Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law & Security Review

The present looks nothing like the Jetsons: Deceptive design in virtual assistants and the protection of the rights of users

Silvia De Conca

Assistant Professor in Law & Technology, Transnational Legal Studies Department, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Deceptive design
Dark patterns
Virtual assistants
GDPR
Consumer protection

ABSTRACT

The increasing popularity of virtual assistants (VAs) raises concerns about deceptive design (also referred to as dark patterns), that is, design tricks to influence users into buying or engaging more, hijacking their decision-making capability. The article argues that some recurring responses and prompts recited by VAs amount to deceptive design, matching some well-known dark patterns such as Price Comparison Prevention and Misdirection. It analyses the challenges of applying the EU consumer and personal data protection laws to deceptive design in VAs, exploring provisions on unfair practices and consumers' rights (UCPD, CRD), fairness of personal data processing (GDPR), as well as the new rules on digital services (DSA) and the proposals for Data Act and AI Act. While the current legal framework offers a sufficient starting point to address the complexity and sophistication of deceptive design in virtual assistants, additional guidance on its application and interpretation is necessary to guarantee a high level of protection of the rights and interests of VA users. This article contributes to the ongoing debate on the regulation of deceptive design, and brings attention to the specific challenges deriving from virtual assistants due to the use of vocal interface, which can bring a paradigmatic shift from the legal perspective.

1. Introduction

The term Virtual Assistant (VA) indicates a software that allows users to operate smart devices via voice commands. VAs are embedded into smart speakers, smartphones, and many other appliances, and are marketed to consumers as the personal assistant that will simplify the life of the whole family. The attentive readers have probably recognized them already: they are known to consumers as Amazon Alexa and Google Assistant.[1]

VAs receive data from sensors (detecting voices and sounds from the environment) and use Natural Language Interface (NLI) to make users access services (online stores, search engines, social networks, etc.) using voice commands. The vocal

[1] At the time of writing, Amazon Alexa and Google Assistant, respectively embedded into the Echo and Nest/Home product lines, are the dominant players in the European Union. The third most popular VA in the European Union is Apple's Siri, but this latter presents a more limited range of abilities and, for this reason, is deemed less relevant for the purposes of this work.

E-mail address: s.deconca@vu.nl

https://doi.org/10.1016/j.clsr.2023.105866

interface and all the other capabilities of VAs are powered by machine learning and by the collection of large amounts of personal data through said sensors and the sensors of connected IoT devices. To build a long-term relationship with the users, VAs have been designed to prompt and influence individuals to talk to them (dialogically) to purchase products, hear the news, listen to music, and visit web pages, consequently sharing data on a constant basis. To influence individuals, the user interface (UI) and the user experience (UX) of VAs are organized using techniques commonly indicated with the term 'deceptive design'. Deceptive design was first identified by UX designer Harry Brignull in the early 2010s. Initially, Brignull used the term 'dark patterns' to indicate "tricks used in websites and apps that make you do things that you didn't mean to, like buying or signing up for something."[2] Deceptive design – or dark patterns –[3] is used to influence individuals into buying products or services they would not otherwise purchase, sharing more data than intended, locking them into subscriptions, preventing product comparison, and so on. Online, we run into it daily, for example when a cookie pop-up has a very prominent, bright 'accept' button and a grey, small 'reject' one.

Deceptive design has recently gone under scrutiny by EU and national authorities, who are set to prevent companies from bypassing existing consumer and data protection laws. The European Data Protection Board (EDPB), for example, defines dark patterns as "interfaces and user experiences implemented on social media platforms that lead users into making unintended, unwilling and potentially harmful decisions in regards of their personal data. Dark patterns aim to influence users' behaviours and can hinder their ability "to effectively protect their personal data and make conscious choices", for example by making them unable "to give an informed and freely given consent". This can be exploited in several aspects of the design, such as interfaces' colour choices and placement of the content."[4]

The European Consumer Association (BEUC) defines them as "online interface or a part thereof that via its structure, function, or manner of operation, subverts or impairs the autonomy, decision-making, or choice of recipients of the service. (…) Specific features of "dark patterns" include relying on user interfaces to influence and manipulate users, subverting intent or preferences and abusing knowledge of human behaviour to predict decisions of users and influence them. Dark patterns can be data-driven and personalised or implemented on a more general basis, tapping into heuristics and behavioural biases, such as default effects or scarcity biases."[5] Currently, there is no consensus on a universal definition or taxonomy of deceptive design techniques, due also to the variety of forms in which it manifests among different websites and interface designs. There have been several attempts at cataloguing and organizing deceptive design techniques, and it is possible to identify a core of dark patterns that recur in every taxonomy, such as Nagging or Misdirection (defined in Section 2 below). Dark patterns can be grouped based on the mechanisms they leverage or on the effects they have on users, but what they all have in common is that a certain design solution hijacks the decision-making and autonomy of individuals, leading to behaviours that, in the absence of the dark pattern, the individuals would not carry out. As will be explained throughout this article, the characteristics of deceptive design make it difficult to regulate. Influencing or persuading customers for marketing purposes is not new, and it is tolerated up to a certain point: drawing the line between acceptable design and harmful, deceptive design is no simple task.

Mindful of the characteristics of VAs and deceptive design, I argue that the use of deceptive design in VAs creates gaps in the protection of the rights and interest of users. Consequently, this article answers the following research question: "How suitable are the EU consumer and personal data protection regimes – also in combination with new adjacent Regulations – to protect the rights of users of VAs vis-à-vis deceptive design?".

To answer the research question, this article maps the deceptive design techniques deployed by Virtual Assistants to influence users into sharing more data and buying products, exploring how existing (and incoming) secondary EU laws apply to them.

The contribution of this article is twofold. First and foremost, it identifies the ways in which deceptive design is used in Virtual Assistants at the beginning, during, and at the end of a conversation with users, using mostly vocal interfaces. The second contribution lies in assessing the application of a selection of legal provisions, focusing in particular on articles from the Unfair Commercial Practice Directive, Consumer Rights Directive, General Data Protection Regulation, Digital Services Act, and the proposals for a Data Act and AI Act. VAs serve as case study to analyse the challenges posed by deceptive design to existing (and prospect) EU secondary laws, discussing their application and coordination. Additionally, the legal analysis of this article exemplifies the challenges that vocal interface poses for the law in general, as this latter often still operates under the assumptions that legally relevant relationships are mediated by written text, or at least digital displays. Voice interaction, however, is a paradigmatic shift, and while the abovementioned laws still apply to it, they also require adjustments and additional guidelines, to ensure that individuals using voice interaction are as protected as those using a screen.

Section 2 explains how VAs work and introduces the deceptive design techniques used in VAs, mapping them against similar existing categories of deceptive design: Nagging, Privacy Zuckering, Misdirection, Disguised Advertising, Price Comparison Prevention, and Roach Motel. Section 3 discusses the application of relevant provisions from European secondary legislative tools. Section 3 touches upon, among others, the traditional distinction between average and vulnera-

[2] Brignull, https://www.deceptive.design last accessed 4 April 2023.

[3] Throughout this work the two terms are used as synonyms, although deceptive design is currently becoming more popular among designers and experts and it is, therefore, preferred in this article too.

[4] EDPB, Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognize and avoid them, 14 March 2022, p.7.

[5] BEUC, "Dark Patterns" and the EU Consumer Law Acquis – Recommendations for better enforcement and reform, BEUC-X-2022-013, 7 February 2022, p. 5.

ble consumers (UCPD), the identification of the limits of space and time when the consumer is provided information via voice interface (CDR), the principle of fairness in the GDPR, and is complemented by an analysis of the new Digital Services Act, and the proposals for Data Act and AI Act.

2. Virtual assistants

VAs receive commands and complete tasks prevalently via vocal interaction. Users say the wake-word – "Hey, Google" or "Alexa" – followed by a command. If no wake-word is detected, the VA remains semi-dormant: in this state, if the device has a screen, images are displayed on rotation. On Google Nest devices these include pictures from photo-albums pre-selected by the user, or so-called 'memories' (old pictures taken on the same day in the past). In the case of Alexa, the rotation includes 'suggestions' about news or products, curated by Amazon.[6]

If the wake-word is detected, the device starts recording every sound within the range of its sensors, streaming it to the Cloud, where the vocal command is transcribed, analysed, and stored.[7] After the task is carried out and completed, the VA goes back to the semi-dormant state. For example, if the user wants to ask the VA about traffic, they say the wake-word, then ask: "How is traffic?". The VA repeats the command ("Checking for traffic") to make sure it was understood correctly. The request of the user is recorded and sent to Cloud servers. There, the voice command is translated into text and analysed for keywords. Once the keyword is identified, the appropriate app is opened, or the information is searched on the Internet. As the information is retrieved, the VA translates it from text to speech, and recites it back to the user: "Today, traffic along the route is intense".

VAs use machine learning to deduce preferences from repeated and routine behaviours, and compile a detailed, granular profile of each user. The profiles are used to personalise the service, but also for marketing, data brokerage, and advertising.[8] The intelligence of VAs depends entirely on the collection, processing, and storing of personal data.

2.1. *Why VAs work*

The vocal interface presents some advantages, but also some limitations, compared to the visual one.

For example, the vocal interface is intuitive and easy, but *volatile* and *linear*: this means that users can easily interact with a VA, but the ways in which VAs present information does not allow users to go back and re-read something, and users must wait for the VA to complete its sentences before being able to go forward. Additionally, with the sole voice interaction users receive less contextual clues about the website from which the VA is retrieving information, and research points in the direction of users being unaware sometimes that some apps downloaded on the smart speakers are managed by third parties, not by Amazon or Google.[9]

Design can leverage the affordances of vocal interaction, to build a user-VA long term relationship. This is important for the producers of digital products or services in general. Digital products and services have a relatively short 'shelf-life', while at the same time it can take a few years before the stream of revenues stabilises. The long-term relationship of a brand with its customers becomes fundamental to ensure that producers gain a dominant position in the market, increase and stabilise their revenues, lock customers into their product lines, or gain share value before being bought by bigger competitors or going public.[10] VAs are no exception. Additionally, the user-VA relationship can be exploited later in time. Small, apparently insignificant actions build up over time, creating dependence and trust, that can be turned into leverages by the VA manufacturer. For example, if several suggestions are welcomed and valued by the users, after an initial time the VA can start offering products from sponsorships or content behind paywall.[11] This is enhanced by profiling, that allows the producers of VAs to monetize the user's habits, identifying vulnerabilities: What circumstances make the user more likely to buy? What words are more inviting for the user? What emotions is the user experiencing?

It is telling, in this regard, that both Google and Amazon are developing and patenting software for emotion and health recognition from voice data. In the Amazon's patent application, an example is shown in which the user has a cold, and the VA offers a recipe for chicken soup, or the 1-hour delivery of cough drops purchased through the proprietary online store.[12]

Since the late 1990s, computer and design experts have been perfecting techniques and strategies to optimise the experience of users with digital and online products. The optimization can aim at making the overall experience more pleasant or efficient for users, but also at persuading users into a behaviour, like accepting a cookie consent pop-up or placing in-game purchases. Certain characteristics enhance the potential influence that a machine can exercise on its users: psychological cues exploit the instinctual positive predisposition of individuals vis-á-vis similar individuals, em-

[6] 'Alexa Talks Politics, but Avoids Republicans, Democrats, and Trump' (VentureBeat, 6 November 2018) https://venturebeat.com/2018/11/06/alexa-talks-politics-but-avoids-republicans-democrats-and-trump/> accessed 22 October 2020.

[7] EDPB, Guidelines 3/2022 (n. 4).

[8] Umar Iqbal and others, 'Your Echos Are Heard: Tracking, Profiling, and Ad Targeting in the Amazon Smart Speaker Ecosystem' (arXiv, 20 February 2023) <http://arxiv.org/abs/2204.10920> accessed 4 April 2023.

[9] Kentrell Owens and others, 'Exploring Deceptive Design Patterns in Voice Interfaces', Proceedings of the 2022 European Symposium on Usable Security (ACM 2022) https://dl.acm.org/doi/10.1145/3549015.3554213> accessed 4 April 2023.

[10] Arvind Narayanan and others, 'Dark Patterns: Past, Present, and Future: The Evolution of Tricky User Interfaces' (2020) 18 Queue 67.

[11] Owens and others (n 8). A similar experience is also described by a Reddit user: https://www.reddit.com/r/alexa/comments/12ehy67/find_my_phone_skill_no_longer_free/ accessed 23 June 2023.

[12] 'Amazon's Alexa Can Now Act on "Hunches" about Your Behavior' [2018] DigiTechNews https://digitechnews.net/amazons-alexa-can-now-act-on-hunches-about-your-behavior/> accessed 26 February 2020.

pathic signals, or affiliation to the same groups. Language can generate a positive predisposition (imagine a pop-up message complimenting a gamer for a well-played match) or convey the impression of personality, and so on.[13]

NLI is a powerful tool to convey language and psychological cues: the witty and personalised responses of Alexa or Google Assistant generate into users a sense of communality and affiliation, and the answers contain positive reinforcements and praises. Slang, mannerisms, sarcasm, and other vocal cues induce users to attribute a personality to the VA, and reinforce the impression of interacting with a peer. Alexa's jokes and sarcastic answers, as well as its carefully written background stories, give it personality. The language of Google Assistant, Alexa, and even Siri, is never demanding or too direct, but always empathic and cordial.[14] Devices that are connected to the internet have an advantage, as they can have access to updated information and personalise their outputs. Persuasion is also enhanced when a machine's suggestion occurs in the right place and at the right time.[15] Devices that are mobile, like smartphones, or embedded into the environment, like the IoT in the smart home, have an inherent advantage. Connectivity, interaction, placement, personalization: these features are at the very basis of the design of VAs, making them potentially very persuasive.

VAs can be habit-forming products, and 'hook' users into a long-term relationship, creating a form of addiction via the User Experience, not just the interface: users develop the habit to open an app or use a device every time they feel basic needs (e.g. if they are bored or want to check the weather), thanks to the easiness of use that will satisfy their necessity without much effort. From there, they will obtain a positive feeling or something perceived as a reward (such as the joy of watching a funny video, buying something, or hearing a joke), and they will come back, sharing more data, liking more posts, buying more products, and so on in a loop.[16]

The habits formed by VAs are not neutral: they are infused with the business model of the producers. Google's business model, for instance, relies on data brokerage: maximising the interactions is functional to increasing the amount of data collected. In the case of Alexa, although its business model is not clear yet, online purchasing on the Amazon platforms plays a role.[17] When a user asks Alexa to buy something, Alexa suggests first an 'Amazon's choice' item from the Amazon online store.[18] The selection of those products is at complete discretion of Amazon, and users only receive a vocal description of the offer. In these occasions, business interests and market mechanisms enter the conversations users entertain with VAs.

It is important to point out that the technical and design features enabling the positive user experience are the same that enable the producers to maximise profits.[19] The core service is inextricably connected to marketing and advertising. The issue is where and how should the regulator trace the line between acceptable marketing practices implemented at the level of UI/UX, and undesired, harmful, deceptive design practices. Where the threshold lies between deceptive and acceptable design, and what harms need the intervention of regulators remain unanswered questions. A possible approach would be to consider deceptive design inherently incompatible with user autonomy, since it tries to suppress, constraint, or leverage the decisions of users, so that the interests of another party (a service provider or other company) can prevail. As will be explained below, this brings the risk of over-inclusiveness and would impair the development of VAs and many more digital products.

At the same time, deceptive design raises serious ethical and normative issues. Scholarship has already highlighted that interfering with the decision-making of individuals affects their autonomy and self-determination and, consequently, human dignity.[20] When VAs' persuasive techniques interfere with individuals and their decisions to harvest more data, or to make them purchase more, individuals are datafied. This consideration is not unique to VAs, as it can apply to other digital services and products, such as social media platforms. VAs, however, present two peculiarities that make them particularly interesting (and worrying) from a deceptive design perspective. First, VAs' use of the vocal interface allows them to granularly profile every inhabitant in the house – identified thanks to their individual voice profiles in a way that a smartphone or smart TV (without VA) cannot do.

Second, VAs are placed in the heart of the private sphere of individuals: the home. Besides the loss of autonomy, users lose control over the home environment, leading to the erosion of the private sphere. VAs also contribute to the blurring of the boundaries between online and offline. Before VAs, off-platform tracking meant that a company such as Google or Facebook could track individuals on other websites. With VAs, it means that companies can track users in their daily, offline, lives. Deceptive design is a tool to make sure users keep the VAs in their home and even depend on them, enabling the opening of new windows into their private sphere.[21]

[13] BJ Fogg and GE Fogg, Persuasive Technology: Using Computers to Change What We Think and Do (Morgan Kaufmann 2003) 136.

[14] Silvia De Conca, 'The Enchanted House: An Analysis of the Interaction of Intelligent Personal Home Assistants (IPHAs) with the Private Sphere and Its Legal Protection' (Tilburg University 2021).

[15] Fogg and Fogg (n 12).

[16] Nir Eyal, 'Here's How Amazon's Alexa Hooks You' (Medium, 5 September 2019) https://medium.com/behavior-design/heres-how-amazon-s-alexa-hooks-you-1b46ee9c92f6> accessed 15 April 2020.

[17] Eugene Kim, 'As Amazon Floods the Market with Alexa Devices, the Business Model Is Getting Fresh Scrutiny' [2019] CNBC https://www.cnbc.com/2019/09/28/amazon-alexa-growth-has-investors-questioning-the-business-model.html> accessed 15 April 2020.

[18] Go Shopping with Alexa - Amazon Alexa (2020) https://www.youtube.com/watch?v=iQD2waZNCao>. Accessed 4 April 2023.

[19] Marijn Sax, 'Optimization of What? For-Profit Health Apps as Manipulative Digital Environments' [2021] Ethics and Information Technology.

[20] Daniel Susser, Beate Roessler and Helen F Nissenbaum, 'Online Manipulation: Hidden Influences in a Digital World' (2019) 4 GEO. L. TECH. REV.

[21] De Conca (n.14)

2.2. *VAs and deceptive design*

VAs are an excellent example to highlight the difficulty of regulating deceptive design in digital products or services because of the way in which deceptive design, exploitative profiling, and normal functionality build on each other. The profiling that enables a VA to give accurate answers to its users also enables the producers to profile them for marketing and to identify 'weaknesses' that will make the user engage more. The user is not only the target audience of the service, but also the means for the corporate's ends. The affordances of vocal interface can be used to improve functionality and, at the same time, diminish the control and autonomy of users.

What the VA says (or displays) is a mix of pre-programmed and personalized outputs. For this reason, the responses it gives to different users might have some parts in common, but also differ in part. The responses also change over time, due to machine learning. It would be impossible to identify all the answers that a VA gives its users, but it is possible to notice some recurring elements and similar categories of replies. This legal analysis is based on common replies and prompts, identified as follows. An initial exploration of the official promotional material and official customer support of Google and Amazon allowed me to pinpoint some very common replies given by the respective VAs. These were: the replies in which the VA suggests users that it can be used for additional tasks too, commercial offers showcasing only one selected product for sale, and (on devices equipped with a display) the suggestion to look at 'trending topics'. The latter was also described and discussed in a number of news articles.[22] The first two replies were also registered in a study by Owens et al. concerning VAs and deceptive design (to date, the only study available on the subject).[23] The same study also identified additional replies considered deceptive, and these have been added to this analysis too. This combination of sources gave me a short but consolidated list of instances in which the behaviour of the VA appears deceptive. To observe as many concrete examples as possible, I resorted again to the information made available on the VAs official tech support forums, and to the questions shared by users on the Reddit threads dedicated to smart speakers (r/smarthome, r/googlehome, r/alexa, r/amazonecho). Based on the knowledge gained from the other sources, I searched on said subreddits using the keywords "By the way", "You can also ask me", "Alexa spam", "Alexa advertising", "Amazon's choice", "Other people also asked", and "Marketing". Doing so, I encountered hundreds of users' comments describing variations of the same feature, or of the same type of reply given by their VAs. Besides offering a user-centric perspective, the uniformity and consistency of the descriptions posted by users helped identifying recurring sentences spoken by the VA. This multi-focal approach allowed me to overcome the difficulty deriving from the ever-learning, personalized voice interface of VAs, and identify multiple ways in which VAs try to influence users. Finally, I grouped the identified instances into a few main categories, based on when they happen and the type of interface used: vocal prompts given at the beginning of a conversation with the users; visual prompts given while the VA is dormant; and strategic replies occurring during or at the end of a conversation.

As explained in Section 1, deceptive design is a term coined to indicate UI/UX design solutions that deceive and trick users of websites or apps. The diffusion of VAs with their vocal interface begs the question whether the aforementioned VA replies and prompts belong to the categories of deceptive design identified so far in relation to websites. The answer is yes. I compiled a list of deceptive design techniques used in e-commerce and personal data processing context, based on frequently cited taxonomies of deceptive design techniques (3 made by European, USA and international institutions, and 3 academic, well-known among the CHI and legal communities).[24] Based on the descriptions provided in the taxonomies for each dark pattern, I compared the VAs replies and excluded the deceptive designs that did not seem to: cause the same or a similar effect on users; leverage similar design solutions; were incompatible with voice interface. Based on this analysis, I believe the abovementioned types of prompts and replies given by VAs match the following types of deceptive design: Nagging, Privacy Zuckering, Disguised Advertising, Misdirection, Price Comparison Prevention, and Roach Motel (as described below). One preliminary conclusion, as that point, was that the identified VA deceptive replies represent a new form in which these established types of deceptive design manifest in connection with NLI, as they have in common the effect on users, the purposes and, in some cases, the modalities to achieve such purposes (although with some differences due to the vocal interface).

Nagging consists of sending pop-ups and prompts frequently and/or at inconvenient times, to exasperate a user into taking a particular action.[25] A typical example is the invitation to purchase a license that pops up every time the file-compression software WinZip is used. As explained above, when the user asks: "How is traffic?" the VA first repeats the command. At this point, the VA might add a suggestion. For example, the VA might say: "Checking for traffic. By the way, you can now ask me to turn off the lights. Just say 'turn off the lights in …'".[26] The inputs aim at informing the users of ad-

[22] The exact sources for the deceptive replies are indicated in the rest of the article, every time they are discussed.

[23] Owens and others (n 8).

[24] Institutional: European Data Protection Board (EDPB), Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them (2022); Federal Trade Commission, Bringing Dark Patterns to Light Staff Report (2022); OECD, Dark commercial patterns (2022). Academic: Christoph Bösch et al. 2016. Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns. 2016, 4 (2016), 237–254.; Colin M. Gray, et al. 2018. The Dark (Patterns) Side of UX Design. In Proceedings of the ACM on Human-Computer Interaction (ACM 2018); Arunesh Mathur et al. 2019. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 81 (nov 2019), 32 pages. https://doi.org/10.1145/3359183.

[25] Arunesh Mathur and others, 'Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites' (2019) 3 Proceedings of the ACM on Human-Computer Interaction 1.

[26] This reply is the most reported among the undesired Alexa features. The example is inspired from a post left by a user on reddit: https://www.reddit.com/r/amazonecho/comments/shxpn7/i_dont_understand_what_she_meant_by_this_and_she/ accessed 8

ditional ways in which the VA can help them, and users cannot skip them, but they must wait for the VA to finish reciting them, because of the linearity of vocal interface. They are often considered frustrating[27] – but nevertheless users might remember the prompt next time they forget to switch off the light or go homeware shopping and see smart light fixtures. They give users the trigger to open an app of the VA, or the affiliated online store, similarly to what push notifications do on smartphones. Periodically suggesting users what else their VAs can do helps building a long-term relationship, leveraging psychological cues and hooking individuals into using the device more.

Users have also reported that sometimes the VA might persistently ask to register or subscribe to a service, without leaving the possibility to move forward and complete a task until the user has agreed.[28] This practice falls within Nagging too.[29] It is important to point out that this is not a feature included by the producers, but is controlled by the app developers. Due to the inherent lack of context of the vocal interface, users often might not understand whether the app they have opened on a VA is from a third-party or from Amazon/Google. Furthermore, due to the linearity of vocal interface, users might have no choice but to proceed with registering or subscribing, before being able to even just close the app.

Increasing the amount of data shared by users is also the main purpose of **Privacy Zuckering**. This technique, named after Facebook's founder,[30] focuses on making users share more data via engagement, for instance by posting more, as shown in the picture below where a pop-up on Instagram has a prominent button to make users share their stories on Facebook too. Incidentally, the prominence of the share button also integrates another technique, Misdirection (as explained below).[31]

The aforementioned "By the way" input given by VAs at the beginning of a conversation can generate more engagement, which translates into more data being shared by users. Furthermore, sometimes the VA 'beeps' without having been activated. This sound informs the user that there is a notification in one of the apps. After hearing the sound, users can awake the VA and ask what that sound was.[32] Privacy Zuckering helps producers collecting more data or directing them to in-app purchases.

Written suggestions and invitations, displayed on screens also prompt users to share more data. VAs equipped with a display show images on rotation while dormant (similarly to screensavers on old computers). The rotation includes pictures (very common on Google Home) but also curated 'suggestions' (this is particularly so in the case of Amazon). The curated suggestions are indicated as 'trending topics': a variety of news, products on Amazon, or things that users can ask to Alexa. Each suggestion is accompanied by a sentence inviting users to ask more about the suggested topic or product. For example, the screen can show news about climate change, with the text 'Try "Alexa, tell me more about climate change"'[33]. Google Home also provides similar inputs on display, but they only appear if the user swipes right or left on the display and scrolls through them, not as a default rotation. The screen offers an advantage compared to the sole voice interaction, because of the visual prompts, such as smoke coming out a chimney with the climate change trending news.

These suggestions also have other implications in terms of influencing users. Terms such as '*trending* topics' or '*Amazon's choice*' evoke trust and leverage the desire to partake: many people read this news or bought this product, so you should too. In reality, it is not clear where those topics are trending, or what factors determined the choice or ranking. Amazon does not disclose based on what data the rankings are made, or the criteria for curating the news nor, most importantly, the sources.[34] Amazon highlights Alexa's objective and impartial nature.[35] Indeed, the news shown are not directly and expressly about politics. When asked, many VAs present themselves as non-political, or reply with a deflecting joke or a self-celebratory answer. For instance, at the question "Alexa, are you a Democrat/Republican?" Alexa replies (apparently) funny things, such as "When it comes to politics, I support good platforms, like myself", or "When it comes to politics, I like to think big. We should be funding deep space exploration. I would love to answer questions from Mars!".[36] This is a normal strategy to avoid upsetting users from one political faith or another.

However, the way in which the trending topics (or the product suggestions, such as books or music) are curated is most likely not neutral. First, at the time of writing the owner of Amazon also owns a famous international news outlet, which might make it easier for Amazon to display more news from that source, without clearly indicating it.[37] Second, certain topics, such as the climate change example, are inherently political, and displaying them at strategic moments, for example nearing an election, is not neutral either. Furthermore, the answers are infused with the company's ideology, such as the role of private companies in space exploration: consider that the owner of Amazon also owns an aerospace company, Blue Origin, and two years after those answers were recorded, he took his first suborbital space ride on one of his rockets.[38]

June 2023. Google has a similar reply too, that says: "Would you like to know …" or "Other people also asked to…".

[27] As emerges from many comments on the many reddit threads that can be found using the keyword "By the way" to search the reddit pages dedicated to Alexa, on r/amazonecho or r/alexa.

[28] Owens and others (n 8).

[29] Depending on how it manifests it can also amount to so-called Forced Action.

[30] See https://www.deceptive.design accessed 4 April 2023.

[31] This pop up was seen on Instragram on 4 April 2023.

[32] Owens and others (n 8).

[33] 'Alexa Talks Politics, but Avoids Republicans, Democrats, and Trump' (n 6).

[34] Aleks Krotoski, 'Where Does Amazon's Alexa Get Her News from?' Financial Times (10 January 2020) https://www.ft.com/content/eea6df18-fcbc-11e9-a354-36acbbb0d9b6>.

[35] The Amazon Blog: Day One, 'Alexa, Tell Me about the Election' (The Amazon Blog: Day One, 18 September 2019) https://blog.aboutamazon.com/devices/alexa-tell-me-about-the-election> accessed 22 October 2020.

[36] Christopher Ojeda, 'The Political Responses of Virtual Assistants' [2019] Social Science Computer Review.

[37] Aleks Krotoski (n 28).

[38] Paul Rincon, 'Jeff Bezos launches to space aboard New Shepard rocket ship' BBC (20 July 2021) <https://www.bbc.com/news/science-environment-57849364> accessed 4 April 2023.

Disguised Advertising is a message that looks like a suggestion or user-generated content but is actually advertising or sponsored content.[39] VAs have been reported to complete the requests of users in a way that aims at influencing them towards purchasing products or services, or installing an app. When a user asks Alexa to buy something in a generic way ("Alexa, I want to buy a food processor"), Alexa suggests first an 'Amazon's choice' item at a certain price.[40] The selection of those products is at complete discretion of Amazon and, when the VA recites the offer orally, it is impossible to distinguish sponsorships or advertising from genuine popularity rankings. According to Amazon's website, Amazon's choice products are selected based on, among others, customers' ratings and the amount of returns for the same product. They also appear to be only PRIME products, which might have consequences in terms of competition, since PRIME products are only those participating to the Fulfilled-By-Amazon (FBA) program, which is offered to sellers for a fee.[41]

It is safe to assume that the item's price has been optimised. Whether the optimization was based on the profile of the user (price personalization) or on the demand-offer mechanism (dynamic pricing) is not disclosed.[42] This lack of transparency is worsened by the volatility of the vocal interface, that does not allow users to look at multiple choices on a screen, to read twice, or verify if it is a sponsored product, and by the fact that the vocal interface might create a sense of urgency in users.[43] At the same time, by leveraging the fact of being in the right place at the right time, a VA can administer these commercial offers, maximising the purchasing potential of, for example, moments of need, frustration, or elation.

These strategic replies can also take more deceiving forms. Owens et al. reported that sometimes the VA can give mismatched replies, in particular regarding apps. For instance, when the user asks the VA to list all the apps already installed, it offers new apps to download: "Here are a few popular ones. I've got one called NewsUpdate, want to try it? Or you can ask for more options."[44]

If the 'Producer's choice' product or the app is suggested by a VA because of a paid sponsorship or a business agreement between the VA producer and the food processor trader, then the suggestion disguises an advertising.

Misdirection "uses visuals, language, and emotion to steer users toward or away from making a particular choice".[45] Design experts call visual interference a specific Misdirection technique very common on websites.[46] Visual interference relies on shapes, colors, and other visual components displayed on a webpage to make certain information more prominent and make them prevail over others, so that the users will be steered towards them. The image below shows an example of Misdirection/visual interference on a popular clothing website: the button to accept cookies is black and very visible, while the option to select which cookies to install is grey and less visible.[47]

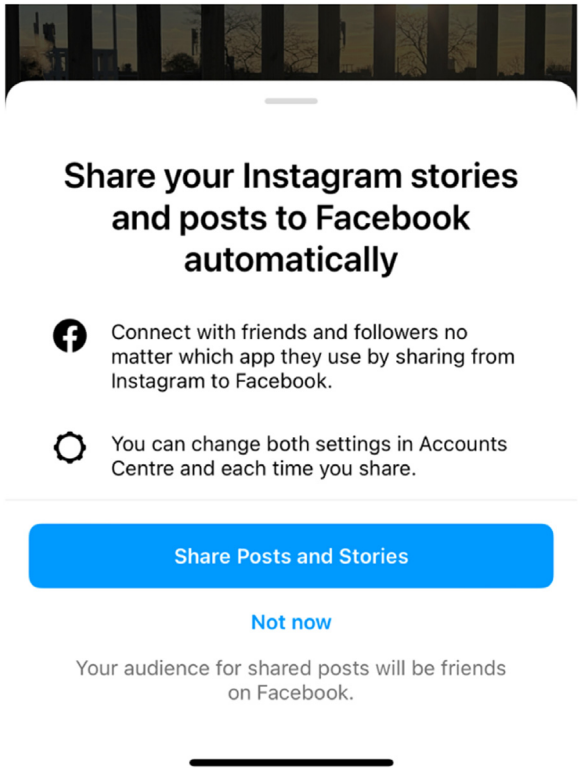

Fig. 1 – An example of Privacy Zuckering (combined with Misdirection).

Fig. 1 and 2. VAs can suggest alternative options when users ask for certain information, using spoken words. Owens et al. report the VA replying to a request to know the time with: "It is 2:45 pm. Also you have some notifications would you like to check them?". When the VA suggests a different action than the one requested by the user it can amount to Misdirection. While it has not been recorded, This type of reply can potentially be used to suggest an alternative based also on business arrangement made with partner companies. To date, this has not been recorded, but should this happen, these strategic replies could also amount to Disguised Advertising.

Due to the vocal interface of VAs, it seems only natural to indicate these suggestions with the term vocal interference. Vocal interferences steer the user from one choice to another, leveraging contextual elements and the voice interface. Vocal interference happens when the user gives a command to the VA, and the VA tries to deflect it by counter-proposing a differ-

[39] Mathur and others (n 21).
[40] Go Shopping with Alexa - Amazon Alexa (n 16).
[41] See Amazon's official PRIME and FBA websites.
[42] This is likely also the pricing mechanism of the Amazon web shop, from which the product is purchased.
[43] Owens and others (n 8).
[44] ibid 6.
[45] OECD, 'Roundtable on Dark Commercial Patterns Online – Summary of Discussions' (DSTI/CP(2020)23/FINAL , 19 February 2012), p. 13.
[46] Colin M Gray and others, 'Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective', Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (ACM 2021) https://dl.acm.org/doi/10.1145/3411764.3445779> accessed 4 April 2023.
[47] This is the cookie consent pop up of Buienradar.nl on 4 April 2022.

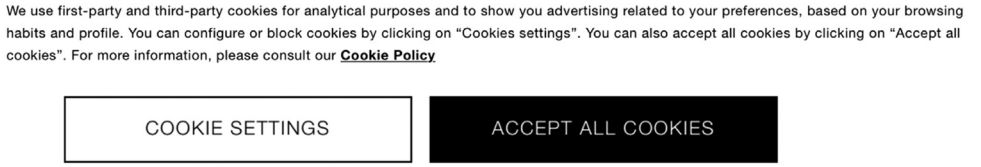

Fig. 2 – An example of Misdirection, specifically visual interference.

Table 1 – Mapping VA deceptive design against already identified, common techniques.

VA deceptive design	Popular Deceptive Design techniques					
	Nagging	Privacy Zuckering	Disguised Advertising [if service or product is sponsored]	Price Comparison Prevention	Misdirection	Roach Motel
"By the way, did you know that you can also use me for…" –"Other users also asked for"	X	X				
Asking users to register or subscribe to a service before being able to move on in a dialogue or close the app	X					
Trending topics on VAs with display with "try asking …" prompt		X	X			
User: "I want to buy a food processor" VA: "I have found a Producer's choice food processor for you, do you want to buy it?"			X	X		
User: "Tell what apps are installed" VA:"Here are some popular apps, do you want to download them?"			X		X	
User: "What time is it?" VA: "It's 1.45pm. Would you like to open the notifications?"			X		X	
VA: "To unsubscribe, please open the app on your smartphone"						X

ent task or command ("Tell me the time" - "Do you also want to open your notifications?"), or when the VA offers new apps to download instead of the app specifically requested by the user.

Price Comparison Prevention consists of organizing the information (usually on screen) in a way that makes it harder to compare the prices of similar items.[48] For example, online stores can display next to each other two liquid soaps from two different brands, indicating for one the price per litre, and for the other the price per bottle of 250 ml. The consumer might be tricked into buying the bottle because it has the lowest price without realizing it is more expensive per litre. The lack of contextual clues due to the vocal interface, paired with the use of tempting language (e.g. 'Producer's choice') makes VA users vulnerable to Price Comparison Prevention. The fact that a user can complete a purchase without even opening the website of an online store, based exclusively on what information the VA recites, makes Price Comparison Prevention particularly easy to implement. Price Comparison Prevention, in combination with vocal interference, can lead users to purchasing goods or services at a price that, without the interference of deceptive design, they would have not paid.

Finally, **Roach Motel** indicates those services for which registering or subscribing is very easy, but de-registering or unsubscribing requires several, complicated steps. It is used to deter users from abandoning a service.[49] VAs can strategically deflect the requests of users to unsubscribe from a service. If asked, the VA replies that to unsubscribe from some third-party services it is necessary to open an app on the smartphone, or go to a website.[50] This might appear like a necessity dictated by the vocal interface limitations, but since users can subscribe to services using only a voice command, there seems to be no technological reason why they cannot unsubscribe using only voice too. The fact that VAs allow users to register or subscribe to services (offered by the producers themselves or by third parties) with the sole voice interaction, but then redirect users to a website or smartphone app to cancel the registration or subscription, seems a clear example of Roach Motel, most likely hidden behind the excuse of the limitations of the vocal interface.

The table below offers a roundup of the VA deceptive designs, mapped against the techniques identified in the popular taxonomies (Table 1).

[48] Mathur and others (n 21).

[49] ibid.

[50] Owens and others (n 8).

3. Legal analysis of VA deceptive design

This section is dedicated to mapping the relevant legal provisions, analysing if and to what extent they are applicable to VA deceptive design. Recently, the consumer protection framework, in the form of the Unfair Commercial Practice Directive (UCPD) and at the Consumer Rights Directive (CRD)[51] has been indicated as effective to regulate some common deceptive designs of online websites. This position has been embraced by the European Commission, and has been supported by national consumer protection authorities.[52] At the same time, the European Data Protection Board has also focused on the relationship between deceptive design and the General Data Protection Regulation (GDPR) .[53] The application and effectiveness of these two regimes in regulating deceptive designs are still under discussion, and the features of VAs raise additional doubts about their adequacy. For the sake of completeness, the Digital Services Act (DSA) and the proposals for Data Act and for AI Act are also analysed, since they contain provisions expressly regulating certain forms of deceptive interfaces and 'manipulative' AI, respectively.[54]

3.1. *Deceptive design and consumer protection*

It has been explained that deceptive design in VAs operates often in combination with extensive profiling, and this combination is an important factor to be considered when assessing the potential influence that VAs exercise on users. Profiling is the result of commercial surveillance, where the digital footprint of individuals is analysed to determine preferences, behaviours, and how prone individuals are to purchase something.[55] Since the early 2000s, the convergence of marketing with profiling and targeted advertising has been often indicated as a form of manipulation of consumers, aiming at finding how vulnerable individuals are and to what, to leverage such vulnerabilities to apply pressure, and hide information or rival goods, reducing choice.[56] In the digital market, these circumstances concurred to exacerbate the asymmetry of information and the power imbalance between producers and consumers.[57] To remedy this, European consumer protection authorities started working on adjusting the consumer law *aquis*, enforcing it on forms of manipulative or unfair commercial practices online.[58] Consequently, even before Brignull coined the term dark patterns, some deceptive design practices were already censored by national consumer protection authorities, and indicated as unfair commercial practices (in accordance with the terminology of consumer protection law).[59]

With the recent diffusion of the terms deceptive design and dark patterns, there is increased awareness on the significance, scale, and features of these design practices (and of their interaction with other practices, such as profiling). The current efforts to update and reform the European consumer law *acquis* are, therefore, a good starting point for this legal analysis. According to the 2021 European Commission's Notice on the interpretation and application of the UCPD: "**The principle-based provisions and prohibitions in the UCPD can be used to address unfair data-driven business-to-consumer commercial practices** in addition to other instruments in the EU legal framework, such as the ePrivacy Directive, the GDPR or sector-specific legislation applicable to online platforms"[60] (emphasis in original).

The UCPD applies to unfair 'business-to-consumer commercial practices', i.e. actions, omissions or other conducts, representations, advertising, marketing communications, done by a trader before, during, or after the sale or supply of a product or service to a consumer (UCPD art. 2(d) and 3). The UCPD combines a general prohibition with rules to de-

[51] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive'), OJ L 149, 11.6.2005 [consolidated version of 28 May 2022], p. 22–39.

[52] BEUC, '"DARK PATTERNS" AND THE EU CONSUMER LAW ACQUIS: Recommendations for Better Enforcement and Reform' (2022) BEUC-X-2022-013; Forbrukerrådet, 'DECEIVED BY DESIGN: How Tech Companies Use Dark Patterns to Discourage Us from Exercising Our Rights to Privacy' (2018).

[53] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, p. 1–88.

[54] Respectively: Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (PE/30/2022/REV/1), OJ L 277, 27.10.2022, p. 1–102; Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), COM/2022/68 final; Proposal for a Regulation of the European Parliament and of the Council on Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, COM/2021/206 final.

[55] Shoshana Zuboff and Karin Schwandt, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power (Profile Books 2018).

[56] Kayleen Manwaring, 'Will Emerging Technologies Outpace Consumer Protection Law? The Case of Digital Consumer Manipulation' (2018) Competition and Consumer Law Journal 26(2).

[57] Ryan Calo, 'Digital Market Manipulation' (2013) 82 George Washington Law Review 995.

[58] Natali Helberger et al., 'EU Consumer Protection 2.0. Structural asymmetries in digital consumer markets' (2021). BEUC [BEUC-X-2021-018].

[59] For example, in 2013 the Netherlands Authority for Consumer and Market fined the low-cost commercial aviation company Ryanair for automatically adding a travel insurance to the purchase of plane ticket, hiding the 'Do Not Insure Me' option in a long list of countries of origin so that users could not easily remove the insurance from their carts. See the 2020 Guidelines of the Netherlands Authority for Consumer and Market 'Guidelines on the Protection of the online consumer Boundaries of online persuasion'.

[60] Notices from European Union Institutions, Bodies, Offices And Agencies, Commission Notice, Guidance on the interpretation and application of Directive 2005/29/EC of the European Parliament and of the Council concerning unfair business-to-consumer commercial practices in the internal market (2021/C 526/01). p. 99.

termine which commercial practices are unfair. This core of provisions is complemented by a list of practices that are always prohibited, in Annex I.

Article 5(1) establishes the general prohibition of unfair commercial practices. A commercial practice is unfair if it: i) goes against professional diligence; or ii) "materially distorts or is likely to materially distort the economic behaviour with regard to the product of the average consumer whom it reaches or to whom it is addressed, or of the average member of the group when a commercial practice is directed to a particular group of consumers" (UCPD, art. 5(2)(b)). The UCPD also clarifies that materially distorting the consumer's economic behaviour means impairing the ability "to make an informed decision, thereby causing the consumer to take a transactional decision that he would not have taken otherwise" (UCPD art. 2(e)).

The 2021 Commission's Notice answers positively to the question about whether the UCPD applies to deceptive design. In general, a dark pattern falls within the UCPD if it materially distorts (or is likely to distort) the economic behaviour of an average or vulnerable consumer, or it could be considered to breach the professional diligence, be a misleading or aggressive practice, based on its specific characteristics and on contextual circumstances (including the average or vulnerable consumer benchmark).[61]

The two main elements to consider for the application of the UCPD to deceptive design are the threshold of professional diligence and the benchmarks of the average and vulnerable consumers. **Professional diligence** is defined as the standards, skills, and care reasonably expected from a trader, in line with the honest market practices and general good faith (UCPD, art. 2(h)). Professional diligence amounts to normative values adopted in a specific industry or sector.[62] Recently, there have been calls to action within the design world against deceptive design, and more attention is being paid to ethical design principles.[63] Broader terms such as 'professional diligence' are in the process of being defined with regard to design (where relevant), as shown by the reference, made by the Commission, to the very recent principles of ethical design.[64] For the application of the UCPD, however, what matters is the *professional diligence of the trader*, not of the designers that made the trader's website. If an online website or platform is acting as a trader (i.e. offering goods and services for sale to consumers), professional diligence should be interpreted as "the standard of special skill and care which *a trader may reasonably be expected to exercise towards consumers*, commensurate with honest market practice and/or the general principle of good faith in the trader's field of activity."[65] (emphasis added). If the sector standards do not include ethical design values, it can be argued that specific deceptive designs go against professional diligence (honesty and good faith). Nevertheless, the trend seems to be that the principles of ethical design are being incorporated as part of the professional diligence of a trader (operating online). Currently, the ethical design initiatives present some shortcomings. They have not graduated yet into official sector guidelines, which creates uncertainty. If it becomes generally acknowledged in the industry that the standards of UI and UX design include ethical design principles, a specific deceptive design technique might breach professional diligence under article 5 UCPD.[66] This evaluation should be done on a case-by-case basis, or at least within a certain industry or sector.[67] This also implies a lack of harmonization amongst sectors or industries. Furthermore, in many sectors there are multiple initiatives promoting ethical guidelines, with vaguely formulated principles, such as "A designer accepts a professional obligation to further the social and aesthetic standards of the community. (…) A designer shall act in keeping with the honour and dignity of the profession" .[68] Deceptive design is a tool used to hide a clash of interests between the provider of a product or service and the users. Such clash can hardly be solved with general principles and vague formulations, whose interpretation is left to the designers or providers of products, especially in the light of the fact that the features that enable VA deceptive design are the same that also enable the full functionality of many digital products. Sector-specific codes of conduct are necessary to bring clarity, harmonise values and standards, and support authorities in the identification of the duties of the operators within a sector.[69] A concrete attempt in this sense has been done by the Netherlands Authority for Consumers and Markets (ACM), that in March 2023 published guidelines for the companies to protect consumers online. The Guidelines identify in a comprehensible and simplified language what rules must be respected when offering products for sale online, what information are provided to consumers, and how.[70] The guidelines mention VAs once, together with other products and services, to remind traders that the rules of consumer protection apply in those cases too.

The **average and vulnerable consumer** benchmarks move from the assumption that, unless certain conditions interfere with their decision-making capability, an individual can assess the information received and make a beneficial, rational, economic decision. The average consumer test has been used by national courts and the CJEU to assess the misleading or aggressive nature of business-to-consumer practices.

[61] BEUC (n 45).

[62] Hans-W Micklitz, Norbert Reich and Peter Rott, Understanding EU Consumer Law (Intersentia 2009); BEUC (n 45).

[63] Narayanan and others (n 9).

[64] Notices from European Union Institutions, 2021/C 526/01, p. 101.

[65] Notices from European Union Institutions, 2021/C 526/01.

[66] idem

[67] Leiser, Mark, 'Dark Patterns: The Case for Regulatory Pluralism between the European Unions Consumer and Data Protection Regimes' in Eleni Kosta, Ronald Leenes and Irene Kamara (eds), Research Handbook on EU Data Protection Law (Edward Elgar Publishing 2022) https://www.elgaronline.com/view/edcoll/9781800371675/9781800371675.00019.xml> accessed 4 April 2023.

[68] International Council of Design, 'Best practice paper: model code of professional conduct for designers', amended in 2011.

[69] MR Leiser and Mireille M. Caruana, 'Dark Patterns: Light to Be Found in Europe's Consumer Protection Regime' [2021] Journal of European Consumer and Market Law 237.

[70] Available at <https://www.acm.nl/en/publications/information-for-companies/acm-guideline/guidelines-protection-online-consumer> accessed on 4 April 2023.

The CJEU defines the average consumer as a "reasonably well-informed and reasonably observant and circumspect"[71] consumer, or one that "is reasonably well-informed and reasonably observant and circumspect, taking into account social, cultural and linguistic factors."[72] If the business targets specific groups (for example commercial offers concerning sewing machines are targeted at tailors and sewists) the average consumer benchmark is assessed based on that group. The approach of the Court has been critiqued, based as it is on the idealization of an individual that has read the labels, inspected a good, done quite a bit of 'due diligence', and somehow even expects that, to a certain extent, the commercial offer might try to influence them.[73] Additionally, the CJEU has incorporated the average consumer test in different ways, sometimes focusing on the 'abstract' interpretation of average consumers, as an ideal benchmark, discussing for instance how technologically savvy they can be expected to be, or the asymmetry of information existing between them and the business. Other times, the interpretation of average consumer has focused on a specific context and on the exact circumstances in which a consumer would operate.[74]

There are conditions that make the consumer less able to assess the information and make a rational choice.[75] Such conditions make the individual a vulnerable consumer, and the UCPD enumerates them: "mental or physical infirmity, age or credulity" (UCPD, art. 5(3)). The business practices of the trader, in some cases, must be assessed not based on the (foreseeable) effect they would have on an average consumer, but on a consumer belonging to a vulnerable group.

With the advent of data-driven profiling, the differentiation of consumers into average and vulnerable has been the object of criticism. This approach shows its limitations in a context in which consumers are individually and collectively profiled with the precise intent to detect vulnerabilities that make them more prone to purchase a product or engage with an online service.[76] Some scholars have even proposed to consider every consumer a vulnerable consumer.[77] The European Commission seems to have adopted a somehow in-between position, with regard to deceptive design. The 2021 Commission Notice explains that the vulnerable consumer benchmark is to be interpreted as dynamic and situational.[78] Consequently, a business-to-consumer practice might be assessed for its potential unfairness from the perspective of an individual consumer that might have been rendered vulnerable to take certain transactional decisions by the circumstances existing in a specific context (while the same consumer might not be vulnerable in other contexts).

The Commission aims for a case-by-case analysis based on how specific deceptive design techniques affect an individual consumer or consumers in general. These benchmarks remain, however, difficult to apply to personalised services, because profiling can narrow a target group up to the point of targeting individuals,[79] or can render any consumer a vulnerable one. And how should courts or national authorities concretely assess the likelihood of a dark pattern to materially distort the consumer's economic behaviour, when even behavioural sciences and psychology are not completely sure of how deceptive design influences individuals? The argument can be made that the vocal interface of VAs reduces the decision-making capability even of an average consumer, due to its volatility, linearity, and to the lack of context. This is particularly true in the case of the prompts and replies given by VAs, also due to the asymmetry of information and lack of technological literacy of users. While these questions are not necessarily new, as shown by the aforementioned early literature on consumer manipulation, the diffusion of IoT makes them more important than ever, especially if we want to maintain a harmonised consumer protection throughout the Union.

The concept of **transactional decision** is also interesting with regard to deceptive design, because its scope is particularly wide. Transactional decisions, in fact, are not only the entering into a contract or purchasing a product, but include the decision to retain or dispose of a product, or exercise a contractual right. The concept is relational in nature, because identifying the transactional decision taken by the consumer requires a fairness assessment based on the business-to-consumer relationship, and circumstances existing at the moment in which the decision was taken.[80] Accordingly, the decision to continue using a service, for example by engaging with it, opening a website, scrolling through a feed, clicking on a picture or advertising, is a transactional decision.[81] Impairing the consumer's choice so that the consumer keeps scrolling the social media feed, or watching videos on a platform, could be considered an unfair commercial practice if it is the result of a material distortion of the average or vulnerable consumer's economic behaviour. The question is whether talking more, or more often, with Google Assistant or Alexa

[71] CJEU, Case C-210/96 'Gut Springenheide and Tusky v Oberkreisdirektor des Kreises Steinfurt', 16 July 1998.

[72] CJEU: Joined Cases C-54/17 and C-55/17, Autorità Garante della Concorrenza e del Mercato v Wind Tre SpA and Vodafone Italia SpA, 13 September 2018 [ECLI:EU:C:2018:710]; Case C-310/15, Vincent Deroo-Blanquart v Sony Europe Limited, 7 September 2016 [ECLI:EU:C:2016:633]; CJEU, Case C-632/16, Dyson Ltd and Dyson BV v BSH Home Appliances NV, 25 July 2018 [ECLI:EU:C:2018:599]; Case C-484/08, Caja de Ahorros y Monte de Piedad de Madrid v Asociación de Usuarios de Servicios Bancarios (Ausbanc), 29 October 2009 [ECLI:EU:C:2009:682]; Case C-611/1426, Canal Digital Danmark A/S, 26 October 2016 [ECLI:EU:C:2016:800]; Case C-122/10, Konsumentombudsmannen v Ving Sverige AB, 12 May 2011 [ECLI:EU:C:2011:299]; Case C–435/1119, CHS Tour Services GmbH v Team4 Travel GmbH, September 2013 [ECLI:EU:C:2013:574].

[73] Vanessa Mak, 'De gemiddelde consument: Van fictie naar feit'. Ars Aequi. 2017;(7), pp. 592-599.

[74] Hanna Schebesta and Kai P. Purnhagen, 'An average consumer concept of bits and pieces: Empirical evidence on the Court of Justice of the European Union's concept of the average consumer in the UCPD', Wageningen Working Papers in Law and Governance 2019/02.

[75] CJEU, Wind Tre (n. 57).

[76] N Helberger and others, 'Choice Architectures in the Digital Economy: Towards a New Understanding of Digital Vulnerability' (2022) 45 Journal of Consumer Policy 175.

[77] Federico Galli, Algorithmic Marketing and EU Law on Unfair Commercial Practices, Springer, 2022.

[78] Notices from European Union Institutions, 2021/C 526/01.

[79] Mak (n. 58).

[80] Willett, C. Fairness and Consumer Decision Making under the Unfair Commercial Practices Directive. J Consum Policy 33, 247–273 (2010).

[81] Notices from European Union Institutions, 2021/C 526/01.

also constitutes a transactional decision. If it does, even the prompts that tell users to ask the VA for a joke, a challenge, or to tell the news, should be analysed in terms of compatibility with the UCPD. The fact that some deceptive design techniques don't lead to a purchase does not mean that they cannot be monetised, directly or indirectly, immediately or at a later time, by the VA producer. For example, the prompts suggesting users to try other functions ('You can also use me to…' or 'Try asking…') do not generate immediate revenue, but multiply the occasions for users to share usage data, that can be used to refine the profiles of users for targeted advertising, and that can also be sold to data brokers and advertisers. As such, it seems reasonable to consider the user engagement deriving from them as transactional decisions. Being prompted to talking more to a VA is no different than being prompted to scrolling the feeds of a social media or clicking on other user's posts, both already indicated as transactional decisions by the Commission: it remains to be determined past which threshold they are the result of a material distortion of the economic behaviour of a consumer, and whether the average or vulnerable benchmark applies.

3.1.1. *Deceptive design as misleading or aggressive practices*

The UCPD further specifies two categories of practices that are always unfair (and therefore always contrary to professional diligence): misleading and aggressive practices.[82]

Misleading practices consist of providing false information or, through presentation and any other element of the practice, deceive the consumer about important elements of the overall business-to-consumer relationship, so that the consumer enters into a transactional decision that, otherwise, they would have not carried out (UCPD, art. 6). The elements about which the consumer might be deceived include, among others, the existence or nature of the product, the direct or indirect sponsorship of the product, the trader's commitment, motives, and the nature of the transaction, the price and how to calculate it, the consumer's rights. A misleading action can also be an omission that "in its factual context, taking account of all its features and circumstances and the limitations of the communication medium" (such as limitations of space or time, UCPD, art. 7) leaves out important information, necessary to the average consumer to take a decision and, as a result, the consumer takes a transactional decision that they would have not taken under different circumstances. Misleading omissions include cases in which the trader "hides or provides in an unclear, unintelligible, ambiguous or untimely manner such material information (…) or fails to identify the commercial intent of the commercial practice if not already apparent from the context" (UCPD, art. 7(2)).

In particular with regard to an *invitation to purchase* (i.e. a commercial communication that enables the consumer to make a purchase, indicating the product and price in a way that is typical of a commercial communication), omitting certain material information also results in a misleading practice, if the information cannot be deduced from the context. Such information includes, among others, the characteristics of the product, its price, the identity and motives of the trader, the right of withdrawal (where applicable) (UCPD, art.7(4)).

When prompting users to purchase a good, VAs become the medium for an invitation to purchase. They should clearly indicate if the product is a sponsorship, if the price is the result of personalization, and should refrain from framing the product in a way that induces pressure buying or that makes users engage in transactional decisions (e.g. purchase the food processor) that, in other circumstances, they would not take. The fact that this information might be available on the webshop or on the connected smartphone app is not enough: the information must be provided in a timely manner under art. 7(2) UCPD, also via email, and *before the purchase is complete*; making it available only if the user picks up another device and opens a website, does not seem to satisfy this requirement.

The UCPD obligation to provide information overlaps with the provisions of CRD detailing the information and rights to which consumers are entitled in the case of distance contracts. According to art. 6(1) CRD, *before the consumer is bound* by a contract, certain information must be provided, among which there are: the main characteristics of the goods, the identity of the trader, whether the price was personalised using automated decision-making, the right of withdrawal of the consumer. When the distance contract is concluded via an online marketplace, additional information should be provided by the marketplace, including general information on the main parameters determining how the offers are ranked when presented to the consumer and the relative importance of those parameters as opposed to others (positioned in a specific section of the online interface, directly and easily accessible from the page where the products are displayed), and how the online marketplace and the trader selling them share the obligations deriving from the contract (CRD, art. 6a).

The aforementioned examples of VA Disguised Advertising might be considered misleading under art. 6 because they deceive (or are likely to deceive) the consumer about the sponsorship of the product, if it has been sponsored by a business partner of Amazon.

If the information is not false but incomplete, the VA reply might be a misleading omission. Here is where the vocal interface plays an important role. Articles 7(1) UCPD affirms that in evaluating whether important information has been omitted, the intrinsic limitations of the medium used to deliver the commercial communication must be taken into consideration. The medium could be a computer screen, in which case the omission of the important information is unjustifiable, but it might also be only voice. Users can complete a purchase with Alexa or Google Assistant entirely via vocal interface, without seeing a web page. The limitations of space and time of vocal interface seem to be acknowledged by the Modernisation Directive (albeit only with regard of the possibility not to provide consumer with the model form to exercise the right of withdrawal).[83] Recital (41) expressly mentions among the means of distance communication with limited space or

[82] Aggressive practices are less relevant fort his article, since they use threats, coercion, physical force, or exploit a position of power of the trader of the consumer (undue influence), to limit the choice or alter the conduct of the consumer (UCPD, art. 8).

[83] Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the bet-

time "voice operated shopping assistants" (a.k.a. VAs). Consistently, the Commission affirms that, based also on the existing case-law of the CJEU, "'the means of distance communication which allows limited space or time' referred to in Article 8(4) are those that do not allow for layered provision of information (via, for example, expandable headings or hyperlinks, directing consumers to a more detailed presentation of the relevant information)".[84]

Article 8(4) CRD, for offers made through means limited in space or time, establishes that: "the trader shall provide, on or through that particular means prior to the conclusion of such a contract, at least the pre-contractual information regarding the main characteristics of the goods or services, the identity of the trader, the total price, the right of withdrawal, the duration of the contract and, if the contract is of indeterminate duration, the conditions for terminating the contract", while the complete information shall be provided elsewhere, for example via email "in a way appropriate to the means of distance communication used in plain and intelligible language." (CRD art. 8(1))

This reasonably applies to those VAs embedded into devices without a display, and therefore in that case traders operating via a VA can only display the limited set of information determined by art. 8(4) CRD and should provide additional information via email before the purchase is completed. If the VA is embedded in a device with a display, the limitation should not apply. If Alexa or Google Assistant were to list orally all the information concerning the purchase, it might result in a very long message, and users might not pay attention. The overall user experience and the principles of UI and UX favour seamless, less obstructive interfaces. At the same time, this creates more occasions for applying VA deceptive design techniques. The necessities of consumer protection and the dominant optimization ideology underlying UI/UX are in conflict, in this case, and need to be reconciled. More indications on how to layer the provision of information to consumers via vocal interface would be beneficial for the application of the UCPD and CRD. To eliminate the deceptive component from VAs commercial offers, it is fundamental to ensure to deliver immediately, via vocal interface, the minimum information necessary for the consumer to decide if the purchase is convenient, and specifically: price and the parameters that determined it, identity of the traders, parameters for the ranking or selection of the offer, additional costs, shipping costs and methods, and rights of the consumers. Additional measures would also ensure a layered but effective delivery of information to consumers: for example, before the purchase is completed, the VA could recite a warning message informing the consumer that the offer comes with important information, and how to access them (with vocal interaction too, if possible, using a specific command such as "read the additional information").

The discipline of the UCPD is completed by Annex I, containing a list of 31 practices prohibited *tout court*, such as "Displaying a trust mark, quality mark or equivalent without having obtained the necessary authorisation", or "Promoting a product similar to a product made by a particular manufacturer in such a manner as deliberately to mislead the consumer into believing that the product is made by that same manufacturer when it is not."

The Commission expressly points out that some dark patterns fall within the list of practices that are always prohibited, contained in Annex I to the UCPD: Bait and Switch, specific forms of Nagging, Visual Interference, Disguised Advertising and Pressure Selling and Limited Time offers.[85]

Some of the practices listed in Annex I, however, require the *intent* to materially distort the economic behaviour of consumers *in a specific way*. This is the case of, for instance, Annex 1(18), that reads: "Passing on materially inaccurate information on market conditions or on the possibility of finding the product *with the intention of inducing the consumer to acquire the product at conditions less favourable than normal market conditions*". Characterizing an item as selling fast amounts to Scarcity, a well-known dark pattern. Scarcity, however, might be used to push the consumer to purchase a product, but not necessarily at less favourable conditions. In general, it is well-known that many deceptive designs are the result of A/B testing, to see which interfaces work best, but that does not necessarily imply the specific intent of distorting the economic behaviour of the target users.[86] Would this imply that, even though such technique materially distorts the economic behaviour of a consumer, it does not fall under the list of Annex I? And should it not be automatically prohibited by Annex I, how could it be evaluated in the light of articles 5–9 UCPD? Due to the uncertainties indicated above, the result of this assessment might vary greatly. The risk is that, depending on the interpretation of the aforementioned provisions, some deceptive design techniques might not fall under the scope of the UCPD. To determine whether (or which) deceptive designs are an unfair commercial practice, further guidance is needed, at European Union level.

Consider the hypothetical case of a VA showcasing a product, and not disclosing that it has been sponsored by another company, instead referring to it as a deal or a producer's choice (Disguised Advertising and/or Price Comparison Prevention). The argument could be made that they are unfair business-to-consumer practices, expressly prohibited under Annex I no.5, if the company is reasonably aware that the products are, in reality, not available at that price.[87] Because many online retailers use dynamic or personalised pricing, it could be difficult to argue their awareness concerning the price.

Personalised or dynamic pricing are not prohibited by the UCPD, but trigger the obligation, for the trader, of informing the consumer of the fact that the price was personalised.[88] Doubts might emerge when the price is personalised and also adjusted to the demand (dynamic): in this case, depending on which parameter is more prevalent to determine the price, the obligation to inform the user might or might not be triggered. According to the Commission, in the case of these and other data-driver business-to-consumer practices, the GDPR and e-Privacy Directive play an important role too, although no clar-

ter enforcement and modernisation of Union consumer protection rules, (PE/83/2019/REV/1), OJ L 328, 18.12.2019 p. 7–28.

[84] Notices from European Union Institutions, 2021/C 526/01.

[85] Idem.

[86] Narayanan and others (n 9).

[87] Notices from European Union Institutions, 2021/C 526/01.

[88] Notices from European Union Institutions, 2021/C 526/01.

ifications are offered with regard to the coordination between the two regimes.

3.2. VA deceptive design and data protection: the GDPR

VAs process the personal data of users, such as voice biometrics (also protected as a special category of data)[89] and requests and behaviours inside the home, profiling users for a variety of purposes. These activities fall within the scope of the GDPR.

Profiling is defined in the GDPR as: "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements" (art. 4(4)). For the GDPR profiling is an evaluation of the individual, in the light of a pre-established purpose.[90] The GDPR regulates profiling as any other type of processing of personal data via automated means: all the main provisions apply to it, including the general principles (art. 5), the legal bases for the lawfulness of processing (art. 6), the data subject rights and the obligations of controllers and processors. In addition, if the profiling is solely automated and produces legal or similarly significant effects, the specific discipline of art. 22 also applies. It is safe to assume that the profiling carried out by VA producers rarely passes the threshold of generating legal or similarly significant effects, but it is not excluded that it can happen. With regard to targeted advertising in particular, it has already been noted by the Article 29 Working Party that the assessment of the application of art. 22 should be made based on factors such as the intrusiveness of the profiling process (for example whether the data subject has been tracked across different websites, devices and services), the way the advert is delivered, and the fact that it is based on vulnerabilities of the targeted data subjects. Additional factors determining the application of art. 22 would be the fact that differential pricing based on profiling or automated decision results in prohibitively high prices (excluding certain individuals from purchasing something), or that some adults might be in a vulnerable condition or belong to a vulnerable group. These circumstances can become relevant: the combination of VA profiling and deceptive design can be evidence to determine whether the profiling or automated decision has had a legal effect or a persistent and severe impact on the data subject's rights and interests (on a case-by-case basis).[91]

When assessing the compliance of processing activities with the GDPR, the first step is to establish whether they abide to the fundamental principles enumerated by art. 5 GDPR, among which is the principle of fairness: "Personal data shall be: (a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency')". The rest of this section will focus on the principle of fairness and its implementation in relation to deceptive design in VAs. This choice is due to the fact that the very existence of deceptive design implies an imbalance of power between the producers of VAs and the users, which is arguably incompatible with fairness. This approach is also confirmed in the Guidelines concerning deceptive design in social media platforms, where the EDPB expressly connects the principle of fairness and the provision concerning data protection by design and by default with the aim to – among others – remedy power imbalances.[92] Since the Guidelines, however, only discuss fairness and deceptive design very briefly, this section offers a more in-depth perspective.

Fairness plays an important role in discussing the general compatibility of deceptive design with the GDPR. Fairness means that "personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading to the data subject."[93] The EDPB affirms that fairness is the benchmark to distinguish a deceptive interface from one that is neutral or even user empowering.[94] In the view of the EDPB, even if dark patterns can comply with the other GDPR principles, they are generally incompatible with fairness, as this latter is an overarching principle. The position of the EDPB refers specifically to those dark patterns used at interface level by social media platforms, to steer users into consenting to the processing, but it offers insights on the role of fairness also in relation to deceptive design of VAs.

In the GDPR, the principle of fairness plays a double role: procedural, and substantive.[95] Procedural fairness, in particular in combination with the principles of transparency and lawfulness, implies that the controllers comply with a series of requirements, measures, and duties, in a manner that ensures the general principles governing processing are implemented effectively. For instance, procedural fairness manifests jointly with transparency in the information requirements established by article 13 and 14 GDPR. Fairness as a procedural benchmark is what the EDPB refers to when it affirms that the principle can be used to assess the existence of a dark pattern in interfaces. It plays a role in clarifying that some cookie-consent banners and privacy policies might appear formally compliant with the principles of transparency and lawfulness but, if designed using deceptive techniques, they are unfair.

In its substantive role, fairness is used to perform a balancing exercise between potentially conflicting interests: the interests of the controllers must be balanced against the interests and rights of the data subjects, their expectations, and

[89] De Conca (n 20).

[90] Isak Mendoza and Lee A Bygrave, 'The Right Not to Be Subject to Automated Decisions Based on Profiling' in Tatiani Synodinou and others (eds), EU Internet Law: Regulation and Enforcement (Springer 2017).

[91] Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0, Adopted on 20 October 2020, p. 22.

[92] EDPB, Guidelines 3/2022.

[93] EDPB, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, 13 November 2019, para. 64.

[94] EDPB, Guidelines 3/2022.

[95] Gianclaudio Malgieri, 'The Concept of Fairness in the GDPR: A Linguistic and Contextual Interpretation', Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (ACM 2020) https://dl.acm.org/doi/10.1145/3351095.3372868> accessed 4 April 2023.

the possible consequences suffered by them.[96] Similarly to the consumer protection regime, the GDPR presupposes an asymmetric relationship between the data subjects and the controllers, and fairness is the tool to mitigate it. Its roots lie in the Roman institute of *bona fide*, good faith, that presupposes that the parties of a contract will act to pursue their own interests but taking into consideration – and protecting – also the interests of the counterparties.[97]

The substantive role of fairness, revolving around the mitigation of power asymmetries, offers a framework for the assessment of deceptive design techniques of VAs. Deceptive design in VAs enhances the asymmetry of power between data subjects and controllers in three ways:

- it conditions users into engaging more with the VA. The prompts and replies create new occasions for data collection, because at every engagement the data subjects share more personal data;
- it is based on the granular profiling of data subjects, a profiling often focused on vulnerabilities;
- it makes users addicted, dependable and influenceable, often leveraging said vulnerabilities, and this in turn reinforces the two factors listed above. It traps users in a loop of passively sharing their data while creating better ways to make them share more.[98]

This type of profiling appears inherently incompatible with the principle of fairness as the benchmark for substantive 'correctness' in the relationship between controllers and data subjects. This approach aligns with the position of the EDPB that lists the power imbalance between data subjects and controllers as one of the key elements that should be taken into consideration when applying data protection by design: "Power balance should be a key objective of the controller-data subject relationship. Power imbalances should be avoided. When this is not possible, they should be recognised and accounted for with suitable countermeasures."[99]

What is, however, the extent of substantive fairness? The GDPR mentions (or hints at) the fair balance only in a few provisions (e.g. Article 6, Recital 47). Besides those, fairness intended as a fair balancing of rights and interests should be incorporated in the processing, in line with article 25(1) of the GDPR, concerning data protection by design. Under this provision, controllers must implement at design level the principles of the GDPR – *including the principle of fairness* – in an *effective manner*, in accordance with its requirements, to protect the rights of the data subjects. Data protection by design applies at every stage of the processing, from the determination of the means to when the processing itself takes place. The technical and organizational means to implement the GDPR in the design can be evaluated by the controllers based on the costs, state of the art, but also based on the "nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing". Fair design and implementation of the GDPR in the processing, through all its stages, intended as the respect of the data subject's rights and interests according to good faith, is an obligation of the controllers.

In the case of VAs, the controllers have designed a loop, in which the processing of personal data is used to profile data subjects to create replies that make individuals purchase goods or engage more with the device when, in the absence of such prompts, they probably would not. It can be argued that even the fact that deceptive design in VAs creates the occasion for additional data collection, prompting users to engage more with the VA, is not compatible with the principle of fairness. The fact that fairness applies to the way in which personal data are collected was expressly established by the 1980 OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data. There, it is affirmed that personal data "should be obtained by lawful and fair means" (OECD Guidelines, art. 7). In the early days of data protection, the focus was specifically on a fair collection of data, and only subsequently, the fairness principle was expressly extended to the entirety of the processing.[100] Creating occasions for harvesting personal data via deceptive design is not a 'fair means' of obtaining data.

VAs deceptive design should not be compatible with the principle of fairness intended as a substantive fair balancing of conflicting rights. The principle of data protection by design (art. 25(1) GDPR) imposes VA producers to implement fairness and the other GDPR principles at every level of the VA system architecture. While this approach would provide extensive protection to data subjects, there is a risk of over-inclusiveness whereby anything a VA suggests or says to a user offers an occasion for (potentially unfair) data collection. Many of the messages that VAs recite to users – especially those meant to make the users discover features or use them more – would be considered against the GDPR. In other words, Alexa and Assistant might remain silent very often. This could restrict the ways in which VAs engage with users because, as explained, the preconditions for deception and those for normal functionality are the same, particularly in VAs, where the provision of the service depends so deeply on profiling. This extensive and pervasive application of the GDPR would require a radical change in the way VAs are designed, and such changes need time and come at a cost (if possible at all).

Finally, at the time of writing the EDPB and some national Data Protection Authorities have only focused on deceptive design techniques operating at interface level, specifically in the context of UI of websites and social media platforms. There is a clear necessity for additional guidelines concerning the application of the GDPR (and principle of fairness in particular) to deceptive design in general, or at least in a wider variety of online products and services, and of interfaces (such as the vocal one).[101] The EDPB must clarify whether fairness

[96] Damian Clifford and Jef Ausloos, 'Data Protection and the Role of Fairness' (2018) 37 Yearbook of European Law 130; Malgieri (n 80).
[97] Malgieri (n 80).
[98] De Conca (n 20).
[99] EDPB, Guidelines 3/2022, p. 10.
[100] With the 1981 Council of Europe Convention for the protection of individuals with regard to the processing of personal data, (also called Convention 108+ after its modernisation in 2018).
[101] The Vocal Prompts that begin with "By the way, you can also ask me to…" have been consistently identified by users as 'spamm-y' and annoying. Even though it is outside of the scope of this article,

is violated when the user would have not shared those data if not for the deceptive design; furthermore, the appropriate measures to implement data protection by design can be evaluated based on the state of the art which, it might be argued, should also include principles of ethical design. In this regard, it should be pointed out that the data protection and consumer protection regimes intertwine: both make use of the principle of fairness, although the relationship between fairness in consumer protection and fairness in data protection is still being discussed. Consumer protection and data protection can rely on each other regarding, amongst others, the existence of professional diligence. The scholarship has already highlighted the possibility that a breach of the GDPR obligations might amount to lack of professional diligence under the UCPD :[102] a more effective protection of individuals against deceptive design requires the EU authorities to expressly coordinate the GDPR with the consumer protection regime.

3.3. *New rules: DSA, the proposals for data act and AI act*

The increasing interest surrounding deceptive design in general has influenced the preparatory works of three new EU legislative tools, namely the Digital Services Act (DSA)[103] and the draft proposals for Data Act and for AI Act. These laws include provisions directly tackling deceptive design, within the respective scopes of application. These provisions represent a good step towards regulating deceptive design in general, but might be of limited relevance with regard to VAs, for the reasons explained below.

3.3.1. *The DSA*

The DSA regulates the provision of digital services by online intermediaries and platforms, to foster the internal digital market and develop a healthy and safe online environment. Article 25 DSA, significantly titled "Online interface design and organization", is particularly relevant for this analysis. According to it: "1. Providers of online platforms shall not design, organise or operate their online interfaces in a way that deceives, manipulates or otherwise materially distorts or impairs the ability of recipients of their service to make free and informed decisions". Recital (67) clarifies that the provision refers precisely to dark patterns and other deceptive designs at interface level.

Article 25 DSA appears very important for regulating deceptive design in general, but it likely won't apply to VAs. This is due to its limitation in scope, namely the fact that the prohibition is addressed to online platforms. Online platforms are "a hosting service that, at the request of a recipient of the service, stores and disseminates information to the public" (unless the service is ancillary to, or only a minor feature of, another service) (DSA, art. 3(1)(i)). Examples of online platforms are social networks and online marketplaces for the sale of goods or services (Recital 13). A VA, for instance Amazon Alexa, could fall within the definition of online platform, but solely in relation to voice-based online purchases, to posting content on social media platforms via the VA, and other few platform-related activities. The use of deceptive design in other apps or functionalities of the VAs would not be covered by the DSA.

Setting aside the limitation in scope, it must be pointed out that Article 25(2) specifies that the general prohibition of deceptive design does not apply to those cases covered by the UCPD and by the GDPR. The letter of the provision specifically says practices 'covered' not 'prohibited' by the UCPD and GDPR. This raises questions about its practical application. With regard to the UCPD, it is likely that the DSA will complement it, since it regulates the interface of online marketplaces, while the UCPD prohibits unfair and misleading practices of traders that, among others, operate *in* the marketplaces. From the letter of the article, however, it is not clear whether the DSA acts as a 'last resort', a residual clause for those deceptive design techniques that are not tackled by the other laws. If that is not the case, producers and platforms could invoke the application of the UCPD and GDPR, claim that their design practices are compatible with them, and therefore elude regulation completely.[104] On the other hand, the GDPR applies whenever personal data are processed, and VAs process large amounts of personal data for their regular operations. Consequently, the DSA and GDPR significantly overlap. Coordination between the two appears strongly necessary to avoid uncertainty and loopholes. As a final remark, for VAs, besides art. 25, the provision of the DSA regulating the disclosure of how products or information are organized and delivered to users could also apply (recommender systems). This means that Amazon might have to make available to its Alexa users information about how an 'Amazon choice' product is selected and how 'trending topics' are ranked (art. 27).

3.3.2. *The data act*

In February 2022 the European Commission proposed a draft Regulation on harmonised rules on fair access to and use of data, the so-called Data Act. Building on the 2018 Free Flow of Non-Personal Data Regulation,[105] the Data Act aims at "ensuring fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data."[106] Its material scope includes both personal and non-personal data generated using a product or related service, but not data inferred or deduced from usage data. The Data

the opportunity to extend the provisions of the e-Privacy Directive concerning spam and robocalls to VAs should also be investigated by experts, national and European authorities. Similarly, it should also be investigated whether the prompts and replies of VAs amount to advertising, as such regulated by other provisions of the DSA.

[102] Philipp Hacker, 'Manipulation by Algorithms. Exploring the Triangle of Unfair Commercial Practice, Data Protection, and Privacy Law' [2021] European Law Journal eulj.12389.

[103] A similar provision concerning gatekeepers is also inserted in the DMA, which falls outside of the scope of this article.

[104] Mark Leiser, and Cristiana Santos, 'Dark Patterns, Enforcement, and the emerging Digital Design Acquis: Manipulation beneath the Interface' (April 27, 2023). Available at SSRN: https://ssrn.com/abstrac=4431048t.

[105] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (PE/53/2018/REV/1), OJ L 303, 28.11.2018, p. 59–68.

[106] Accompanying memorandum to the Data Act, p. 2.

Act entitles users that generate the data (by interacting with digital products) to access and share/transfer data to other operators, complementarily to the GDPR portability right. The providers and producers who collect and process said data have a corresponding set of duties to ensure they don't abuse their position of data holders, to enable the fair circulation of data among different stakeholders and mitigate the concentration of large datasets within a few big actors in the European digital market. Notably, the Data Act covers data generated by using IoT, and virtual assistants expressly fall within its scope (art. 7(2)).[107]

The Data Act tackles deceptive design too. Article 6 establishes the obligations of the natural or legal persons that are at the receiving end of a data sharing (data recipients, designated by the users). Among the obligations, art. 6(2)(a) prohibits the recipients to "coerce, deceive or manipulate the user in any way, by subverting or impairing the autonomy, decision-making or choices of the user, including by means of a digital interface with the user".[108] This provision expressly prohibits using deceptive design in UI/UX, however its scope is very limited: it only applies to data recipients, when users share or transfer them their IoT usage data. The provision is also formulated in very broad strokes, and does not offer much to clarify which deceptive design techniques would fall under its scope. It is also unclear whether the prohibition means that recipients cannot use the data received to deceive, coerce, or manipulate data users, or if it applies in general.

It should also be pointed out that in a subsequent version of the Data Act (December 2022), an additional provision expressly prohibits data holders to "coerce, deceive or manipulate in any way (…) by subverting or impairing the autonomy, decision-making or choices of the user or the data subject, including by means of a digital interface"[109] to hinder the exercise of the rights to access, use, and share usage data, established by art. 4 of the Data Act. This provision, while narrow in scope, is very important, because it clearly prohibits deceptive design that could bypass the Data Act and prevent users from exercising their rights. While it remains very limited in scope and application, once the Data Act will enter into force such a provision might make a difference for those users who decide to share or transfer data from one VA provider to another. Nevertheless, the Data Act does not seem to offer users any protection against the deceptive design prompts and replies discussed in Section 2, due to its specific scope.

3.3.3. *The AI act*

The AI Act is a proposed Regulation aiming at regulating the placing in the market, putting into service, and use of so-called AI Systems (i.e. machine learning and logic- or knowledge-based software capable of generating predictions, recommendations, or various types of content, AI Act art. 1, 3(1)).[110]

Article 5(1)(a) expressly prohibits AIs that manipulate individuals using "subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's or a group of persons behaviour by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision they would not have taken otherwise in a manner that causes or is likely to cause that person, another person or group of persons significant harm". The same article also prohibits an AI system "that exploits any of the vulnerabilities of a person or a specific group of persons, including characteristics of such individual's or group of persons' known or predicted personality traits or social or economic situation, age, physical or mental ability" that accidentally or intentionally causes – or is reasonably likelty to cause – the person or a group significant harm (AI Act, art. 5(1)(b)).

Regarding VAs, their persuasive capabilities are enhanced and sometimes unlocked by the machine learning behind the vocal interface, and by profiling. However, it is important to clarify what is the role of AI vis-à-vis the manipulation or exploitation: an extensive interpretation would consider the prohibitions of art.5(1)(a)-(b) applicable even if the role of the AI is only ancillary, for example because profiling enables a better or more effective manipulation. A narrow interpretation would only apply those prohibitions when the AI is the primary source of the manipulation or exploitation. This is an important distinction, because in many cases deceptive outcomes can be obtained with very basic techniques, using combinations of words and visual components, and the AI can serve only as a catalyst or to pre-determine vulnerable profiles. Recital (16) seems to indicate a narrow approach, where it specifies that: "The intention to distort the behaviour may not be presumed if the distortion results from factors external to the AI system which are outside of the control of the provider or the user, such as factors that may not be reasonably foreseen and mitigated by the provider or the deployer of the AI system. In any case, it is not necessary for the provider or the user to have the intention to cause the significant harm, as long as such harm results from the manipulative or exploitative AI-enabled practices." It seems that a lot will depend on the assessment of whether it is reasonably foreseeable that a harm might derive from the combination of ai with deceptive design. Finally, the Recital expressly affirms that the prohibition is complementary to the UCPD, and that "lawful commercial practices, for example in the field of advertising, that are in compliance with Union law should not in themselves be regarded as violating prohibition."

Furthermore, the scope of art. 5 is limited by a very high threshold: only those manipulative or exploitative techniques that cause or are likely to cause *significant* harms are prohibited. This seems an unlikely scenario or, at best, a redundant

[107] Art. 2(4) defines VAs as "software that can process demands, tasks or questions including based on audio, written input, gestures or motions, and based on those demands, tasks or questions provides access their own and third party services or control their own and third party devices". This is the first time VAs are officially defined in a law in the EU.

[108] In the subsequent version of 8th December 2022, art. 6 was modified to protect both users and data subjects, when these are not the same person, from deception or manipulation.

[109] Proposal for Data Act, version of 8 December 2022, art. 4(2a), <https://data.consilium.europa.eu/doc/document/ST-15035-2022-INIT/en/pdf> accessed 4 April 2023.

[110] Please note that at the time of writing, the text of the AI Act is still being negotiated in the relevant EU settings. The analysis is based on the draft including the Compromised Amendments as approved by the EU Parliament on 9 May 2023.

precaution, since we can safely affirm that such a product would be illegal in most Member States in any case, and it implies that the prohibition will probably not apply to VAs in practice.

Asides from the limitation of scope, the choice of vague and undefined (undefinable) terms, such as "subliminal techniques" and "a person's consciousness" are not clarified by the article, nor by Recital (16), which only adds a rather confusing reference to brain-computer interfaces, and affirms that: "AI systems deploy subliminal components individuals cannot perceive or exploit vulnerabilities of individuals and specific groups of persons due to their known or predicted personality traits, age, physical or mental incapacities, social or economic situation. They do so with the intention to or the effect of materially distorting the behaviour of a person and in a manner that causes or is likely to cause significant harm to that or another person or groups of persons, including harms that may be accumulated over time". The references to impairing the decision-making of individuals aligns the AI Act and DSA to the UCPD, which is a step in the right direction to make sure these provisions converge. Nevertheless, while Recital (16) helps putting into context the words 'subliminal' and 'consciousness', their concrete interpretation might raise doubts on its judicial applicability and leave room for vastly different interpretations among national courts.

Finally, Art. 5(b) lists the vulnerabilities whose exploitation by AI system is prohibited (if it causes, or is likely to cause, harms): age, disabilities (as defined by Dir (EU) 2019/882), predicted personal or group traits, and social and economic circumstance. This formulation is welcome, because it detaches from the average versus vulnerable consumer benchmarks and acknowledges the role of profiling for vulnerabilities, at least in part (socio-economic circumstances). The examples listed by Recital (16) align with the Article 29 Working Party that, with regard to targeted advertising and the GDPR, already affirmed it is possible, or even necessary, to consider a wide range of vulnerability factors, either at individual or group level, based on the reality of profiling practices.[111] Recital (16) even goes beyond the Working Party, affirming that the harm might result from the accumulation, over time, of multiple manipulations. This is a potentially revolutionary introduction, especially since the GDPR so far was not able to tackle the effects over time of profiling.[112] It will be very interesting to see whether this reference will remain in the final text of the AI Act, and if and how it will be interpreted and enforced.

4. Conclusions

This article explains how deceptive design is used in VAs to influence users into engaging more or purchasing goods, to favour the commercial interests of producers to the detriment of the interests and autonomy of users.

More specifically, VAs use sounds, voice, and visuals to prompt users into using the VA, and administer them commercial offers during a conversation. Many digital products and services are designed to persuade users, but VAs present some peculiarities. The vocal interaction offers some affordances that visual interfaces don't have: VAs can leverage language, tone, and their presence inside the home, to enhance the persuasive capability and build an emotional connection with the users. At the same time, with the vocal interface it is more difficult for users to obtain contextual clues and distinguish advertising from personalised suggestions, and users find themselves forced to wait for the VA to complete a message or listen to a promotional offer before being able to move on.

These prompts and replies match the strategies deployed (and the effects induced) by some famous deceptive design techniques: Nagging, Privacy Zuckering, Misdirection (Vocal Interference), Disguised Advertising, Price Comparison Prevention, or Roach Motel.

The second part of the article assesses the application of the existing consumer and personal data protection EU regimes. The combination of granular profiling, vocal interaction, and placement inside the home raises doubts on several fronts with regard to the application of the Unfair Consumer Practice Directive, Consumer Rights Directive, General Data Protection Regulation, Digital Services Act, and proposals for Data Act and AI Act. This does not mean that these tools do not apply, or that users are not already protected (at least in part). The combination of the analysed tools has the potential to offer a baseline of protection to individuals. What emerges is that:

- more guidelines are necessary to clarify which VA deceptive design go against professional diligence or amounts to a misleading practice under the UCPD;
- the significant use of profiling by VA producers further contributes to undermining the average and vulnerable consumer benchmarks in the consumer protection regime, especially because of the volatility and linearity of vocal interface and because VAs are located inside the home of the users;
- the obligations to provide information within the limitations of space and time dictated by the vocal interface (under the UCPD and CRD) should include also information concerning how an offer is ranked or selected, how the price is determined (e.g. personalized or dynamic pricing). This information must be provided before the purchase is completed, in a layered manner;
- under the GDPR, VA deceptive design appears inherently contrary to the principle of fairness, because it creates unfair means to collect the data (thanks to Privacy Zuckering and prompts to engage users more). This can potentially lead to significant over-inclusiveness, since most VA functions would be incompatible with the GDPR, and therefore requires careful consideration;
- due to the inter-relation between vocal interface and profiling, and due to the business models and operations of the producers of VAs, it is necessary to coordinate the application of the UCPD, CRD, GDPR, DSA, and AI Act, and the respective national and European authorities, to avoid gaps and ensure the protection of EU citizens is not jeopardised;

[111] Article 29 Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (WP 251rev.01)'.
[112] De Conca (n. 14).

- the vocal interface requires additional guidelines and clarity on how to concretely apply the existing (and impending) rules to make sure users of VAs receive the same levels of protection of users of visual interfaces.

Beyond the immediate reality of VAs, this article shows that the vocal interface creates a paradigmatic shift in how users interact with digital products and services, and even with the environment around them. The impact of this shift on the existing legal provisions protecting individuals and their rights remains uncharted territory. This article offers a starting point for further research and reflections, identifying the main frictions between some characteristics of the vocal interface – specifically its volatility, linearity, and lack of contextual cues – and the law. The law often relies on written words (more recently on displays and screens) to make sure that individuals would follow up on their promises, remember their obligations and, not less important, understand the solemnity and importance of a act. Today a purchase made via vocal interface leaves a trace, it can still be recorded in the logs of the device and confirmed with an email. The problem, however, lies in the fact that the vocal interface is based on enhancing the user experience with a seamless, frictionless interaction. The lack of friction takes away solemnity and removes the time to evaluate, assess, and reflect before a decision.[113] This matters, even if the decision is about purchasing a blender.

Vocal interface brings with it an array of additional challenges for the law. Individuals can perceive pressure when they are asked if they want to complete a purchase after just a brief description, or might be induced in confusion due to the lack of contextual cues that occurs when it is always the same voice offering users different options or apps, regardless of whether they are provided by different companies. Voice interface means that before taking a decision users cannot scroll a website, open another offer, compare prices, go back to a previous page, re-read. Existing laws were made keeping in mind written text, paper ledgers, or at least monitors and digital documents. In some cases, this underlying assumption means it is not enough to simply apply existing provisions to vocal interface.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The author is very grateful to the colleagues from the Amsterdam Law & Technology Institute, from the Boundaries of Law research group at VU, Anne-Jel Hoelen, and the anonymous reviewers for their valuable feedback and support.

[113] Arnout Terpstra and others, 'Improving Privacy Choice through Design: How Designing for Reflection Could Support Privacy Self-Management' [2019] First Monday https://journals.uic.edu/ojs/index.php/fm/article/view/9358> accessed 4 April 2023.