally proactive (autonomous decision-making and actions made of behalf of the human, which may be overridden). Please refer to Figure 2, which presents our typology.



**Figure 2:** Typology of human realism and technologically-facilitated intelligence.

### *2.1.1 Deepfakes: High Human Realism and High Technologically-facilitated Intelligence*

Deepfakes are firstly positioned within the high human realism, high technologically-facilitated level of intelligence quadrant. As deepfakes can be constructed to be highly believable, albeit inauthentic, audio-visual content which manipulates the facial, vocal, and bodily features of humans (Whittaker et al., 2020), we propose that deepfakes possess high human realism. Through their utilization of AI to recognize a human's facial features and expression patterns, and via deep learning to optimize output by training on external data fed into the system (Kietzmann et al., 2020), deepfakes are conceptualized to possess a high level of cognitive-facilitated intelligence rather than other advanced forms of intelligence such as emotional or social intelligence (Kaplan & Haenlein, 2019). Deepfakes are therefore more proactive in accordance with the Proactive-Interactive-Passive (PIP) framework of Letheren et al. (2019), as while deepfakes require human input to create, they proactively use AI to act on the human's behalf to generate synthetic representations created from inputted data.

### *2.1.2 Positioning Existing Marketing Practices: Human Realism and Technologically-facilitated Intelligence*

We next compare deepfakes to existing marketing practices which fall within other quadrants of our proposed typology but differ based upon their levels of realism and use of technologically-facilitated intelligence. Chatbots attempt to emulate human characteristics, for example by having humanized voices and visual animations, however such synthetic representations of human characteristics by chatbots have been shown to produce negative evaluations and the 'uncanny valley' effect (Ciechanowski, Przegalinska, Magnuski, & Gloor, 2019), therefore they may possess low perceived human realism. However, chatbots possess a high level of technology-facilitated intelligence as they can utilize AI and machine learning to automate customer service and communications by employing language-processing tools to understand customer requests and respond appropriately (Luo, Tong,

Fang, & Qu, 2019). Chatbots are therefore more interactive in accordance with the PIP framework (Letheren et al., 2019), as they provide information and suggestions.

Native advertising possesses high human realism as it emulates authentic and realistic content which makes it difficult for customers to recognize it is an advertisement (Campbell & Evans, 2018). As native advertising is often produced by social media influencers (Kay, Mulcahy, & Parkinson, 2020), relying on the human to generate content rather than utilizing AI, it is conceptualized to possess low technologically-facilitated intelligence. Native advertising is therefore passive in accordance within the PIP framework (Letheren et al., 2019), as the technologies used to create and disseminate the endorsed content (e.g. cameras, photo and video editing software, social media platforms) are static. Therefore, there is greater reliance placed upon the human to generate the output.

Lastly, the animated nature of cartoon spokescharacters means that they possess low human realism. Indeed, realism is generally not the aim of cartoon spokes-characters, which are instead designed to act as symbols, and even evoke nostalgia for childhood (Hosany, Prayag, Martin & Lee, 2013). Spokescharacters also possess a low level of technologically-facilitated intelligence as they lack self-agency being entirely designed, animated, scripted, and programmed by human visual designers.

The proposed typology does possess several caveats which will now be acknowledged. Although deepfakes are beginning to become more sophisticated and harder to distinguish from authentic content (Kietzmann et al., 2020), many deepfakes are constructed with lower human realism, (particularly for parody purposes), with the alteration being clearly visible. In addition, while deepfakes are conceptualized to possess high intelligence, they are not automated and do require human intervention (e.g. to collect optimal training data of the source and target and input it, in addition to converting and

exporting the video and conducting further visual editing if needed). In addition, recall that deepfakes do not currently possess more advanced forms of AI, such as the simulated emotional or social intelligence found within human-inspired or humanized AI (Kaplan & Hainlein, 2019). Furthermore, if other marketing practices were to increase levels of realism via technologically-facilitated intelligence such as chatbots and spokescharacters, we argue this would signify these practices being or emulating being ‘deepfaked’. In summary, the typology developed is not meant to be an exhaustive depiction of all deepfake, chatbot, native advertising and cartoon spokescharacter content, as shifting can certainly occur between the quadrants, and other examples may also be found which fit within these quadrants. The typology instead aims to differentiate deepfakes from commonly deployed marketing practices which use realism and technologically-facilitated intelligence to better understand their unique contribution to the marketing discipline. Now that deepfakes have been introduced, conceptualized and differentiated in comparison to other marketing practices, the following sections outline their current and potential positive and negative impacts.

## *2.2. Deepfakes – A New Hope for Marketing?*

From the perspective of ‘a new hope’, deepfakes can be utilized to create content which can positively persuade or connect with customers. For example, deepfakes have been used to reach voters speaking a different dialect such as in the video released by Indian president Manoj Tiwari which depicted him speaking another dialect and reached approximately 15 million people (Jee, 2019). In another instance, social marketing applications of deepfakes are emerging, such as in the Malaria Must Die campaign. The campaign uses symbolism and appeals that transcend cultural barriers, depicting former English footballer David Beckham seemingly speaking nine different languages to promote a petition to end malaria (Malaria Must Die, 2019).

Customers can also create deepfakes, for instance, using personalized media creation that uses face-swapping technology or even creating their own AI avatars. For example, users can face-swap with actors in scenes from movies and TV shows by using *Zao*, a popular Chinese app, and post their creations onto social media. Deepfakes therefore offer new opportunities for businesses and organizations to personalize online experiences and enhance customer engagement. Such an opportunity was realized when Doritos partnered with the *Sway* app to create the first AI dance challenge to engage with customers. Users filmed themselves striking a collection of dance poses, which the app transposed via deepfake over the body of a professional dancer emulating the dance moves which Lil Nas X and Sam Elliot performed in the Doritos Super Bowl commercial. Users were then encouraged to share their professional dance moves using #CoolRanchDance via social media (Williams, 2020). Deepfakes can also augment the personalization of online customer experiences such as clothes shopping. With the recent development of AI-generated promotional models by companies such as *Artificial Talent*, online clothes retailers could eventually allow customers to generate their own personalized avatar and deepfake their facial characteristics onto it to visualize prospective clothing purchases.

### 2.3. *Deepfakes – a Phantom Menace for Marketing?*

Alternatively, deepfakes can be argued to be a ‘phantom menace’, negatively influencing both customers and businesses. Given deepfakes possess low barriers to creation, have great potential to persuade the public and can be easily shared in this new digital era (Kietzmann et al., 2020), these forged entities can be very dangerous indeed. Disconcertingly, the growing simplicity involved in creating convincing deepfakes, combined with our increasingly digitally-documented lives, will heighten the potential for deepfakes to be used for malicious purposes such as disinformation or ‘fake news’, blackmail, intimidation,

sabotage, harassment, defamation, revenge pornography, identity theft, and bullying (Chesney & Citron, 2019; Kietzmann et al., 2020).

Deepfakes present profoundly problematic implications for social media and political marketers. Fake news proliferates widely on social media and studies have assessed how text-based disinformation messages are accepted within a social media context (Pennycook & Rand, 2019). However, as visual stimuli can evoke emotional and cognitive arousal more effectively than text (Sherwin, Feigenson, & Spiesel, 2006), deepfakes are a more potent and persuasive tool for disinformation purposes which may even result in the public losing trust in factual information (Chesney & Citron, 2019). Indeed, Vaccari and Chadwick (2020) found that political deepfakes do not need to deceive to do damage – instead, their presence creates confusion, negatively influencing norms and civic culture in a way that could potentially play a significant role in undermining democracy. In further evidence of the potential for deepfakes to menace businesses and customers, Facebook recently removed over 900 counterfeit accounts which circulated pro-Trump messages as they utilized fake profile pictures generated using deepfake technology (Nuñez, 2019). Given the emergence of these practices, countries such as China and the United States are exploring legislative measures to mitigate negative impacts of deepfakes. China for instance has declared it a criminal offense to publish a deepfake without providing disclosure of its artificial nature (Woollacott, 2019). Comparatively, the United States Senate recently passed the Deepfake Report Act. This Act requires the Department of Homeland Security to publish an annual report on ‘digital content forgery’ – digital content crafted using AI which is designed to mislead (Deepfake Report Act, 2019).

In summary, these discussions highlight the potential for deepfakes to have a positive *and* negative impact on customers and businesses. Understanding this juxtaposition, we next

argue how balance within the deepfake ‘force’ is needed by applying the theoretical lens of balanced centrality.

### **3. The Need for Balance in the Deepfake Force: A Conceptual Model**

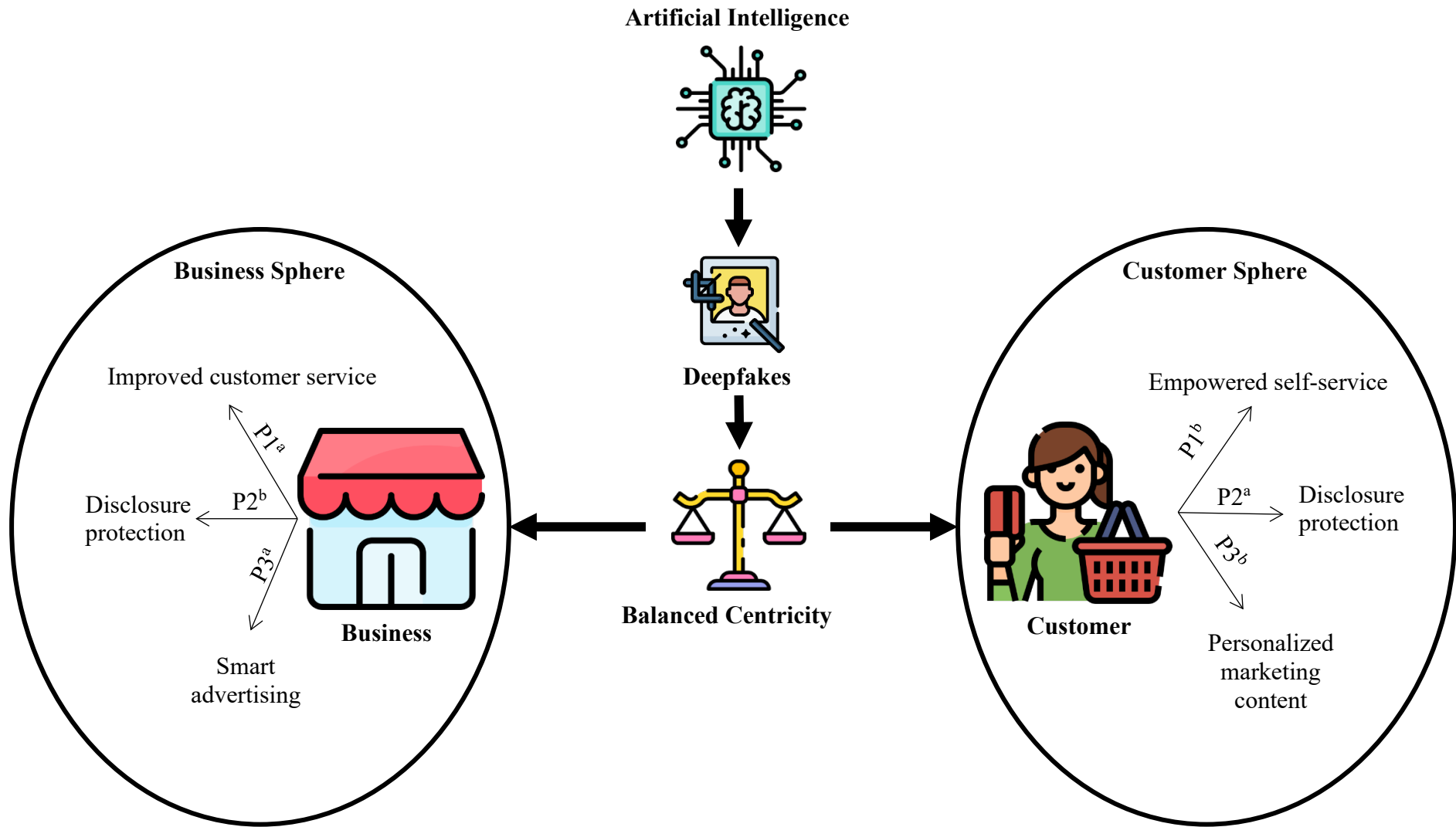
In this paper, we argue for the need to balance the benefits between business and customers to ensure all can benefit from the deepfake application of AI. Arguments regarding the need for balance in marketing theorizing is not new, but it is yet to be extended into the application of AI and deepfakes. This debate regarding the need for balance is evidenced in discussions regarding the limitations of perspectives such as customer centrality, whereby the focus of marketing on benefitting the customer (e.g. ‘customer is king’ or ‘the customer is always right’; Shah, Rust, Parasuraman, Staelin, & Day, 2006). Instead, others such as Gummesson (2008) urge marketing scholars to accept the complexity of marketing and embrace and theorize towards a balanced centrality, whereby all stakeholders benefit, and thus in turn create a more sustainable and harmonious economy. It is in line with the theoretical perspective of balanced centrality that we suggest deepfakes can assist by balancing the benefits for business and customers, which is what we demonstrate in the propositions arising from our conceptual model.

Currently, in opposition to the central tenets of balanced centrality, it could be argued that emergent empirical studies into deepfakes are predominantly customer-centric, with a distinct focus on understanding how customers might interact with deepfake media in an effort to begin to measure their negative impacts (Dobber et al., 2020; Vaccari & Chadwick, 2020). In comparison, there is only limited and emerging commentary discussion of the business implications of deepfakes (see Kietzmann et al., 2020; Kwok & Koh, 2020; Whittaker et al., 2020), with some reports suggesting they may result in deep losses for businesses (Helms & Mutter, 2020). We argue that theorizing and studying deepfakes through a balanced centrality lens will help businesses (and customers) in understanding how

all stakeholders, rather than one in isolation, can benefit or be protected from the negative impacts of deepfakes. Balanced centrality therefore provides research motivation and direction to ensure that the positive and negative impacts of deepfakes are understood within both the customer and business sphere.

In line with this prior argument, the current paper proposes three main propositions, which are broken into the business and customer sphere based upon the notion of balanced centrality, as shown in Figure 3. The top of our model depicts our earlier discussion in Section 2 which explained that AI powers deepfakes, and hence this is the first connection drawn within our model. Next is the association between deepfakes and balanced centrality. As argued in Section 3, there is a need for balanced centrality to ensure that both customers and businesses can benefit, but also just as importantly, be protected from deepfakes. The relationship drawn between deepfakes and balanced centrality provide the grounding for our propositions and their variations which are broken up according to the business and customer sphere. Next, the first proposition relating to customer service and self-service is discussed.





**Figure 3.** Conceptual model of deepfakes and their implications for businesses and customers.

### *3.1. Proposition 1 – Customer and Self-service via Deepfakes*

#### *3.1.1 Improved Customer Service in the Business Sphere*

The broader AI literature provides support for deepfakes enhancing customer service. Scholars argue that AI will begin – and continue to – replace repetitive service tasks and undertake analytical tasks, whereas the proficiency of AI in empathetic tasks is limited (Huang & Rust, 2018). For example, AI-generated advertisements are not as effective at conveying emotional or hedonic appeals when compared to human-created advertisements, and people may form negative evaluations of AI because of their inability to experience or feel (Bakpayev et al., in press).

As shown in recent studies of AI, customers have a negative perception of the use of ‘machines’ (Luo et al., 2019). Individuals are also known to possess fundamental biases against algorithms which could impede the acceptance of AI within customer service contexts, a phenomenon referred to as algorithm aversion (Dietvorst, Simmons, & Massey, 2015). Individuals prefer to rely on humans rather than algorithms even when this results in an inferior overall outcome, as there is a greater intolerance of mistakes made by algorithms when compared to humans, with a greater likelihood of abandoning an algorithmically based decision-maker rather than a human alternative (Dietvorst et al., 2015). In addition, individuals place less trust in algorithms when they are used to complete subjective tasks, which require interpretation and intuition and emotional perception, as algorithm aversion is primarily driven by the perception that algorithms lack such human abilities (Castelo, Bos, & Lehmann, 2019).

Whilst studies such as Huang and Rust (2018) suggest AI cannot be emotive (e.g. express empathy), we suggest AI via deepfakes may bridge this gap. This aligns with suggestions that AI can help interactions between machine and human, as well as human-to-human interactions (Wilson & Daugherty, 2018). When algorithms are perceived to have

high affective human-likeness (e.g. creativity and emotional sensitivity) algorithm aversion has been shown to have a reduced effect within more emotionally-orientated and intuitive tasks (Castelo et al., 2019).

Deepfakes can potentially be used to enhance the naturalness of artificial agents and improve their ability to build simulated empathy and emotional connection. Deep neural networks, which facilitate the creation of deepfakes, can be trained to recognize and learn to represent emotions based on inputted facial data (Lai & Lai, 2018). Deepfakes are already being used to impart emotional experiences within customer contexts. The Dalí Museum in Florida adopted deepfake technology to bring tourists face-to-face with Dalí, who appears as a life-size incarnation and speaks with the visitor to better forge an emotional connection between the visitor and his artworks (The Dali Museum, 2019). Deepfake technologies therefore grant artificial agents the power to not only establish a form of emotional connection but can imbue them with a simulation of emotional intelligence, enabling them to adapt and better satisfy human desire for a more natural emotional experience.

Imagery can help to capture attention, increase memorability, influence attitudinal changes, and evoke cognitive and emotional arousal to become more persuasive (Seo & Dillard, 2019; Sherwin et al., 2006). Facial expressions serve a vital role in coordinating social interactions, serving as a visual tool which helps the message recipient understand the communicator's emotions, beliefs and intentions (Keltner & Haidt, 1999). Positive emotions such as enthusiasm, amusement, and attachment love facilitate greater acceptance of weak persuasive messages (Griskevicius, Shiota, & Neufeld, 2010), therefore the addition of emotional warmth to messaging may help to improve its persuasiveness. Using deepfakes to visualize positive emotions such as facial expressions could therefore improve the persuasiveness of communication over text-based alternatives and help to deliver the message as it was intended.

We therefore propose that deepfakes could be used to emotionally augment AI-powered service agents (i.e. chatbots and virtual assistants), allowing them to possess greater ability to simulate emotions, thereby enhancing perceived emotional intelligence and adapting in real-time to human needs for empathy and understanding.

**P1<sup>a</sup>:** Deepfakes integrated into AI powered services (e.g. chatbots, virtual assistances) will increase perceptions of their emotional intelligence

### *3.1.2 Empowered Self-Service Customer Sphere*

Another benefit which may accrue from deepfakes and AI for customers is greater empowerment of self-service within the marketplace. As André et al. (2018) point out, AI will be able empower customers through lowering search, transaction, and decision-making costs. Further, AI may also benefit vulnerable customers – individuals who experience a state of powerlessness that hinders consumption goals and circumstances which negatively affect perceptions of self (Baker, Gentry, & Rittenburg, 2005). Therefore, AI may help to alleviate such powerlessness and allow vulnerable customers to take advantage of market opportunities and be supported to deal with complex marketplaces and exchanges where they may have limited experience or self-efficacy (Libai et al., in press).

It could therefore be posited that applications of deepfakes via AI could empower vulnerable customers to use self-service technologies (SSTs) by providing them with increased levels of self-efficacy, confidence and trust. Indeed, the concept of customer empowerment has been extensively investigated in reference to self-service technology (SSTs). This is demonstrable through technological designs which shows increased levels of empowerment through self-efficacy, confidence, and trust lead to greater use of digital services. For example, customers of m-health services are known to use them more if they have a greater sense of empowerment through heightened self-efficacy (Schuster, Proudfoot, & Drennan, 2015). Other studies also show that if customers feel empowerment in digital and

self-service settings through increased levels self-confidence, or the sense of power or autonomy, they have an increased likelihood toward adopting mobile banking (Chaouali, Souiden, & Ladhari, 2017). Perceived empowerment has also been shown to lead to increased trust within virtual environments (Füller, Mühlbacher, Matzler, & Jawecki, 2009).

We therefore extend these findings from prior technologies to suggest that deepfakes may be able to replicate and improve such outcomes. For instance, individuals diagnosed with amyotrophic lateral sclerosis (ALS) may rely on a text-to-speech interface to facilitate verbal communication as the condition progressively inhibits voluntary muscle control and can lead to total paralysis. Using generative AI to learn from voice samples provided by those diagnosed with ALS, *Project Revoice* ([www.projectrevoice.org](http://www.projectrevoice.org)) creates personalized synthetic voices for vocally-paralyzed clients which allows them to communicate with others using their own voice. Similarly, deepfakes can be used to empower the marginalized to allow them to share their stories using authentic emotion without fear of identification.

*Welcome to Chechnya*, a documentary which explores the persecution of LGBTQ people in Chechnya, uses deepfakes to anonymize the subjects of the documentary by transplanting the faces of New York-based activists onto the faces of the Chechen documentary subjects. The use of deepfakes to anonymize, instead of utilizing conventional anonymization tools within documentaries such as pixilation, blurring, darkened rooms, and altered voices, allows the documentary subjects to convey richer emotion (Richards, 2020). Deepfakes can therefore be used to provide the disempowered with the confidence to communicate without fear of identification while also preserving the emotional richness of their messages. Based upon the prior discussion of prior studies and the practical application examples of deepfakes, we propose the following:

**P1<sup>b</sup>:** Deepfakes integrated within self-service technologies will increase a customer's sense of empowerment (self-efficacy, confidence and trust) and in turn their usage.

### *3.2. Proposition 2 – Protecting Customers and Businesses from Deepfakes via Disclosure*

The next proposition of our conceptual model suggests the need to protect the customer and business sphere from the deviant use of deepfakes. Recall previously in Section 2.2 which discussed that deepfakes can potentially be a ‘phantom menace’ for marketers (business) and customers. Further, to combat this threat, policy makers are currently discussing regulatory measures such as disclosure of deepfake content creation to address this issue. Disclosure has been of interest to marketers, particularly regarding sponsorship (Campbell & Evans, 2018) and social media influencer marketing (Kay et al., 2020). Recent developments such as Facebook banning AI-manipulated videos (deepfakes) as a result of their potential influence within the 2020 US Presidential Election (Hern, 2020) have drawn the attention of regulators and policy makers regarding the need for deepfake disclosure. Whilst drawing this attention, these discussions of the integration of disclosure regulations for deepfakes have yet to be empirically supported or understood. We therefore turn to the broader disclosure literature to propose future research directions regarding the impact of disclosure on deepfakes, which can contribute important theoretical and policy insights.

#### *3.2.1 Disclosure of Deepfakes to Protect the Customer Sphere*

In the literature, studies have evidenced that disclosure of advertising has assisted customers in overcoming the persuasiveness of a message (Boerman, Willimsen, & Van Der Aa, 2017; Campbell, Mohr, & Verlegh, 2013; De Jans, Cauberghe, & Hudders, 2019). For example, Campbell and colleagues’ (2013) study shows that disclosure helps customers overcome covert marketing efforts and correct persuasive knowledge. This most likely aligns with the perspective of policy makers and regulators that introducing disclosure measures will help protect the customer sphere. Thus, given the evidence of prior disclosure research, we propose the following, relating to deepfakes and the customer sphere:

**P2<sup>a</sup>:** Disclosure in deepfakes will protect customers from the persuasion of deepfakes.

### *3.2.2 Disclosure of Deepfakes to Protect the Business Sphere*

Research also hints that the business sphere can benefit through disclosure. Kay and colleagues' (2020) study evidenced that when social media influencers disclosed sponsorship, customers perceived products as more attractive and were more likely to purchase. Thus, in line with these findings, it could be suggested that a balance could be struck for the business and customer sphere by the introduction of disclosure measures for deepfakes. However, disclosure is not only important from business-controlled messaging, but also protection from the threat they may pose to their brands and products via fake news and fake posts using deepfakes. Fake news can target brands, such as a malicious news article which intentionally misquoted Pepsi's CEO Indra Nooyi to say that Donald Trump's supporters should not purchase Pepsi, which led to calls for boycotts and potentially influenced the 3.75% drop in Pepsi's share price (Liffreing, 2016). In 2019, a company was defrauded of \$243,000USD when AI-enabled voice software was used to convince the CFO that the CEO was requesting the money be transferred (Stupp, 2019). Deepfakes can therefore be constructed to add artificial realism towards malicious agendas, meaning that businesses may need to begin to formulate strategies to protect themselves from potential reputational damage, for example, in the event of a malicious deepfake video targeting a senior executive going viral on social media. Therefore, based upon the prior discussion, and in line with the findings of benefits of disclosure for business outcomes from Kay and colleagues (2020), we propose:

**P2<sup>b</sup>:** Disclosure in deepfakes will enhance business outcomes and protect them from the deviant/destructive use of deepfakes.

### 3.3 *Proposition 3 – Smart and Personalized Advertising via Deepfakes*

We propose that deepfakes will provide opportunities for the business sphere regarding smart advertising, advertising which takes advantage of integrated customer tracking solutions to create commercials and advertisements which are adapted to the preferences of the public or individuals (Culic, Radovici, & Rasu, 2020). Deepfakes can already be used to develop personalized videos to send to customers throughout their customer journey. Personalized deepfake videos using an AI presenter can be created via *Synthesia* ([www.synthesia.io](http://www.synthesia.io)), where organizations can submit text to be converted into personalized videos such as appointment reminders, welcome messages, and abandoned shopping cart reminders, using the customer's name and native language. Personalized internet content has been shown to create positive evaluations (Kalyanaraman & Sundar, 2006), reduce information overload, and increase user satisfaction (Liang, Lai, & Ku, 2008). In addition, anthropomorphized content has been shown to have a positive influence upon customer behavior, where products possessing human-like features are evaluated more positively (Aggarwal & McGill, 2007), and the presence of anthropomorphism within branding can have positive influence on marketing outcomes such as brand love (Rauschnabel & Ahuvia, 2014). As deepfakes are already being implemented to personalize marketing communications and can be used anthropomorphize content via automatic generation of human characteristics such as faces and voices, deepfakes will potentially enhance the authenticity and effectiveness of advertising messages (Kietzmann et al., 2020).

#### *3.3.1 Smart Advertising and Creating the Ideal Message Source for the Business Sphere*

The first benefit to business, smart advertising, refers to AI-based advertising which allows “consumer-centered, data-driven and algorithm-mediated brand communication” (Li, 2019, p. 333). Research suggests AI is assisting in the curation and refinement of content (Kietzmann, Paschen, & Treen, 2018), targeted and personalized advertising in real time, and



ROI on advertisement expenditure (Davenport, Guha, Grewal, & Bressgott, 2020). Indeed, scholars such as Davenport and colleagues (2020, p. 35) pose the question which encapsulates the benefits of smart advertising, “would [these] advertising dollars be required in the future, wherein firms may be better predicting customers’ preferences, and thus would not need to advertise as much?”.

Creating an ideal message source is central to the development of source credibility. An individual’s tendency to accept information depends on the perceived credibility of the communication source (Hovland, Janis, & Kelley, 1953). One such practice employed by marketers to adopt an ideal message source is the utilization of celebrity endorsement for product or brand promotion. Celebrities are considered as effective endorsers as they are viewed to be trustworthy, believable, persuasive, and likeable (Freiden, 1984). The value of celebrity endorsement is realized through the creation of an associative link between the brand and celebrity, however this also creates inherent risk (Till, 1998). As endorser credibility can directly influence brand credibility and indirectly influence customer-based brand equity (Spry, Pappu, & Cornwell, 2011), negative publicity acquired by the celebrity can be associated by customers with the endorsed brand (Till, 1998). To alleviate the risks associated with celebrity endorsement, marketers can instead opt to utilize spokescharacters for promotion. Through visual design or animation, marketers can control spokescharacters and how they promote products or brands to instill trust towards the character and in turn brand attitude (Garretson & Niedrich, 2004). Given that cartoon spokescharacters act not as themselves, but are specifically animated to serve as a symbol and embodiment of brand characteristics (Stafford, Stafford & Day, 2002), they are inherently less realistic and authentic than a non-cartoon endorser who exists outside of their brand role. While less human realism may be preferable in many instances, in some cases the level of authenticity offered by a human endorser may be desirable. We propose that marketers can use deepfakes

to create a synthetic endorser, striking the balance between a human endorser and controllable spokescharacter, to create an ideal message source.

Lil Miquela is one example of a synthetic endorser, who is constructed using computer-generated imagery (CGI). Boasting 2.7 million followers on Instagram, Lil Miquela engages in endorsement of brands such as Samsung, Calvin Klein, Dior, and Prada (Powers, 2019). Synthetic endorsers such as Lil Miquela have distinct advantages for brands, as they stick to the script, can be aesthetically perfected to suit the brand's objective, have their personal narrative and life drama designed by the brand, all without the uncontrollable risks or issues which a human endorser can introduce (Koh & Wells, 2018; Powers, 2019). Synthetic endorsers therefore retain the controllable aspects of conventional spokescharacters, but with an enhanced ability to embody human qualities similar to those of the target audience, and hence a greater sense of authenticity and relatability. For instance, a personal narrative or life story, beliefs, personality, and physical attractiveness can all be programmed into a synthetic endorser to create an ideal message source to fit the brand's audience. In turn, synthetic endorsers can feel real, as an individual's love or liking of a fictional character can be attributed to greater perceived realness (Gardner & Knowles, 2008). Beyond CGI, the use of deepfakes to produce an ideal message source is beginning to be realized by companies such as *Artificial Talent* (<https://artificialtalent.co>), who use deepfake technology to create AI-generated fashion models – giving brands the ability to generate synthetic, region-specific promotional models who wear the clothing of the fashion brand using customizable poses, all without incurring the costs associated with employing human models, such as agency fees, photographers, and beautification services.

As noted in the previous sections, deepfakes offer unique opportunities to change the appearance of individuals in marketing messages and further to reduce the 'noise' of advertising by appropriately targeting messages. Not only will deepfakes leverage the

benefits of AI for targeting customers, it will also be able to refine messages and message sources (e.g. promotional models and spokescharacters) to the preference of the customer. Thus, given the previously noted advantages of AI and those combined with the advantages of deepfakes, it is plausible to suggest that deepfake-powered advertising could be superior to purely AI-based advertising messaging.

**P3<sup>a</sup>:** Deepfake-enhanced advertising messages will be more effective than purely AI-based advertising messages.

### *3.3.2 Improved Advertising Authenticity and Personalization for the Customer Sphere*

The broader AI literature outlines numerous advantages which technological advances will provide customers, particularly regarding their customer experience, and we transfer this thinking to their experience of advertising. AI-powered chatbots can be viewed as personal assistants as they assist customers with requests and automate customer services (Luo et al., 2019), which importantly demonstrates that AI can work for the customer and not just the company, enhancing authenticity and believability. The integration of deepfakes can possibly enhance the perception of authenticity and personalization of advertising messaging, all of which are argued to be important for enhancing messaging effectiveness (Beverland, Lindgreen, & Vink, 2008). For instance, artificial tools such as Google Lens and Amazon Style Snap identify fashion items from videos or photos viewed by customers, providing information about how they can be worn and where they can be purchased (Lui, 2019). Deepfakes may also be increasingly deployed within the tourism industry to enable more vivid visualization. For example, co-creation opportunities between companies and potential tourists could occur if social media users were given super-resolution images of the tourist destination and the ability to deepfake themselves into the destination image – allowing them to better visualize a potential tourist experience (Kwok & Koh, 2020). Using the examples of fashion and tourism, we propose that deepfakes can create more personalized advertising by

allowing for more realistic visualization of the customer within the consumption experience and improving their ability to imagine how they will use products, which has been identified as an important assistance for customers understanding their purchases (Mulcahy, Letheren, McAndrew, Glavas, & Russell-Bennett, 2019). We therefore propose based upon the prior discussion:

**P3<sup>b</sup>:** Deepfake-enhanced advertisements will improve perceptions of advertising authenticity and personalization, as well as enhance a customer's ability to imagine use of products.

#### **4. Agenda for the Future**

By design, our review of extant work in deepfakes has taken a predominately marketing focus to ensure the conceptual model and propositions are actionable, but also to uncover the gaps in the marketing literature which provide new and exciting topics for exploration. From this perspective, we provide the below specific research questions using the aforementioned areas of smart advertising and personalization, customer and self-service, and disclosure to provide a research agenda for marketing scholars' investigation of deepfakes. Please see Table 1.

**Table 1.**

## Overview of marketing practice, propositions and research questions for deepfakes.

Area of Marketing Practice	Proposition(s)	Research Questions
Customer and Self-Service (P1)	<p>P1<sup>a</sup>: Deepfakes integrated into AI powered services (e.g. chatbots, virtual assistances) will increase perceptions of their emotional intelligence.</p> <p>P1<sup>b</sup>: Deepfakes integrated within self-service technologies will increase a customer's sense of empowerment (self-efficacy, confidence and trust) and in turn their usage.</p>	<p>Q1<sup>a</sup>: Can deepfakes of digital service employees enhance customer perceptions of a service organization's customer service?</p> <p>Q1<sup>b</sup>: Do customers perceive deepfake-enhanced digital assistants as superior to other digital assistants (e.g. chatbots?).</p> <p>Q1<sup>c</sup>: How can deepfakes be used to empower vulnerable customers in consumption settings?</p>
Disclosure Protection for Customers and Businesses (P2)	<p>P2<sup>a</sup>: Disclosure in deepfakes will protect customers from the persuasion of deepfakes.</p> <p>P2<sup>b</sup>: Disclosure in deepfakes will enhance business outcomes and protect them from the deviant/destructive use of deepfakes.</p>	<p>Q2<sup>a</sup>: What impact does disclosure of deepfakes have upon customer reactions and business outcomes?</p> <p>Q2<sup>b</sup>: Do certain types of disclosure of deepfakes work more effectively than others?</p> <p>Q2<sup>c</sup>: Does the timing and placement of disclosure for deepfakes impact their effectiveness?</p>
Smart and Personalized Advertising (P3)	<p>P3<sup>a</sup>: Deepfake-enhanced advertising messages will be more effective than purely AI-based advertising messages.</p> <p>P3<sup>b</sup>: Deepfake-enhanced advertisements will improve perceptions of advertising authenticity and personalization, as well as enhance a customer's ability to imagine use of products.</p>	<p>Q3<sup>a</sup>: To what extent can business advertising benefit from deepfakes?</p> <p>Q3<sup>b</sup>: What are the opportunities for personalized deepfake advertising, and how do customers respond to such content?</p>

### 4.1. *Customer and Self-service*

A significant body of research has, and is, continuing to be conducted in the use of AI technology for customer service and self-service. These studies to date have often been focused on machine (robot and chatbot) or human (self-service) input. Deepfakes provide an opportunity to begin to blur these foci, which provides opportunities for marketing scholars to gain insights into the issues and opportunities of blending human and machine. For example,

recall that in our proposed typology we positioned chatbots within the low human realism, high technologically-facilitated intelligence quadrant – it is possible that with the inclusion of deepfake enhancement that human realism and technologically-facilitated intelligence may be improved. To guide the identification of these insights we propose the following research questions:

**Q1<sup>a</sup>:** Can deepfakes of digital service employees enhance customer perceptions of a service organization’s customer service?

**Q1<sup>b</sup>:** Do customers perceive deepfake-enhanced digital assistants as superior to other digital assistants (e.g. chatbots)?

**Q1<sup>c</sup>:** How can deepfakes be used to empower vulnerable customers in consumption settings?

#### 4.2. *Disclosure*

In evaluating deepfakes and their potential to be a ‘phantom menace’, understanding strategies such as the impact of disclosure on their level of persuasion is critical. These have important implications for the development of policy and regulations for deepfakes – specifically the impact which disclosure has on deepfakes attempting to persuade customers. Further, insights could be drawn as to when, how, and where disclosure of a deepfake must be made to mitigate their level of persuasion. Marketing scholars can therefore seek to address the following research questions relating to deepfakes and disclosure:

**Q2<sup>a</sup>:** What impact does disclosure of deepfakes have upon customer reactions and business outcomes?

**Q2<sup>b</sup>:** Do certain types of disclosure of deepfakes work more effectively than others?

**Q2<sup>c</sup>:** Does the timing and placement of disclosure for deepfakes impact their effectiveness?

#### 4.3. *Smart Advertising and Personalization*

It is widely acknowledged that AI will change the way marketers create content and advertise, and it is hopeful that this will benefit both organizations and customers. Finding the

balance whereby AI enhances marketing and customer outcomes will be challenging. Of importance is to understand how and if deepfakes can potentially be used achieve this balance. To direct marketing scholarship to investigate whether deepfakes may achieve this balance, the following research questions are proposed:

**Q3<sup>a</sup>:** To what extent can business advertising benefit from deepfakes?

**Q3<sup>b</sup>:** What are the opportunities for personalized deepfake advertising, and how do customers respond to such content?

## **5. Conclusion**

Our key aims in this paper were to define and introduce deepfakes to the marketing literature, differentiate deepfakes from existing marketing practices, and provide a conceptual model which sets an agenda for future research. We have defined deepfakes in a broad sense, providing a series of recent examples to demonstrate the contemporary importance of this topic to marketing scholarship. We proposed a typology which demonstrates that deepfakes provide a unique contribution to the marketing discipline due to their high human realism and technologically-facilitated intelligence. Lastly, we developed and proposed a conceptual framework that highlights the need for balanced centrality, whereby businesses and customers can benefit from deepfakes but also be protected from them. The unresolved questions we provide have been organized in accordance to three identified areas of importance to “The Rise of Deepfakes” and set an agenda to direct future empirical investigations in this domain.

**Acknowledgements:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, 34(4), 468-479.
- André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., ... & Yang, H. (2018). Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions*, 5(1-2), 28-37.
- Baker, S. M., Gentry, J. W., & Rittenburg, T. L. (2005). Building understanding of the domain of consumer vulnerability. *Journal of Macromarketing*, 25(2), 128-139.
- Bakpayev, M., Baek, T. H., van Esch, P., & Yoon, S. (in press). Programmatic creative: AI can think but it cannot feel. *Australasian Marketing Journal*.
- Beverland, M. B., Lindgreen, A., & Vink, M. W. (2008). Projecting authenticity through advertising: Consumer judgments of advertisers' claims. *Journal of Advertising*, 37(1), 5-15.
- Boerman, S. C., Willemsen, L. M., & Van Der Aa, E. P. (2017). "This post is sponsored": Effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook. *Journal of Interactive Marketing*, 38, 82-92.
- Campbell, C., & Evans, N. J. (2018). The role of a companion banner and sponsorship transparency in recognizing and evaluating article-style native advertising. *Journal of Interactive Marketing*, 43, 17-32.
- Campbell, M. C., Mohr, G. S., & Verlegh, P. W. (2013). Can disclosures lead consumers to resist covert persuasion? The important roles of disclosure timing and type of response. *Journal of Consumer Psychology*, 23(4), 483-495.
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809-825.
- Chan, C., Ginosar, S., Zhou, T., & Efros, A. A. (2019). Everybody dance now. *Proceedings of the IEEE International Conference on Computer Vision*. 5933-5942.
- Chaouali, W., Souiden, N., & Ladhari, R. (2017). Explaining adoption of mobile banking with the theory of trying, general self-confidence, and cynicism. *Journal of Retailing and Consumer Services*, 35, 57-67.
- Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*, 98(1), 147-155.
- Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, 539-548.
- Culic, I., Radovici, A., & Rusu, C. (2020). Smart advertising system. In *Commercial and Industrial Internet of Things Applications with the Raspberry Pi* (pp. 125-162). Berkeley, CA: Apress.
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24-42.
- De Jans, S., Cauberghe, V., & Hudders, L. (2019). How an advertising disclosure alerts young adolescents to sponsored vlogs: The moderating role of a peer-based advertising



- literacy intervention through an informational vlog. *Journal of Advertising*, 47(4), 309-325.
- Deepfake Report Act of 2019, S. 2065, 116th Cong., 1st Sess. (2019).
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126.
- Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2019). Do (microtargeted) deepfakes have real effects on political attitudes?. *The International Journal of Press/Politics*, 1-23.
- Freiden, J. B. (1984). Advertising spokesperson effects: an examination of endorser type and gender on two audiences. *Journal of Advertising Research*, 24(5), 33-41.
- Füller, J., Mühlbacher, H., Matzler, K., & Jawecki, G. (2009). Consumer empowerment through internet-based co-creation. *Journal of Management Information Systems*, 26(3), 71-102.
- Gardner, W. L., & Knowles, M. L. (2008). Love makes you real: Favorite television characters are perceived as “real” in a social facilitation paradigm. *Social Cognition*, 26(2), 156-168.
- Garretson, J. A., & Niedrich, R. W. (2004). Spokes-characters: Creating character trust and positive brand attitudes. *Journal of Advertising*, 33(2), 25-36.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Proceedings of the Neural Information Processing Systems Conference*. 2672-2680.
- Griskevicius, V., Shiota, M. N., & Neufeld, S. L. (2010). Influence of different positive emotions on persuasion processing: A functional evolutionary approach. *Emotion*, 10(2), 190-206.
- Gummesson, E. (2008). Extending the service-dominant logic: from customer centricity to balanced centricity. *Journal of the Academy of Marketing Science*, 36(1), 15-17.
- Helms, K. J., & Mutter, D. (2020). Deepfakes could mean deep losses for businesses if not prepared. Retrieved from:  
[http://www.prweb.com/releases/deepfakes\\_could\\_mean\\_deep\\_losses\\_for\\_businesses\\_if\\_not\\_prepared/prweb17431805.htm](http://www.prweb.com/releases/deepfakes_could_mean_deep_losses_for_businesses_if_not_prepared/prweb17431805.htm)
- Hern, A. (2020, January, 07). Facebook bans 'deepfake' videos in run-up to US election. *The Guardian*. Retrieved from:  
<https://www.theguardian.com/technology/2020/jan/07/facebook-bans-deepfake-videos-in-run-up-to-us-election>
- Hosany, S., Prayag, G., Martin, D., & Lee, W. Y. (2013). Theory and strategies of anthropomorphic brand characters from Peter Rabbit, Mickey Mouse, and Ronald McDonald, to Hello Kitty. *Journal of Marketing Management*, 29(1-2), 48-68.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*, New Haven, CT: Yale University Press.
- Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.

- Jee, C. (2020). An Indian politician is using deepfake technology to win new voters. Retrieved from: <https://www.technologyreview.com/f/615247/an-indian-politician-is-using-deepfakes-to-try-and-win-voters/>
- Kalyanaraman, S., & Sundar, S. S. (2006). The psychological appeal of personalized content in web portals: Does customization affect attitudes and behavior?. *Journal of Communication*, 56(1), 110-132.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.
- Kay, S., Mulcahy, R., & Parkinson, J. (2020). When less is more: the impact of macro and micro social media influencers' disclosure. *Journal of Marketing Management*, 36(3-4), 248-278.
- Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition & Emotion*, 13(5), 505-521.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat?. *Business Horizons*, 63(2), 135-146.
- Kietzmann, J., Paschen, J., & Treen, E. (2018). Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. *Journal of Advertising Research*, 58(3), 263-267.
- Kietzmann, T. C., McClure, P., & Kriegeskorte, N. (2019). Deep neural networks in computational neuroscience. In *Oxford Research Encyclopedia of Neuroscience*. Oxford, UK: Oxford University Press.
- Koh, Y., & Wells, G. (2018). The making of a computer-generated influencer. *The Wall Street Journal*. Retrieved from: <https://www.wsj.com/articles/the-making-of-a-computer-generated-influencer-11544702401>
- Kwok, A. O., & Koh, S. G. (2020). Deepfake: a social construction of technology perspective. *Current Issues in Tourism*, 1-5.
- Lai, Y. H., & Lai, S. H. (2018). Emotion-preserving representation learning via generative adversarial network for multi-view facial expression recognition. *13th IEEE International Conference on Automatic Face & Gesture Recognition* (pp. 263-270), Xi'an, China: Conference Publishing Services.
- Letheren, K., Russell-Bennett, R., Mulcahy, R., & McAndrew, R. (2019). Rules of (household) engagement: technology as manager, assistant and intern. *European Journal of Marketing*, 53(9), 1934-1961.
- Li, H. (2019). Special section introduction: Artificial intelligence and advertising. *Journal of Advertising*, 48(4), 333-337.
- Liang, T. P., Lai, H. J., & Ku, Y. C. (2006). Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems*, 23(3), 45-70.
- Libai, B., Bart, Y., Gensler, S., Hofacker, C. F., Kaplan, A., Kötterheinrich, K., & Kroll, E. B. (in press). Brave new world? On AI and the management of customer relationships. *Journal of Interactive Marketing*.

- Liffreing, I. (2016). Fake news sites spark Pepsi boycott, with other brands in crosshairs. Retrieved from: <https://www.campaignlive.com/article/fake-news-sites-spark-pepsi-boycott-brands-crosshairs/1416027>
- Lowry, B. (2017). Carrie Fisher's death renews questions about digital re-creation. Retrieved from: <https://edition.cnn.com/2017/01/16/entertainment/carrie-fisher-digital-recreation/index.html>
- Lui, M. (2019). The robot wears Prada: what happens when AI starts giving out fashion tips? Retrieved from: <https://theconversation.com/the-robot-wears-prada-what-happens-when-ai-starts-giving-out-fashion-tips-125140>
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6), 937-947.
- MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. *Journal of Marketing*, 75(4), 136-154.
- Malaria Must Die. (2019, April 08). David Beckham speaks nine languages to launch malaria must die voice petition. [Video file]. <https://www.youtube.com/watch?v=QiiSAvKJIHo>
- Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof*, 23(3), 255-262.
- Mende, M., Scott, M. L., van Doorn, J., Grewal, D., & Shanks, I. (2019). Service robots rising: How humanoid robots influence service experiences and elicit compensatory consumer responses. *Journal of Marketing Research*, 56(4), 535-556.
- Mulcahy, R., Letheren, K., McAndrew, R., Glavas, C., & Russell-Bennett, R. (2019). Are households ready to engage with smart home technology?. *Journal of Marketing Management*, 35(15-16), 1370-1400.
- Nuñez, M. (2019). Facebook removes hundreds of fake pro-Trump accounts using AI-generated profile photos. Retrieved from: <https://www.forbes.com/sites/mnunez/2019/12/20/facebook-removes-hundreds-of-fake-pro-trump-accounts-using-ai-generated-profile-photos/#170794956175>
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39-50.
- Powers, K. (2019). Virtual influencers are becoming more real—here's why brands should be cautious. Retrieved from: <https://www.ama.org/marketing-news/virtual-influencers-are-becoming-more-real-heres-why-brands-should-be-cautious/>
- Rauschnabel, P. A., & Ahuvia, A. C. (2014). You're so lovable: Anthropomorphism and brand love. *Journal of Brand Management*, 21(5), 372-395
- Saito, Y., Takamichi, S., & Saruwatari, H. (2017). Statistical parametric speech synthesis incorporating generative adversarial networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(1), 84-96.
- Schuster, L., Proudfoot, J., & Drennan, J. (2015). Understanding consumer loyalty to technology-based self-services with credence qualities. *Journal of Services Marketing*, 29(6/7), 522-532.

- Seo, K., & Dillard, J. (2019). The persuasive effects of two stylistic elements: Framing and imagery. *Communication Research*, 46(7), 891-907.
- Shah, D., Rust, R. T., Parasuraman, A., Staelin, R., & Day, G. S. (2006). The path to customer centricity. *Journal of Service Research*, 9(2), 113-124.
- Sherwin, R. K., Feigenson, N., & Spiesel, C. (2006). Law in the digital age: How visual communication technologies are transforming the practice, theory, and teaching of law. *Boston University Journal of Science & Technology Law*, 12(2), 227-270.
- Spry, A., Pappu, R., & Cornwell, B. T. (2011). Celebrity endorsement, brand credibility and brand equity. *European Journal of Marketing*, 45(6), 882-909.
- Stafford, M. R., Stafford, T. F., & Day, E. (2002). A contingency approach: The effects of spokesperson type and service type on service advertising perceptions. *Journal of Advertising*, 31(2), 17-35.
- Stupp, C. (2019, August 30). Fraudsters used AI to mimic CEO's voice in unusual cybercrime case. *The Wall Street Journal*. Retrieved from: <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402>
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). The MIT Press.
- The Dali Museum. (2019, May 08). Behind the scenes: Dali lives. [Video file]. <https://www.youtube.com/watch?v=BIDaxl4xqJ4>
- Till, B. D. (1998). Using celebrity endorsers effectively: Lessons from associative learning. *Journal of Product and Brand Management*, 7(5), 400-407.
- Tractica. (2019). Artificial intelligence software market to reach \$118.6 billion in annual worldwide revenue by 2025. Retrieved from: <https://tractica.omdia.com/newsroom/press-releases/artificial-intelligence-software-market-to-reach-118-6-billion-in-annual-worldwide-revenue-by-2025/>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1-13.
- van Esch, P., Black, J. S., Franklin, D., & Harder, M. (in press). AI-enabled biometrics in recruiting: Insights from marketers for managers. *Australasian Marketing Journal*.
- Whittaker, L. W., Kietzmann, T. C., Kietzmann, J. & Dabirian, A. (2020). “All around me are synthetic faces”: the mad world of AI-generated media. *IT Professional*, 22(5), 90-99.
- Williams, R. (2020). Doritos lets mobile fans create Super Bowl dance videos with AI app. <https://www.mobilemarketer.com/news/doritos-lets-mobile-fans-create-super-bowl-dance-videos-with-ai-app/571361/>
- Wilson, H. J., & Daugherty, P. R. (2018). Collaborative intelligence: humans and AI are joining forces. *Harvard Business Review*, 96(4), 114-123.
- Woollacott, E. (2019). China bans deepfakes in new content crackdown. Retrieved from: <https://www.forbes.com/sites/emmawoollacott/2019/11/30/china-bans-deepfakes-in-new-content-crackdown>