Received: 11 September 2020 | Accepted: 28 April 2021

DOI: 10.1111/eulj.12389

VARIETY: ORIGINAL ARTICLE

WILEY

Manipulation by algorithms. Exploring the triangle of unfair commercial practice, data protection, and privacy law

Philipp Hacker*

Abstract

The optimisation of sales practices in consumer markets through machine learning not only harbours the potential to better match consumer preferences with products, but also risks to facilitate the exploitation of consumer weaknesses discovered via data analysis. More specifically, recent technological advances have brought us to the edge of mind-reading technologies, which automatically analyse mental states and adapt offers accordingly, in potentially manipulative ways. This article shows that, in market contexts, the challenges of manipulation by algorithm necessitate an integrated understanding of unfair commercial practice, data protection, and privacy law. It maps the interactions between these contiguous yet distinct fields of law, and draws on economics and computer science to develop a novel framework to deal with algorithmic influence. Furthermore, it critically discusses the Commission proposals for the Digital Services Act and the Artificial Intelligence Act, and suggests to complement them with more broadly applicable measures to mitigate algorithmic manipulation.

1 | INTRODUCTION

For centuries, human beings have observed one another to evaluate how their counterparties think or feel, and to strategically use that knowledge. Recent breakthroughs in machine learning, however, have taken that analytical potential to a new level. At a fast pace, mind-reading technologies are being developed that automatically analyse

* Prof. Dr. Philipp Hacker, LL.M. (Yale), Chair for Law and Ethics of the Digital Society, European University Viadrina, European New School of Digital Studies, Frankfurt (Oder). This paper benefitted from valuable comments by Prof. Hans-W. Micklitz, Prof. Sandra Wachter and audiences at Humboldt University of Berlin, the University of Graz and the European New School of Digital Studies, as well as very helpful suggestions by three anonymous reviewers and the Editor in Chief. Funding information: AXA Research Fund, Grant/Award Number: AXA Postdoctoral Scholarship. Funding information: AXA Research Fund, Grant/Award Number AXA Postdoctoral Scholarship.

mental states, such as emotions, based on images, voice or video recordings.[1] What was once considered the exclusive precinct of science fiction novels is now increasingly deployed in online transactions, brick-and-mortar stores, and even in public spaces.[2] If advertising companies are betting on the potential of such technologies to match consumers with products they are in the mood for,[3] privacy activists are sounding the alarm on what they consider as intrusive surveillance of intimate body functions and mental states.[4] At the very least, mind-reading technologies urge us to reconsider more broadly to what extent psychological traits and weaknesses of counterparties may be revealed and acted upon in digital environments.

Generally speaking, such mental states can be exploited in two ways. First, general offers made indiscriminately to the public can be designed to disproportionately affect consumers which exhibit certain biases or emotions. For example, contracts comprising teaser rates followed by steep price increases could be presented to the general public, with only (or at least mostly) heavily myopic consumers accepting the deal.[5] In these cases, vulnerable consumers *self-select* into the contract because certain psychological traits are present. Such offers therefore constitute a psychologically informed screening device for a trader who is uninformed about which consumers belong to what group.[6] Second, offers can be targeted, that is, made exclusively to specific subgroups of consumers which the trader believes will be particularly receptive because they share certain characteristics.[7] Hence, consumers do not self-select, but are *pre-selected* by the trader.[8] That strategy therefore presupposes knowledge, by the trader, of subgroup membership. Such information is increasingly provided by data collection and algorithmic modelling,[9] and it is therefore this type of targeting that the paper predominantly deals with.[10]

While such technologies arguably raise important challenges in a number of legal fields,[11] this paper particularly considers applications in market contexts, and asks to what extent the specific targeting of cognitive traits or emotional states may lead to a manipulation of consumer decision-making which runs afoul of EU market regulation, including the proposed Digital Services Act (DSA)[12] and the proposed Artificial Intelligence Act (AIA).[13] The application of machine learning to marketing and contracting contexts risks to exacerbate exploitative practices as algorithmic models may—in ways intended or not intended by their users—capture and optimise on expected consumer weaknesses in cognitive or emotional domains. However, as this paper shows, current EU law struggles to address these issues within its existing framework. If anything, the case of mind-reading technologies underscores the need for an integrated market order for the digital economy, in

[1] For an overview, see C. Burr and N. Cristianini, 'Can Machines Read Our Minds?' (2019) 29 *Minds and Machines*, 461.

[2] See A. McStay, *Emotional AI* (Sage, 2018), ch. 8; and below, Section 2.1.

[3] A. McStay, 'Emotional AI, Soft Biometrics and the Surveillance of Emotional Life: An Unusual Consensus on Privacy' (2020) 7 *Big Data & Society* 6–7, https://doi.org/10.1177/2053951720904386.

[4] See, e.g. M. Whittaker et al., *AI Now Report* (2018), ainowinstitute.org/AI_Now_2018_Report.pdf, 4, 8.

[5] H. Shui and L. Ausubel, 'Time Inconsistency in the Credit Card Market' (2004) *14th Annual Utah Winter Finance Conference*, https://eml.berkeley.edu/~webfac/dellavigna/e218_f05/ausubel.pdf, 3–4, 9–10 (all websites last visited on 11 March 2021); see also below, n. 167.

[6] On screening generally J. Stiglitz, 'The theory of "screening", education, and the distribution of income' (1975) 65 *American Economic Review*, 283.

[7] This is often called behavioural targeting, see, e.g. F. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Wolters Kluwer, 2015), 15, 47.

[8] See, e.g. R. Calo, 'Digital Market Manipulation' (2014) 82 *George Washington Law Review*, 995, 1003 et seq., and references in n. 15.

[9] See below, Section 2.2.

[10] Unless otherwise noted, 'targeting' refers to this practice. General offers, in turn, are discussed further in Sections 2.1.2 and 3.2.2.1; see also n. 167.

[11] See e.g. with respect to face recognition technology, European Commission, 'White Paper on AI', COM(2020) 65 final, 21–22; Council of Europe, Guidelines on Facial Recognition, T-PD(2020)03rev4, 2021; E. Kindt, 'Having Yes, Using No? About the New Legal Regime for Biometric Data' (2017) 34.3 *Computer Law & Security Review*, 523; see also European Data Protection Supervisor (EDPS) 'Opinion 3/2018 on online manipulation and personal data' (2018); S. Stolton, 'Commission under Pressure in EU Court over "Lie Detector Tech"', *EURACTIV* (8 Feb. 2021), https://www.euractiv.com/section/digital/news/aommission-under-pressure-over-lie-detector-tech-in-eu-courts/.

[12] European Commission, Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act), COM(2020) 825 final.

[13] European Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), COM(2021) 206 final.

which unfair commercial practice, data protection, and privacy law complement and mutually support one another.[14]

Indeed, manipulation by digital technologies is a growing concern among legal scholars.[15] More recently, the proposals of the DSA and the AIA have also taken up worries about manipulation in digital spaces,[16] and may offer a window of opportunity for policy proposals in this respect. This paper adds to this discourse by undertaking a systematic and in-depth assessment from the perspective of EU market law. A particular focus rests on the Unfair Commercial Practices Directive (UCPD), the key EU law instrument for safeguarding informed and rational market decisions. Even the DSA and the AIA, if eventually enacted, will not detract from the centrality of the UCPD for digital manipulation since, in this context, they primarily complement the UCPD with specific provisions concerning platform transparency (DSA) or AI systems impacting consumers outside of commercial practices (AIA). However, as will be discussed, the DSA and the AIA do fill important gaps, and merit close scrutiny concerning any future framework for the mitigation of algorithmic manipulation.

In this endeavour, the article proceeds in four steps. First, it lays conceptual and interdisciplinary foundations for the treatment of algorithmic manipulation by asking how manipulative practices differ from acceptable economic behaviour (Section 2). In this, it draws on insights from behavioural economics and computer science to further refine the concept of manipulation with respect to consumer weaknesses. Second, it explores the limits of harnessing machine learning for emotional and cognitive targeting under the UCPD (Section 3). Third, it turns to data protection law for an analysis of algorithmic manipulation under the General Data Protection Regulation (Section 4). Throughout, the paper stresses the interactions and interdependencies between the different legal instruments considered. Finally, it discusses the recent anti-manipulation initiatives in the DSA and the AIA and makes three concrete proposals to update EU law in view of the challenges of mind-reading technologies, and algorithmic manipulation more generally (Section 5).

2 | INTERDISCIPLINARY AND CONCEPTUAL FOUNDATIONS

When regulating markets and technologies, and analysing concepts like manipulation, legal scholarship cannot isolate itself from insights of neighbouring disciplines investigating these very terms and phenomena. Conceptually, the delimitation of acceptable persuasion from inacceptable manipulation is hotly debated in moral philosophy and jurisprudence.[17] In these debates, one may distinguish two dimensions of manipulation: manipulation as unawareness of influence,[18] and manipulation as the impossibility of rational choice.[19] Normatively, both dimensions are relevant

[14] For a similarly integrated view of EU market law, see also N. Helberger, F. Zuiderveen Borgesius and A. Reyna, 'The Perfect Match? A Closer Look at the Relationship between EU Consumer Law and Data Protection Law' (2017) 54 *Common Market Law Review*, 1427, 1439–1443; F. Costa-Cabral and O. Lynskey, 'Family Ties: the Intersection between Data Protection and Competition in EU Law' (2017) 54 *Common Market Law Review*, 11; P. Hacker, 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under EU Law' (2018) 55 *Common Market Law Review*, 1143; D. Clifford, I. Graef and P. Valcke, 'Pre-formulated Declarations of Data Subject Consent—Citizen-Consumer Empowerment and the Alignment of Data, Consumer and Competition Law Protections' (2019) 20 *German Law Journal*, 679.

[15] See, e.g. Calo, n. 8; K. Yeung, 'Hypernudge: Big Data as a Mode of Regulation by Design', (2017) 20 *Info. Comm. & Society*, 118, 119; A. Nadler and L. McGuigan, 'An impulse to exploit: The behavioral turn in data driven marketing' (2018) 35 *Critical Stud. Media Comm.*, 151, 161; T. Zarsky, 'Privacy and manipulation in the digital age' (2019) 20 *Theoretical Inquiries in Law*, 157, 158; D. Susser, B. Roessler and H. Nissenbaum, 'Online manipulation: Hidden influences in a digital world' (2019) 4 *Georgetown Law Technology Review*, 1; S. Wachter, 'Affinity profiling and discrimination by association in online behavioural advertising' (2020) 35 *Berkeley Technology Law Journal* (forthcoming); L. Willis, 'Deception by design' (2020) 34 *Harvard Journal of Law and Technology*, 115; N. Helberger et al., EU Consumer Protection 2.0, BEUC Report, 2021; J. Laux et al., 'Neutralizing online behavioural advertising' (2021) 58 *Common Market Law Review* (forthcoming), https://ssrn.com/abstract=3822962; see also, more generally, Zuiderveen Borgesius, n. 7; F. Pasquale, *The Black Box Society* (Harvard University Press, 2015).

[16] See, e.g. Recital 57 DSA; Recitals 15, 70 AIA.

[17] See, e.g. C. Coons and M. Weber, 'Introduction', in C. Coons and M. Weber (eds.), *Manipulation: Theory and Practice* (Oxford University Press, 2014), 1.

[18] Susser, Roessler and Nissenbaum, n. 15, 17.

[19] A. Wood, 'Coercion, Manipulation, Exploitation', in C. Coons and M. Weber (eds.), *Manipulation: Theory and Practice* (Oxford University Press, 2014), 17, 35; R. Noggle, 'Manipulative actions: A conceptual and moral analysis' (1996) 33 *Am. Phil. Q.*, 43, 44; see also Th. Hill, *Autonomy and Self-Respect* (Cambridge University Press, 1991), 33.

because each of them suggests an infringement on autonomy, which requires a sufficiently independent formation of preferences, and the possibility to critically and rationally review these preferences.[20] Therefore, this paper will operate with both dimensions of manipulation, distinguishing them whenever necessary.

The focus of this paper, however, will rest on challenges to rational choice. For manipulation by unawareness, disclosure constitutes a natural antidote, conceptually ruling out the hidden nature of influence.[21] Importantly, though, divulging information about influence does not guarantee rational decision-making by the addressee of that information, as is well known from research in psychology and behavioural economics. Therefore, the first part of this paper turns to such insights (Section 2.1) and then to computer science (Section 2.2) to refine the concept of manipulation in digital market contexts.

2.1 | (Behavioural) economics and psychology: the concept of manipulation revisited

From an economic perspective, unconstrained optimisation of machine learning models on consumer behaviour would be a much smaller problem if all consumers were fully rational actors. In this case, hidden influence could be largely remedied by disclosure, and rational choice would rarely be impossible. As research in psychology and behavioural economics has shown, however, actors often process information in imperfect ways, being subject to various forms of cognitive biases (e.g. optimism bias; status quo bias; present bias).[22] Such effects have been demonstrated to arise frequently if the decision-maker resorts to so-called “System 1” processing, a fast, intuitive and heuristic-based type of cognitive response.[23] While biases can generally be overridden by a more reflective and balanced “System 2” response, such cognitive checks and balances are often not applied in a decision situation, leading to suboptimal decision-making.[24] As behavioural contract theory shows, such consumer weaknesses can even lead to so-called “exploitative contracts” in which consumers not only end up paying more than they were initially willing to,[25] but which may also lower social welfare overall.[26]

2.1.1 | Manipulation and rational choice

While marking a *deviation* from the normative decision-making model of rational choice, boundedly rational behaviour has come to be understood as the *standard* form of decision-making in an empirical sense,[27] at least among non-experts such as most consumers.[28] This has important implications for the concept of manipulation. Even if most actors do not act fully rationally in most contexts, manipulation as the impossibility of rational choice cannot be equated with the mere presence of biases or emotions in decision-making. Otherwise, close to all decisions would count as manipulated, depriving the concept of its signalling and distinguishing function. Conversely, however, demanding strict proof that rational choice is *impossible* raises the bar so high that the criterion would hardly ever be shown to be fulfilled in practice. Therefore, conceptually, it should be considered sufficient that rational

[20] G. Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press, 1988), 18; see also Wood, n. 19, 36.

[21] P. Hansen and A. Jespersen, ‘Nudge and the manipulation of choice’ (2013) 4 *European Journal of Risk Regulation*, 3, 19; P. Hacker, ‘Nudging and autonomy. A philosophical and legal appraisal’, in H.-W. Micklitz, A.-L. Sibony and F. Esposito (eds.), *Research Methods in Consumer Law* (Edward Elgar, 2018), 77, 105–111.

[22] See, e.g. C. Camerer et al. (eds.), *Advances in Behavioral Economics* (Princeton University Press, 2004); E. Zamir and D. Teichman, *Behavioral Law and Economics* (Oxford University Press, 2018), 19 et seq.

[23] See, e.g. D. Kahneman, ‘A perspective on judgment and choice: mapping bounded rationality’ (2003) 58 *American Psychologist*, 697, 698 et seq.

[24] See, for an overview, M. Altman, *Handbook of Contemporary Behavioral Economics: Foundations and Developments* (Routledge, 2015).

[25] See, e.g. B. Kőszegi, ‘Behavioral contract theory’ (2014) 52 *Journal of Economic Literature*, 1075, 1104.

[26] P. Heidhues and B. Kőszegi, ‘Naivete-based discrimination’, (2017) 132 *Quarterly Journal of Economics*, 1019.

[27] K. Stanovich, R. West and M. Toplak. *The Rationality Quotient: Toward a Test of Rational Thinking* (MIT Press, 2016), 222 et seq., 235 et seq.; T. Sharot ‘The optimism bias’ (2011) 21.23 *Current Biology*, R941, R942.

[28] For a nuanced discussion of biases among experts, see P. Brest and L. Krieger, *Problem Solving, Decision Making, and Professional Judgment* (Oxford University Press, 2010), 298 et seq.

decision-making is *significantly impaired*, which implies that System 1 processing very likely does not successfully correct System 2 thinking. Not only does "significant impairment" better align with the standard in unfair commercial practice law;[29] it also embodies a more realistic, factually operationalisable threshold that still restricts manipulation to more severe forms of interference. Hence, it is suggested to reserve the concept of manipulation by impairment of rational choice to instances in which specific elements are present which: (i) make it highly likely that rational decision-making of one party is significantly diminished; and (ii) whose relevance for the decision can be, directly or indirectly, attributed to the intervention of the other party. Without such intervention, one party might commit a unilateral mistake, but manipulation typically involves some plan or action on behalf of the other party who stands to benefit.[30] As we shall presently see (Section 2.2), these forms of intervention change in digital environments.

Concerning the first element of the definition, reduced rational decision-making capacity, it is submitted that the required significant impairment may stem from two different sources: the type of internal state of the actor; or the circumstances of the decision. With respect to the former (type of state), there are two main sub-cases which may be said to sufficiently counteract rational decisions in market contexts. First, the bias or emotion may be of such *pronounced extent* that it is highly improbable to be corrected by System 2 (intensity variety). It is now well-established that biases are not uniform, but occur in sizes small and large.[31] The reduced capacity may therefore result from particularly strong cognitive biases,[32] or from "hot" emotional states making a cool assessment unlikely.[33] For example, strong biases or emotions could make consumers underestimate the risks of products (optimism bias; joy), or overestimate their utility (availability heuristic), leading to an inflated willingness to pay.[34] Traders could exploit this by price-discriminating against these consumers, charging them supra-competitive prices.[35] Second, consumers' decision-making capacities may be significantly impaired if offers play on several, *mutually reinforcing* biases or emotions (combination variety). For example, they might be lured, even by catering to smaller biases, into actions that violate their long-term interests; these, in turn, are often difficult to account for because of yet other cognitive or emotional obstacles, such as present bias.[36]

With respect to the circumstantial form of reduced rational decision-making capacity, the targeting of emotions or biases may be taken as an important indication that rational choice is indeed impaired if the influence is exercised in a *complex or unanticipated* situation (complexity variety): there is empirical evidence that emotions, and biases, tend to override rational decision-making particularly in those contexts.[37] This may be an online or even an offline environment with a flurry of competing offers that are difficult to compare.[38] Note, however, that even in such environments, the threshold to manipulation is arguably only passed if complexity is paired with emotional or cognitive weaknesses, even though they need not be as strong as in the former two cases (intensity/combination variety).

[29] See below, Sections 3.2.2 and 5.2.1.

[30] Coons and Weber, n. 17, 10; generally, even intentionality is required, see, e.g. Hill, n. 19; T. M. Wilkinson, 'Nudging and manipulation' (2013) 61 *Political Studies*, 341, 351; but see also n. 53 and accompanying text.

[31] K. Stanovich and R. West, 'Individual differences in reasoning: Implications for the rationality debate?' (2000) 23.5 *Behavioral and Brain Sciences*, 645; Stanovich, West and Toplak, n. 27.

[32] B. Kőszegi, 'Behavioral contract theory' (2014) 52 *Journal of Economic Literature*, 1075, 1104.

[33] G. Loewenstein, 'Emotions in economic theory and economic behavior' (2000) 90 *American Economic Review*, 426, 428 et seq.

[34] See, e.g. A. Ezrachi and M. Stucke, *Virtual Competition* (Harvard University Press, 2016), ch. 11.

[35] A detailed analysis of the many aspects and types of price discrimination transcends the scope of this paper; see, e.g. F. Zuiderveen Borgesius and J. Poort, 'Online price discrimination and EU data privacy law' (2017) 40 *Journal of Consumer Policy*, 347; F. Zuiderveen Borgesius, 'Price discrimination, algorithmic decision-making, and european non-discrimination law' (2020) 31 *European Business Law Review*, 401; and A. Ezrachi and M. Stucke, above A. Ezrachi and M. Stucke, above n. 34.

[36] See, on present bias, D. Laibson, 'Golden eggs and hyperbolic discounting' (1997) 112 *Quarterly Journal of Economics*, 443.

[37] J. Forgas, 'Mood and judgment: the affect infusion model (AIM)' (1995) 117 *Psychological Bulletin*, 39.

[38] Such complexity can also be consciously engineered into a digital environment by so-called 'dark patterns', see, e.g. C. Bösch et al., 'Tales from the dark side: Privacy dark strategies and privacy dark patterns' (2016) 4 *Proceedings on Privacy Enhancing Technologies*, 237; Forbrukerrådet, *Deceived by Design*, Report, 2018, 32 et seq.; M. Martini et al., 'Dark patterns', *ZfRD* (2021), 47.

2.1.2 | General vs. targeted offers

This observation points to a final and important distinction: in the discussed cases, manipulation generally only occurs in case of targeting, not of general offers made indiscriminately to all consumers.[39] There are two reasons for this. First, conceptually and as mentioned in the introduction, targeting in this sense refers to offers that only reach a subgroup of consumers with certain characteristics.[40] Emotional or cognitive targeting forms a subclass of such targeting: it (implicitly or explicitly) measures existing, individual emotions or cognitive traits by technical means and harnesses them by confronting the consumer with content (including prices) optimised in view of the measurement. If these measured characteristics happen to be emotional or cognitive *weaknesses* of the types discussed above (intensity; combination; complexity variety), the practice will be considered manipulative, provided that the relevance of the bias or emotion for the decision at hand can be attributed to the trader. This, in turn, can generally be assumed only if the weakness was specifically targeted,[41] and not if a general offer was made to all consumers of which some happen to exhibit said weakness.

Second, from a legal perspective investigated in detail below,[42] emotional and cognitive targeting—but not a general offer—typically changes the reference actor from the average to the vulnerable consumer, making a finding of UCPD unfairness much more likely. Together, these observations show that digital environments may indeed change the type and effects of manipulative interventions, an issue to which we now turn in more detail.

2.2 | Computer science: Intentional vs. unintentional targeting

As the preceding discussion has shown, manipulation is conceptually only obtained if the hidden influence or the impairment of rational choice of one party can be, directly or indirectly, attributed to the other party. Over centuries, manipulation has generally involved a deliberate and specific intervention in decisions, for example via hypnosis, intoxication, or bribery,[43] rendering a separate discussion of attribution obsolete. This stands to change with the rise of machine learning.

Simply put, in machine learning, an algorithm calibrates a model given a target, the so-called “objective function”.[44] The learning algorithm tweaks the model, by adjusting its internal parameters, to maximise the target value. In marketing, the objective will typically be profit maximisation.[45] Hence, the machine learning model will scrutinise large amounts of (potentially anonymised) customer data to evaluate which offers work best with what types of consumers.[46] Ideally, the model will correctly predict consumer preferences and suggest offers that closely match these preferences, saving search costs for the consumer and enhancing sales and profits for the trader.[47]

However, not all cases of using AI for sales optimisation will be mutually beneficial. Generally speaking, optimisation without constraints mitigating its effect on consumer weaknesses risks unilaterally favouring the trader, as the following sections will show. Importantly, however, one must distinguish between the *intentional* targeting of manipulable traits on the one hand, and manipulation as a mere *side effect* of unconstrained algorithmic optimisation on the other.

[39] A counterexample to this general rule is given in n. 167.

[40] See n. 7 and accompanying text.

[41] See also Sections 2.2 and 3.2.1.2.

[42] See Section 3.2.2.

[43] See Hill (n. 19); Coons and Weber (n. 17), 9 et seq.

[44] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning* (MIT Press, 2016), 79, 107.

[45] S. Lessmann et al., ‘Targeting customers for profit’ (2019) *Information Sciences*, https://doi.org/10.1016/j.ins.2019.05.027.

[46] I. Witten et al., *Data Mining*, (Morgan Kaufmann, 4th edn. 2016), 27 et seq.

[47] V. Marotta et al., ‘The welfare impact of targeted advertising technologies’ (2021) *Information Systems Research* (forthcoming), https://ssrn.com/abstract=2951322, 4.

2.2.1 | Unintentional targeting

First, a model may learn that certain consumers may exhibit the manipulable traits discussed above (intensity; combination; complexity variety), which may be costly for them but beneficial for the trader.[48] In this way, the use of the model may lead to an exploitation of cognitive or emotional weaknesses simply because such actions maximise the objective function of the given model.[49] Such effects can arise even in models that are applied to a heterogeneous pool of consumers, as models analyse and differentially act upon training data encompassing both vulnerable and more rational consumers.[50] Since algorithmic outputs may depend on very specific feature combinations,[51] the same model may optimise on boundedly rational weaknesses in some consumers and furnish mutually beneficial offers facilitating rational choice to others. In this sense, the exploitation of errors and weaknesses of contractual counterparties may indeed be an unintentional side effect of algorithmic profit maximisation.[52] Nevertheless, with respect to the concept of manipulation, the fact that the trader did deliberately employ an unconstrained model, without "checks and balances" for consumer weaknesses, should in principle be sufficient to attribute even the unintentional targeting of manipulable traits to the trader (see for details below, Section 3.2.1.2).[53]

2.2.2 | Intentional targeting

Second, mind-reading technologies increasingly convey capacities to very intentionally tailor actions to cognitive and emotional states of counterparties.[54] The technical basis of this strategy is what media scholar McStay has dubbed "Emotional AI" (also called: affect recognition).[55] It consists in the detection of mental states with the help of machine learning methods, mostly using deep neural networks.[56] Emotional AI is being used for varying purposes, from road safety (surveillance of drivers) all the way to targeted advertisement.[57] For example, Microsoft has integrated emotion detection into its Azure AI suite, which can be embedded in various apps;[58] the same holds true for Amazon Web Services.[59] While this type of psychological analysis, which builds on basic emotion theory,[60] has been criticised, inter alia, as simplistic and unreliable,[61] the technology has already delivered surprising performance levels in some fields,[62] and will likely improve over time. Similar techniques could be used to detect and harness cognitive biases of potential counterparties.[63]

While the potential of machine learning to exploit cognitive and emotional weaknesses of counterparties can be easily derived from the modelling process, empirically validating that such manipulation does take place in practice is much more difficult because the models tend to be hidden from public scrutiny.[64] However, some descriptions of the implementation of potentially manipulative AI can be found in the literature. Emotion recognition is already being

[48] M. Cronin, 'Digital commerce, AI, and constraining consumer choice', *The Ethical Machine*, Shorenstein Center on Media, Politics, and Public Policy, Harvard Kennedy School (19 March 2020), https://ai.shorensteincenter.org/ideas/2019/4/3/digital-commerce-ai-and-constraining-consumer-choice.

[49] Calo (n. 8); P. Hacker, 'Personal data, exploitative contracts, and algorithmic fairness' (2017) 7 *International Data Privacy Law*, 266, 270–274.

[50] Goodfellow, Bengio and Courville, n. 44.

[51] Id., 3 et seq.

[52] Willis, n. 15, Section I.C.

[53] See also Coons and Weber, n. 17, 13 (requiring not intention but only that 'the influencer has no regard for whether the influence *makes sense* to' the manipulated person).

[54] See n. 1.

[55] See McStay, n. 2, ch. 1; R. Calvo and S. D'Mello, 'Affect detection' (2010) 1(1) *IEEE Transactions on Affective Computing*, 18.

[56] Burr and Cristianini, n. 1, 472.

[57] McStay, n. 2, 115 et seq.; McStay, n. 3, 2.

[58] Microsoft Azure, Face, https://azure.microsoft.com/en-us/services/cognitive-services/face/.

[59] AWS, Emotion, https://docs.aws.amazon.com/rekognition/latest/dg/API_Emotion.html.

[60] See P. Ekman, 'Basic Emotions' in T. Dalgleish and M. Power (eds.), *Handbook of Cognition and Emotion* (Wiley, 1999), 45; D. Keltner et al., 'What basic emotion theory really says for the twenty-first century study of emotion' (2019) 43 *Journal of Nonverbal Behavior*, 195.

[61] Whittaker et al., n. 4, 4, 8; L. Barrett et al., 'Emotional expressions reconsidered' (2019) 20 *Psychological Science in the Public Interest*, 1.

[62] See Burr and Christiani, n. 1.

[63] See Stanovich, West and Toplak, n. 27.

[64] See generally Pasquale, n. 15, ch. 3.

put to use by marketers to improve campaigns online and in brick-and-mortar retail, and is increasingly launched in public spaces as well.[65] For example, the company eyeQ has developed a tool scanning the faces of in-store shoppers to analyse emotions and other parameters in real time, and tailoring in-store marketing on this basis.[66] This will arguably become manipulative if emotions acted upon are particularly strong (intensity variety), reinforced by biases (combination variety), or if the decision-making situation in which the targeting takes place is particularly complex (complexity variety, for example, choice overload and difficult comparisons in a store). Another example of profiling consumers with respect to emotions and consecutively targeting them was reported in a leaked document concerning Facebook Australia.[67] Allegedly, the company offered business customers the possibility to target advertisements toward emotionally insecure and vulnerable adolescents.[68] According to the leak, the company has found a way to pinpoint when their younger users feel “worthless”, “overwhelmed”, “stressed”, and “a failure”.[69] If these emotions were strong enough to significantly impair rational decision-making, which does seem plausible, the targeting was tantamount to manipulation under the intensity variety. Hence, emotional and cognitive targeting is not only a theoretical, but already a very practical challenge for market regulation.[70]

2.3 | Conceptual results

In sum, both manipulation as unawareness and manipulation as impossibility of rational choice need to be refined and contextualised. The former can be countered by disclosure, but this does not guarantee adequate consumer decision-making in practice. Neither does emotional or cognitive targeting of consumers, in the view of this paper, render rational choice impossible *per se*. Rather, it becomes manipulative, in the latter sense, if it significantly impairs rational consumer decision-making—either through the type of weakness targeted (intensity; combination variety) or via circumstances rendering rational decision-making highly unlikely (complexity variety). Moreover, the specific manipulative aspect must be attributable to the targeting entity, for example because vulnerable consumers were intentionally selected, or because the trader used an unconstrained model.

Viewed from this perspective, algorithmic modelling risks not only to facilitate preference matching, but also, in the mentioned cases, to take the manipulation of consumer choice to a new level. It exacerbates information asymmetry between traders and consumers, facilitating the targeting of manipulable traits; it extends such targeting to instances in which the trader may “mean no harm” but employs a model that optimises on consumer weaknesses; and it gives rise to complex architectures which may hamper rational decision-making.

As shown, these elements are increasingly buttressed by empirical evidence. The remainder of the article will therefore explore what resources existing EU law holds, particularly in the Unfair Commercial Practices Directive, the General Data Protection Regulation, and privacy law, but also in the DSA and AIA proposals, to safeguard autonomous choice in the face of algorithmic manipulation. The specific challenge lies in mitigating manipulation without, simultaneously, foreclosing the positive preference-matching effects algorithmic modelling may have in consumer markets.

[65] See A. McStay, ‘Empathic media and advertising’ (2016) 3(2) *Big Data & Society*, Article 2,053,951,716,666,868, 2 et seq.; McStay (n. 2).

[66] McStay (n. 2), 119; see also https://eyeq.tech/core-technologies/#emotion (last accessed on June 20, 2020).

[67] Susser, Roessler and Nissenbaum (n. 15), 6.

[68] S. Levin, ‘Facebook told advertisers it can identify teens feeling “insecure” and “worthless”’, *The Guardian* (May 1, 2017), https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens.

[69] Id.

[70] See Cronin (n. 48); D. Clifford, *The Legal Limits to the Monetisation of Online Emotions* (Ph.D. Thesis, Leuven, 2019), para. 6.

3 | ALGORITHMIC MANIPULATION AND THE UCPD

In market transactions, manipulation is chiefly addressed, under EU law, by the Unfair Commercial Practices Directive (UCPD). Algorithmic manipulation, however, poses a double challenge for this legal instrument. On the one hand, from an internal UCPD perspective, it is unclear where the limits of unfairness lie for emotional and cognitive targeting practices. The following sections therefore draw on recent jurisprudence by the CJEU, and on decisions by consumer agencies, to more sharply delineate the threshold for algorithmic manipulation under the UCPD. On the other hand, from the perspective of EU market regulation more generally, the analysis is complicated by the interactions between the General Data Protection Regulation (GDPR[71]) and the UCPD.

To start with, most of the practices of algorithmic influence mentioned in Part 2 fall within the scope of the UCPD. They occur between a trader and a consumer, and constitute, for the most part, commercial practices pursuant to Article 2(d) UCPD as they are employed to market a specific product. For practices within its scope, the UCPD contains different fairness standards, with three levels of concreteness. First, the black list in Annex I comprises practices considered unfair under all circumstances. However, none of the currently listed activities clearly matches cases of algorithmic manipulation in the sense of ML-based cognitive or emotional manipulation. Second, the specific provisions of Articles 6–9 UCPD guard against misleading actions and omissions, as well as aggressive practices. Finally, the general clause in Article 5 UCPD covers unfair commercial practices more generally.

In general, the prohibition of misleading actions and omissions is capable of addressing concerns about manipulation by unawareness of influence. This paper will, nevertheless, focus on aggressive commercial practices (3.1) and the general clause (3.2), for two reasons. First, the issue of misleading actions by digital means has already been dealt with repeatedly in the literature.[72] Second, and more importantly, disclosure of influence does not rule out manipulation as the impossibility of rational choice. While traders can always deflect the charge of misleading consumers by providing more information, Articles 8–9 and 5 UCPD formulate substantial criteria to rein in algorithmic manipulation.[73] From an empirically-informed regulatory perspective, this seems important as a wealth of studies show that information on data processing is flatly disregarded by consumers in the digital economy.[74] Therefore, enforcement activities in the UCPD context should focus on aggression and the general clause rather than misleading practices. In the following sections, the paper therefore seeks to transpose the substantive discussion on the concept of manipulation to Articles 8–9 and 5 UCPD.

3.1 | Aggressive commercial practices, Article 8 UCPD

According to Article 8 UCPD, an aggressive commercial practice consists in the use of a particularly inacceptable form of influence (harassment, coercion or undue influence) which significantly impairs the average consumer's freedom of choice and thereby alters her decision.

[71] This assumes the applicability of the GDPR, see below, Section 4.

[72] See N. Helberger, 'Profiling and targeting consumers in the internet of things', in R. Schulze and D. Staudenmayer (eds), *Digital Revolution: Challenges for Contract Law in Practice* (Hart/Nomos 2016), 135, 145–47; Helberger et al., n. 14; C. Goanta and S. Mulders, '"Move fast and break things": Unfair commercial practices and consent on social media' (2019)8 *Journal of European Consumer and Market Law*, 136, 143; Laux et al., n. 15; for US law, Willis, n. 8.

[73] See F. Caronna, 'Tackling aggressive commercial practices' (2018)43 *European Law Review* 880, 897 et seq.

[74] See, e.g. J. Obar and A. Oeldorf-Hirsch, 'The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services' (2020) 23 *Information, Communication & Society*, 128; O. Ben-Shahar and A. Chilton, 'Simplification of privacy disclosures: an experimental test' (2016) 45(S2) *Journal of Legal Studies*, S41.

3.1.1 | Undue influence and the judgment in *Orange Polska*

As Natali Helberger has argued, breaches of privacy in algorithmic targeting may amount to harassment, for example if the targeting is delivered on personal digital assistants.[75] In the context of algorithmic manipulation, however, the primary legal gateway among the alternatives of Article 8 UCPD is undue influence, being itself defined in Article 2(j) UCPD as the exploitation of 'a position of power in relation to the consumer so as to apply pressure […] in a way which significantly limits the consumer's ability to make an informed decision'. Article 9 UCPD lists a number of criteria that must be taken into account to determine whether undue influence was exercised.

In the recent *Orange Polska* decision, the CJEU ruled that a practice may constitute an aggressive practice if it 'is liable to make that consumer feel uncomfortable and thus to *confuse his thinking* in relation to the transactional decision to be taken'.[76] At first glance, this passage could suggest that the targeting of cognitive bias or emotional weakness should be considered an undue influence, if the effect is so significant as to confuse the consumer. However, such an interpretation of the decision faces two convincing objections. First, according to Article 9(c) UCPD, a situation-specific impairment of the consumer's decision-making capacities must be taken into account if the trader is *aware* of it.[77] This was clearly the case in *Orange Polska*, but as we have seen, the exploitation of emotional or cognitive weaknesses may be an unintentional side-effect of contractual optimisation by means of machine learning. While one might argue that traders should not be rewarded for using inscrutable models, which leave traders unaware of exploitative features, the wording of the UCPD quite clearly says that, for an aggressive practice under Article 9(c), the trader must be "aware", and not "should have been aware".[78]

Even if the trader was aware of the manipulative potential of its model, however, the second and more forceful objection points to the inextricable connection of undue influence with "pressure", according to Article 2(j) UCPD. This entails, in my view, that deception cannot suffice, but that influence must be exerted in a way consciously perceived by the consumer—one cannot be pressured without noticing it. Hence, pressure clearly fails to capture the hidden-influence prong of algorithmic manipulation. It might, however, cover the other prong, the impairment of rational choice. This raises the question whether a mere advertisement or contractual offer may, even if openly targeted at emotional or cognitive weaknesses, be said to exert "pressure".

3.1.2 | Contractual offers and undue influence

The term "pressure" is not defined in the UCPD, but is often said to presuppose the threat of disadvantages for the consumer *outside* of the concrete offer.[79] Therefore, it is submitted that a mere contractual offer cannot, in principle, exert pressure, as it only sets incentives and need not be accepted.[80] The German Federal Court for Private Law (BGH) has endorsed this position in cases concerning contractual bundling practices.[81]

In certain cases of urgent demand, however, one might argue that the contractual offer amounts to more than just an option, which the consumer may accept or reject at will: if the consumer is: (i) in significant need of the offered product; and (ii) has no reasonable alternative to the offered contract. This is precisely the argument advanced in the context of Article 7(4) GDPR:[82] that bundling prohibition is, according to the majority view,

[75] Helberger, n. 72, 156.

[76] Case C-628/17, *Orange Polska*, ECLI:EU:C:2019:480, para. 47 [emphasis added].

[77] See Helberger et al., n. 15, Part I., para. 169, 171.

[78] See Caronna, n. 73, 897 et seq. (demanding intentionality); H. Köhler, in H. Köhler, J. Bornkamm and J. Feddersen (eds.), *UWG* (Beck, 38th edn. 2020) §4a para. 1.59; I. Scherer, 'Die Neuregelung der aggressiven geschäftlichen Handlungen in §4a UWG' (2016) *Gewerblicher Rechtsschutz und Urheberrecht (GRUR)*, 233, 239.

[79] See, e.g. OLG Munich, U 2225/15 Kart, *GRUR* 2017, 1147 para. 190; Scherer, n. 78, 238 et seq.; Köhler, n. 78, §4a para. 1.59.

[80] Köhler, n. 78, §4a para. 1.30; H. Köhler, 'Kopplungsangebote neu bewertet' (2010) *GRUR*, 177, 182; Scherer, n. 78, 239.

[81] BGH, Case I ZR 4/06, *GRUR* 2011, 532 para. 22; Case I ZR 192/12, *GRUR* 2014, 686 para. 38.

[82] See, e.g. F. Zuiderveen Borgesius et al., 'Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation' (2017) 3 *European Data Protection Law Review*, 353, 361.

supposed to apply in such conditions of necessity, and render a consent bundled with a contractual offer invalid, precisely because the data subject does not, in these cases, have a free choice to deny consent if she wants to conclude a contract. Similarly, one might say that, for the UCPD, the line between incentive and pressure has been crossed if the consumer desperately needs the object of the contract. This would imply that, in functioning markets, manipulative targeting does not amount to an aggressive practice, but that it may when competition is substantially weakened and alternatives are not available, for example in the case of significant network effects (e.g. social networks; basic amenities). In this vein, the Italian Competition Authority (ICA) has fined Facebook for its default settings concerning widespread data collection. More precisely, the ICA found that when 'users decide to limit their consent, they are faced with significant restrictions on the use of the social network and third-party websites/apps, which induce users to maintain the pre-selected choice'.[83] In the understanding of the ICA, when contractual choice is steered by specific disadvantages connected to certain choices, this amounts to undue influence.[84]

Again, however, the question remains whether an element of pressure can really be identified. Ultimately, in my view, the more convincing arguments advise against qualifying exploitative contractual offers as such as aggressive commercial practices,[85] even in cases of necessity. First, the pressure is generally not exerted by the trader, as required by Article 2(j) UCPD; rather, it results from the needs and budget constraints of the consumer.[86] Note that this is different in the Facebook case just cited, where Facebook actively threatened a reduced functionality.[87] Second, however, even in these cases the element of pressure would, if anything, derive from the situation of necessity, not from the targeting of emotional or cognitive traits. Therefore, the regime of aggressive commercial practices generally fails to do justice to such marketing actions.

It should be noted, however, that in some cases, pressure might arise if the targeted bias or emotional state makes the consumer feel "locked in", and the trader is aware of this. For example, the offline pressure element from *Orange Polska* (a courier insisting on an immediate contract signature) could be replicated online by sending time-limited offers to consumers who have been explicitly primed beforehand to include the conclusion of the contract in their endowment position. In this situation, the combination of loss aversion[88] (with respect to the contract) with time pressure could pass the threshold of significant disadvantage required for pressure in the sense of Art. 8 UCPD. Such a finding might seem particularly likely if the trader, in addition, consciously triggered a strong psychological craving for the product such that it appears without any reasonable alternative to the consumer. However, in my view, beyond the "irrational attractiveness" of the offer, an additional element of "insistence" by the trader is still needed (as by the courier in *Orange Polska*), e.g. by reference to some vague external threat, social pressure, or a fictitious legal obligation to contract. Barring that, if courts continue to hold the position that the loss of the contractual opportunity itself (even if subjectively exacerbated by loss aversion) does not qualify as a disadvantage in the context of pressure,[89] even these cases will likely be decided under the general clause, to which we now turn.

3.2 | The general clause, Article 5 UCPD

Since manipulative targeting does not, generally, work through "pressure", the general clause of Article 5 UCPD is particularly relevant to determine if such practices are substantially unfair. As is well known, unfairness under the

[83] Italian Competition Authority, 'Facebook fined 10 million Euros by the ICA for unfair commercial practices for using its subscribers' data for commercial purposes', Press Release, 7 December 2018, https://en.agcm.it/en/media/press-releases/2018/12/Facebook-fined-10-million-Euros-by-the-ICA-for-unfair-commercial-practices-for-using-its-subscribers%E2%80%99-data-for-commercial-purposes.

[84] Id.

[85] See also Franz Hofmann, 'Der maßgeschneiderte Preis' (2016) *Wettbewerb in Recht und Praxis* 1974, 1081.

[86] See also Köhler (n. 78), §4a para. 1.73.

[87] Even such an admonishment, however, need not count as pressure, in my opinion; otherwise, differential pricing of different contractual options ('contract menus', a standard and widely accepted practice) would also amount to undue influence.

[88] See, e.g. N. Novemsky and D. Kahneman, 'The boundaries of loss aversion' (2005) 42.2 *Journal of Marketing Research*, 119.

[89] See nn. 79–81.

general clause presupposes, according to Article 5(2) UCPD, a breach of professional diligence (Section 3.2.1), and a material distortion of consumer decision-making (Section 3.2.2).

3.2.1 | Professional diligence

Article 2(h) UCPD defines professional diligence as behaviour 'commensurate with honest market practice and/or the general principle of good faith in the trader's field of activity'. To determine the limits of professional diligence, it must be acknowledged that every type of commercial practice somehow affects consumer decision-making in the interest of the trader. Therefore, a balance of the different involved interests must be struck to determine which type of influence does breach professional diligence.[90] As the CJEU has ruled in *Deroo-Blanquart*, the legitimate expectations of the average consumer act as the primary yardstick in this exercise.[91] In our context, it seems important to differentiate, again, between intentional cognitive and emotional targeting on the one hand and unintentional manipulative side effects of machine learning on the other.

3.2.1.1 | Intentional targeting

As I will argue, the *intentional* targeting of cognitive biases or emotional weaknesses should, in principle, be deemed to violate professional diligence (3.2.1.1.1).[92] Intention, in this context, does not constitute a necessary condition for the breach of professional diligence, as the discussion of unintentional targeting will show.[93] Rather, intention should be deemed sufficient to shift the burden to the trader: It is then incumbent on him to show that the concrete offer or advertisement is not legally reprehensible, for example because the consumer validly consented to the practice (3.2.1.1.2).

3.2.1.1.1 | A breach, in principle

There are three main reasons for deeming the intentional targeting of cognitive or emotional traits, for example via Emotional AI, a violation of professional diligence, at least in principle. First, the interests of the traders to engage in effective advertising or contract design are of significantly reduced weight if potentially manipulative strategies are deployed in an intentional matter. Traders cannot generally expect practices to be condoned under the UCPD when seeking to benefit from significant consumer weaknesses as identified in the foundational part (intensity; combination; complexity variety).

Second, on the consumers' end, such targeting makes it highly difficult to fully foresee the consequences of the contract,[94] be it because important features of the offer are tailored to boundedly rational traits (e.g. over-optimism[95]), or because the trader explicitly benefits from hot emotional states, which similarly displace a cool weighting of costs and benefits.[96] To be sure, there is significant case law dealing with so-called "emotional advertisements," such as the (in)famous Benetton "shock picture campaign".[97] While such practices are now often condoned, even if they are (also) designed to generate sales by spurring emotional responses,[98] emotional targeting is strikingly

[90] H. Köhler, 'Kopplungsangebote neu bewertet' (2010) *GRUR*, 177, 182.

[91] CJEU, Case C-310/15, *Deroo-Blanquart*, ECLI:EU:C:2016:633, para. 34.

[92] In a similar vein, J. Trzaskowski, 'Behavioural innovations in marketing law', in H.-W. Micklitz, A.-L. Sibony and F. Esposito (eds), *Research Methods in Consumer Law* (Edward Elgar, 2018), 316 et seq.

[93] See below, Section 3.2.1.2.

[94] See Köhler (n. 90), 182.

[95] N. Weinstein, 'Optimistic biases about personal risks' (1989) 246.4935 *Science*, 1232.

[96] E. Johnson and A. Tversky, 'Affect, generalization, and the perception of risk' (1983) 45 *Journal of Personality and Social Psychology*, 20; J. Lerner et al., 'Emotion and decision making' (2015) 66 *Annual Review of Psychology*, 799, 803.

[97] See, e.g. T. Wilhelmsson, 'Harmonizing unfair commercial practices law' (2006) 44 *Osgoode Hall Law* Journal, 461, 481; German Constitutional Court, Case 1 BvR 426/02, *NJW* 2003, 1303.

[98] See, e.g. J. Trzaskowski, 'Behavioural Economics, Neuroscience, and the Unfair Commercial Practises Directive' (2011) 34 *Journal of Consumer Policy* 377, 389; BGH, Case I ZR 55/02, *GRUR* (2006) 75; H. Hartwig, 'Der BGH und das Ende des Verbots "gefühlsbetonter Werbung"' (2006) *Neue Juristische Wochenschau NJW* 1326, 1327; H. Köhler, in H. Köhler, J. Bornkamm and J. Feddersen (eds), *UWG* (Beck, 38th edn. 2020) §3 paras. 9.5–9.6; but see H.-W. Micklitz, 'Unfair commercial practices and misleading advertising' in H.-W. Micklitz, N. Reich and P. Rott, *Understanding Consumer Law* (Intersentia, 2009), 61, 103.

different in important ways. It does not necessarily seek to actively stir certain emotions in the addressee, the effectiveness of which is highly unclear and non-personalised; rather, as mentioned in the foundational part, it measures existing, individual emotions by technical means and harnesses them by confronting the consumer with content optimised in view of the measurement. While there is a significant margin of error,[99] the technology can be expected to improve over time, and studies show that it already now works reasonably well in many cases.[100] Hence, whereas consumers are—particularly when endowed with high emotional intelligence[101]—in a position to recognise and withstand traditional emotional advertisements (e.g. shock pictures),[102] they generally cannot guard against the hidden influence of emotional targeting if they are analysed without being conscious of it. While consumers are in principle aware that they are often tracked online and offline, they will not expect, for the time being, to be subjected to automated analysis of cognition or emotion when shopping. If such hidden error detection is coupled with targeted offers, this will arguably amount to a violation of autonomy, and—as behavioural contract theory shows—it may generate individual and potentially also social welfare losses.[103]

Third, importantly, such targeting may infringe the GDPR and the fundamental right of data protection, if personal data is processed, and privacy norms like Article 7 of the Charter of Fundamental Rights (Charter) and the ePrivacy Directive (ePD), which apply to non-personal data as well.[104]

Taking privacy and the GDPR into account

Privacy and data protection implications seem crucial when evaluating cognitive and emotional targeting. However, before taking this point for granted, one must pause to ask whether such considerations may play a role in the realm of the UCPD at all. In EU competition law, for example, it is hotly debated whether privacy and data protection violations may count toward the breach of competition law or not.[105]

In the case of unfair commercial practice law, two arguments could potentially be advanced to isolate that field from privacy and data protection considerations. First, from a systematic perspective, one could argue that privacy aspects were only taken up, in the realm of unfair commercial practice law, in Article 13 ePD—a hybrid norm between unfair competition and privacy law[106]—and therefore cannot have implications for the interpretation of the UCPD (*argumentum ex negativo*). Indeed, privacy is not expressly mentioned in the UCPD. Second, a doctrinal foundation for limiting the influence of the GDPR on the UCPD can be found in Article 3(4) UCPD, which installs the primacy of other EU law 'regulating specific aspects of unfair commercial practices' over the UCPD in case of conflict. On the basis of this *lex specialis* provision, some scholars indeed argue that the GDPR falls outside of the purview of the UCPD and hence cannot be considered for its interpretation in the first place.[107]

However, this conclusion does not seem warranted. First, it is quite doubtful whether the GDPR really can be considered *lex specialis* to the UCPD, given the general-purpose nature of the Regulation.[108] Rather, the two instruments seem to constitute intersecting, and not concentric, circles. The GDPR clearly covers only one aspect of commercial practices—data protection—and generally does not outlaw them *because* they qualify as unfair in the sense of the commercial practice law.[109] Second, interpreting the UCPD in the light of the GDPR precisely aims at avoiding

[99] Barrett (n. 61), 3.

[100] Burr and Cristianini (n. 1), 472 et seq. (e.g. 72% test accuracy).

[101] J. Yip and S. Coté, 'The emotionally intelligent decision maker' (2013) 24 *Psychol. Sci.*, 48.

[102] Trzaskowski (n. 98), 385.

[103] See n. 26.

[104] See O. Lynskey, *The Foundations of EU Data Protection Law* (Oxford University Press, 2015), 90.

[105] See, e.g. Costa-Cabral and Lynskey (n. 14); V. Robertson, 'Excessive data collection: Privacy considerations and abuse of dominance in the era of big data'(2020) 57 *Common Market Law Review*, 161; see also BGH, Case KVR 69/19.

[106] See Recital 41 ePD; EDPB, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (2019) 14; see also below, Section 4.1.2.

[107] H. Köhler, in H. Köhler, J. Bornkamm and J. Feddersen (eds), *UWG* (Beck, 38th edn, 2020) §3 para. 3.10; H. Köhler, 'Durchsetzung der DS-GVO mittels UWG und UKlaG?' (2018) *Wettbewerb in Recht und Praxis* 1269, 1275.

[108] See also H.-W. Micklitz, 'Unfair practices and misleading advertising', in N. Reich et al. (eds.), *European Consumer Law* (Intersentia, 2nd edn., 2014), 82 et seq.

[109] See T. Wilhelmsson, 'Scope of the directive', in G. Howells, H.-W. Micklitz and T. Wilhelmsson (eds), *European Fair Trading Law: The Unfair Commercial Practices Directive* (Ashgate, 2006) 49, 76.

the conflict mentioned in Article 3(4) UCPD. Without conflict, however, there is no primacy of the specific instrument pursuant to said Article. Third, if the GDPR really did take precedence over the UCPD, the application of the latter would be precluded whenever personal data are processed in the digital economy—a consequence that stands in stark contradiction to the case law of Member State courts[110] and to the vast majority of scholarship dealing precisely with this matter.[111]

In my view, it is therefore preferable to integrate privacy and data protection dimensions into the UCPD analysis—just like UCPD considerations influence data protection law.[112] Even if the UCPD does not mention privacy, recitals 6 and 18 explicitly reference the principle of proportionality. According to the CJEU, this implies that, in a horizontal setting, the fundamental rights of different private actors need to be weighed and brought into a proportionate balance.[113] If, therefore, the freedom to conduct a business is considered on the one hand, consumer rights cannot be restricted to an understanding of consumer autonomy as unconstrained choice, but must include other fundamental rights positions, such as privacy and data protection.[114] In fact, many privacy theorists precisely argue that privacy and data protection, ultimately, safeguard individual autonomy.[115]

Taking privacy and data protection into account seems particularly convincing for a general clause like Article 5 UCPD and the definition in Article 2(h) UCPD, which operates with vague terms like "honest market practice" and "good faith".[116] In fact, such an interpretation of the UCPD presents a further, important building block in the development of an integrated market order for the digital economy. If the latter fuses economic transactions and privacy-sensitive data processing, the law should, to the extent that it is doctrinally possible and methodologically sound, mirror this tendency by building further bridges between different legal fields once considered separate.[117]

The effects of privacy and GDPR breaches

This link is important because, for potential plaintiffs, it would present one further route to demonstrating a lack of professional diligence, beyond the often technical and empirically difficult question of *intentional* targeting. If the GDPR applies, a breach may have already been registered by a Data Protection Authority. As a case in point, the processing of sensitive data for targeting may violate Article 9(1) GDPR.[118]

If Article 3(4) UCPD does not bar the GDPR from being considered, the question still remains, however, what exactly a GDPR breach, or the infringement on privacy norms, should imply for professional diligence. In my view, one should distinguish between cases in which the trader is simultaneously the data controller (see Art. 4(7) GDPR) responsible for the GDPR violation/privacy breach, and cases in which he is not. In the former case, there is much to be said for qualifying a GDPR violation or a privacy breach by the trader as contrary to professional diligence, unless the practice concerns a purely formal or petty breach.[119] While not every violation of any law by the trader will constitute a breach of professional diligence, those types of market regulation which are of immediate relevance for the use of the product or for sales and distribution mechanisms should be considered relevant for Article 5(2)(a) UCPD.[120] Concerning the use of models processing personal data, the GDPR, and privacy law more generally, is the epitome of

[110] See, e.g. LG Berlin, Case 16 O 341/15, ECLI:DE:LGBE:2018:0116.16O341.15.0A, para. 49–51; KG Berlin, Case 5 U 9/18, ECLI:DE:KG:2019:1220.5 U9.18.0A, para. 58 et seq.

[111] See, e.g. the references in n. 72.

[112] See, e.g. Art. 21(2) GDPR; Article 29 Working Party, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC', WP 217, 2015, 45–47; see also below, Section 4.

[113] See, e.g. Case C-314/12, *UPC Telekabel*, ECLI:EU:C:2014:192, para. 63.

[114] See also Zuiderveen Borgesius (n. 7), 155.

[115] J. Kupfer, 'Privacy, autonomy, and self-concept' (1987) 24 *American Philosophical Quarterly* 81; P. Bernal, *Internet Privacy Rights: Rights to Protect Autonomy* (CUP, 2014); see also Zuiderveen Borgesius (n. 7), ch. 3.

[116] See A. Röthel and F. Möslein, 'Concretisation of General Clauses' in K. Riesenhuber (ed), *European Legal Methodology* (Intersentia, 2017) 261; see also CJEU, Joined Cases C-240/98 to C-244/98, *Océano*, ECLI:EU:C:2000:346, paras. 22–24.

[117] See the references in n. 14.

[118] See, e.g. Á. Cuevas et al., 'Does Facebook use sensitive data for advertising purposes', *arXiv preprint arXiv:1907.10672* (2019).

[119] See also, for the use of unfair contract terms, BGH, Case I ZR 45/11, *NJW* 2012, 3577, para. 46.

[120] C. Alexander, 'Vertragsrecht und Lauterkeitsrecht unter dem Einfluss der Richtlinie 2005/29/EG über unlautere Geschäftspraktiken', (2012) *Wettbewerb in Recht und Praxis*, 515, 520.

a regulation that directly implements (most likely even horizontally applicable[121]) fundamental rights and specifically controls the traders' activities. Consumers legitimately expect compliance with these important instruments of digital market regulation;[122] therefore, they must be heeded by the trader as a matter of professional diligence. This view is further supported by a recent empirical study which shows that more than 50% of consumers reject targeting by Emotional AI *per se* and another 30% if they are personally identifiable.[123] If indeed, as the CJEU ruled in *Deroo-Blanquart*,[124] consumer expectations are key in determining professional diligence, then these empirical findings imply that privacy and data protection clearly should matter, and strongly suggest a breach of professional diligence in the case at issue.

Matters are more complex, however, if the trader is not the data controller responsible for the GDPR violation or for the privacy infringement. For example, the trader may use a model which was trained on personal data, in breach of the GDPR, by a separate AI developer. In this case, it would probably go too far to demand, as a matter of professional diligence, that the trader checks, in all cases, to what extent the training and other third-party procedures fully complied with the GDPR and EU privacy law. This would result in an excessive burden for the trader, particularly given the complex technical nature of the subject. Moreover, it would be contradictory to hold the trader liable for a GDPR breach under the UCPD if he is not even liable under the GDPR (because of lacking controller status). However, if objective indications are present that any of these norms were breached by the developer, for example because a Data Protection Authority has launched an inquiry into the product, then the trader should be required to undertake further due diligence to ensure that the model was developed in a GDPR- and privacy-compliant way. Hence, in my view, if the trader is not the data controller responsible for the GDPR/privacy breach, professional diligence is only violated if the trader knew or should have known that these norms were not complied with.[125]

Consequences for manipulative targeting

Generally speaking, therefore, the trader can only be faulted for GDPR or privacy breaches, in the context of professional diligence, if he controlled the breach, or should have known of it. In the case of intentional cognitive or emotional targeting, however, it can generally be assumed that the trader voluntarily installed the algorithmic tool and therefore controlled the targeting to an extent that makes him responsible, not only in the sense of joint controllership pursuant to Article 4(7) GDPR,[126] but also for matters of professional diligence. In the case of intentional targeting, the trader must be aware of the privacy implications of emotional and cognitive measurements. As a consequence, the intentional targeting of emotions or cognitive weaknesses should, in general, be qualified as a violation of professional diligence: empirically supported consumer expectations and data protection/privacy interests outweigh the trader's interests in mood- or cognition-based marketing.

3.2.1.1.2 | Defences and the role of consent

While the intentional targeting of cognitive bias or emotions should, in principle, be considered a breach of professional diligence, this conclusion does not hold in all cases and contexts. Under some circumstances, the balance of interests may tip toward the admissibility of cognitive and emotional targeting. The burden of argumentation and proof, however, now rests on the trader.[127] For example, the trader may argue that, despite the targeting of weaknesses, the transaction is in the best interest of the consumer. For instance, if an Emotional AI application detects that a person strongly fears an infection with COVID-19, a trader could offer specific face masks or other, genuinely

[121] See CJEU, Case C-414/16, *Egenberger*, ECLI:EU:C:2018:257, para. 76; Joined Cases C-569/16 and C-570/16, *Bauer and Willmeroth*, ECLI:EU:C:2018:871, para. 89 et seq.

[122] See, e.g. European Commission, 'Attitudes on data protection and electronic identity in the European Union', Special Eurobarometer 359 (June 2011), 74–75 and 148–149; European Commission, 'e-Privacy', Flash Eurobarometer 443 (2016), 4–5.

[123] McStay (n. 3), 9.

[124] See n. 91.

[125] But see Helberger et al. (n. 15), Part I., para. 177.

[126] See particularly CJEU, Case C-40/17, *Fashion ID*, ECLI:EU:C:2019:629, para. 75.

[127] See Helberger et al. (n. 15), Part I., para. 184 et seq.

useful products for dealing with the pandemic. An indication for such mutually beneficial offers would be that consumers do not regret the transaction.

One further important angle would be for the trader to argue that the requirements of professional diligence were met because the consumer *consented* to the targeting practice. This raises two important questions. First, can consumers validly consent to practices that would otherwise be ruled unfair in the sense of Article 5 UCPD, that is, to being manipulated? And if so, second, what specific requirements must be met for such consent to be valid?

Consenting to manipulation

The first question raises the issue of inherent limits of consent. Generally speaking, individual autonomy should enable individual actors to consent, in an informed way, even to practices others might consider manipulative or otherwise deleterious, such as cognitive or emotional targeting. It should be noted at the outset that the purpose of consent is frustrated if manipulation works by *hidden* influence, as consent must be informed, or by the *impossibility* of rational choice, as that deficiency will likely also affect the consent declaration. For example, in case of cognitive deficiencies amounting to serious mental impairments, national rules on the limits of party autonomy may apply (incapacity).[128] Consent remains possible, however, if manipulation occurs through the mere significant diminution, but not total eclipse, of rational decision-making,[129] which does not reach the threshold of incapacity *stricto sensu*. Finally, consent is of obvious importance if the threshold to a breach of professional diligence is crossed because, and only because, of a GDPR violation. In that case, consent according to the GDPR removes that infringement and, as a consequence, the UCPD breach.

Even in these cases, however, the power of individual consent is limited by Member State rules on unconscionability, public policy or *bonos mores* which apply to contracts as well as unilateral declaration such as consent.[130] Therefore, while consumers may in principle consent to practices which are unfair in the sense of Article 5 UCPD, consent will likely be invalid in case of particularly egregious practices considered invalid under national public policy causes. An example may be the targeting of strong negative emotions to offer contracts that impose an existential financial risk on the consumer.

Transposing GDPR rules on consent

More importantly, second, even within these fairly wide bounds of *bonos mores* and incapacity, it seems evident that, even though explicit conditions for consent are lacking under the UCPD, not just any waiver may suffice to exonerate the trader. Therefore, it is submitted that, as a minimum standard, consent meeting the exigencies of professional diligence should conform to the general requirements of the GDPR.[131] Hence, it must be freely given, specific, informed and unambiguous (Art. 4(11), 7(4) GDPR), and freely revocable (Art. 7(3) GDPR). There are two reasons for transplanting this standard from the GDPR to the UCPD in the context of the digital economy.

First, it is the declared aim of the UCPD to safeguard consumer autonomy in market transactions.[132] However, as Recital 7 GDPR clearly shows, consent is thought to be the main tool to facilitate control and data sovereignty—in other words: autonomy—in digital contexts in which personal data is processed. Second, and more concretely, a systemic argument can be deduced from the ePD, which specifies that consent is generally necessary for direct marketing by automatic calling machines, fax or email (Art. 13(1)). Importantly, the definition of consent in Article 2(f) ePD merely references the Data Protection Directive, and hence now, according to Article 94(2) GDPR, the consent regime of the GDPR itself. While cognitive and emotional targeting will often pursue communication channels—from advertisements on adaptive in-store screens to social media—different from the ones mentioned in Article 13(1) ePD, an argument can be made that these novel channels are sufficiently similar to the more traditional forms

[128] In the alternative, such consent could be deemed to not be "freely given" in the sense of Art. 4(11) GDPR; see the next section.

[129] See the discussion in Section 2.1.1.

[130] See, e.g. A. Ohly, *Volenti non fit iniuria* (Mohr, 2002), 408 et seq.

[131] See also Goanta and Mulders (n. 72), 144.

[132] See, e.g. H.-W. Micklitz, 'The general clause on unfair practices' in G. Howells, H.-W. Micklitz and T. Wilhelmsson (eds), *European Fair Trading Law: The Unfair Commercial Practices Directive* (Ashgate, 2006) 83, 102, 104 et seq.

of direct marketing mentioned in Article 13(1) ePD to trigger a need for the protection of individual autonomy via consent. To complicate things further, however, Article 13(3) ePD offers Member States an option to deal with direct marketing by means other than those mentioned in paragraph one. Member States may choose whether they require consent, or a mere right to object. In fact, to the extent that targeting individualises the recipients of sales communications, it must be taken to constitute direct marketing, so that Article 13(3) ePD is directly applicable.

Importantly, however, even if cognitive or emotional targeting does constitute direct marketing in the sense of Article 13(3) ePD, this does arguably not prevent the UCPD from installing different and potentially stricter requirements (such as consent, not merely a right to object) *if* the commercial practice is otherwise deemed unfair according to the standards of the UCPD.[133] Admittedly, this presupposes an assessment of UCPD unfairness independent of the standards of the ePD. This might be thrown into doubt by the fact that a key reason for deeming intentional targeting a breach of professional diligence is the very implication of privacy. However, the privacy concerns raised above differ from the reasons for regulating unsolicited communications in Article 13 ePD, which mainly constitute a bother and a physical infringement of the private sphere of the consumer, but which are not based, in general, on an analysis of cognitive or emotional traits. Moreover, other than Article 5 UCPD, Article 13 ePD does not require at all the finding of material distortion. Therefore, it is submitted that the assessment of Article 5 UCPD in our context is sufficiently different from the risks contemplated in Article 13(3) ePD to justify a standard that deviates from this norm. Hence, for example, consent at a moment in which strong negative emotions dominate might not be "freely given" (cf. Art. 4(11) GDPR), particularly if the trader specifically analyses and targets that type of emotion. As a consequence, professional diligence might still be breached due to a lack of valid consent, despite the merely optional consent framework in Article 13(3) ePD.

Transposing the GDPR standard of consent to the UCPD will not, however, magically solve all problems of consumer autonomy. Rather, from a policy perspective, lessons must clearly be drawn from the widely acknowledged failure of consent to perfectly safeguard autonomous choice in the context of data protection law.[134] Therefore, the GDPR rules on consent should only be taken as a minimum requirement for the sake of professional diligence, not ruling out further policy options discussed in the last part of the article.

Overall, therefore, intentional targeting of cognitive or emotional weaknesses will breach professional diligence unless specific circumstances, such as valid consent, clearly suggest that the exigencies of good faith were honoured. Such an assessment is in line with an older ruling of the CJEU[135] and more recent scholarship[136] stressing that excessive teaser effects of certain offers, which are generally brought about intentionally, must be considered as violations of professional diligence. Under the framework advanced in this article, excessive teaser rates can be considered manipulative (and a breach of professional diligence) if they expressly target vulnerable consumers in emotional or cognitive states significantly diminishing their capacity to rationally evaluate the teaser offer (e.g. strong present bias[137]), and particularly so if the psychological measurement infringes the GDPR. The main problem, then, lies in proving intentionality in a technologically complex machine learning environment. While this may be presumed in Emotional AI actively measuring emotions, it is much more difficult when emotions or cognitive states are merely captured indirectly, via proxy variables or context.

3.2.1.2 | Unintentional targeting and algorithmic auditing

Such unintentional targeting, arising as a side effect of machine learning optimisation, is arguably more difficult to evaluate with respect to professional diligence. The definition of that term mentions good faith.[138] However, as

[133] See G. Howells, 'Aggressive commercial practices' in G. Howells, H.-W. Micklitz and T. Wilhelmsson (eds.), *European Fair Trading Law: The Unfair Commercial Practices Directive* (Ashgate, 2006) 167, 181 et seq.

[134] Zuiderveen Borgesius (n. 7), Chapter 7; P. Blume, 'The inherent contradictions in data protection law' (2012) 2 *International Data Privacy Law*, 26; Hacker (n. 49), 274.

[135] See CJEU, Case 286/81, *Oosthoek*, ECLI:EU:C:1982:438 para. 18;

[136] Köhler (n. 90), 182 et seq.

[137] See, e.g. O. Bar-Gill, *Seduction by Contract* (Oxford University Press, 2012), 52 et seq.

[138] But see Caronna (n. 73), 880, 886.

explained above (Section 2.2.1), traders may act “in good faith” while their models unintentionally manipulate consumers. On the one hand, one could therefore argue that traders cannot be expected to be smarter than their counterparties and to scrutinise highly complex, technical models for potential tendencies of exploitation.

In my view, however, striking an even balance of risk and return in algorithmic offers suggests the opposite. Those who use machine learning techniques to enhance their expected payoffs must bear the burden of inspecting their model for manipulative features, or having them inspected. This is not a trivial task. However, a growing number of companies do offer algorithmic auditing services.[139] Hence, traders must make a serious, good-faith effort to detect and mitigate manipulative tendencies if they wish to use machine learning models. For example, they may have to check if the key predictive features of their model explicitly build on or closely correlate with consumer weaknesses of the three manipulation varieties discussed (intensity; combination; complexity variety). If they fail to substantially audit for unintentional targeting of errors and weaknesses of their counterparties—be it out of ignorance or because it is technically impossible in the model they use—this should be considered a breach of professional diligence under the same conditions as in the case of intentional targeting: *qui habet commoda ferre debet onera*.[140] This implies that, from the perspective of the UCPD, opaque models may be used by traders, but at their own risk.

3.2.2 | Material distortion

Even if intentional or unintentional targeting violates norms of professional diligence, it must also materially distort consumer decision-making in the concrete situation to be relevant under Art. 5(2) UCPD. Notably, this is also the yardstick for prohibited manipulative practices under Article 5(1)(a) and (b) AIA. In Article 2(e) UCPD, material distortion is defined as a commercial practice that ‘appreciably impair[s] the consumer's ability to make an informed decision, thereby causing the consumer to take a transactional decision that he would not have taken otherwise’. The law, hence, does not require an impossibility of rational choice, either;[141] rather, an appreciable impairment suffices. The exact contours of this concept continue to be disputed, but some guidelines can be extrapolated from the jurisprudence of the CJEU and Member State courts. For example, in the context of point 31 of the black list (Annex I to the UCPD: advertisement with non-existing prize), the CJEU ruled in *Purely Creative* that such practices are aggressive—and by extension unfair in the sense of the general clause—because they seek ‘to exploit the *psychological effect* created in the mind of the consumer by the perspective of having won something and to cause him to take a decision which is *not always rational*’.[142] By analogy, the targeting of cognitive biases may be considered unfair if it does displace rational decision-making.[143] Similarly, the BGH has maintained that a significant impairment of consumer's freedom of choice is reached, at the very latest, if the practice ‘completely eclipses the rationality of the consumer decision’.[144] Clearly, marketing practices have always played on cognitive and emotional effects, for example by “framing” (messages positively) or employing ‘mood congruence’ (mentioning products when addressees are expected to be in a receptive mood).[145] Generally speaking, the threshold to material distortion is therefore only passed if these effects are so significant that, in the context of the entire decision, they plausibly prevent an overriding of boundedly rational or emotional System 1 processes by more rational, System 2-based responses (see Section 2.1).

[139] See, e.g. https://www.neurocat.ai/; https://www.taktile.com/.

[140] See, e.g. R. Zimmermann, *The Law of Obligations* (Juta, 1990), 201, 209.

[141] See Section 2.1.1.

[142] CJEU, Case C-428/11, *Purely Creative*, ECLI:EU:C:2012:651, para. 49 [emphasis added].

[143] See also R. Incardona and C. Poncibo, ‘The average consumer, the unfair commercial practices directive, and the cognitive revolution’, (2007) 30 *Journal of Consumer Policy*, 21; A.-L. Sibony, ‘Can EU consumer law benefit from behavioural insights? An analysis of the unfair practices directive’ (2014) 22 *European Review of Private Law*, 901.

[144] BGH, Case I ZR 25/17, *GRUR* 2018, 1063 para. 14 [translation by the author].

[145] G. Bakamitsos and G. Siomkos, ‘Context effects in marketing practice: The case of mood’ (2004) 3 *Journal of Consumer Behaviour* 304.

3.2.2.1 | *Average consumers*

A key question therefore is where this threshold is situated with respect to average consumers, which Article 5 (2) UCPD primarily references. This standard applies to general offers which do not target any specific subgroup of consumers. At least in principle, the notion of the average consumer, the behavioural critique notwithstanding,[146] refers to a reasonably well informed and reasonably observant and circumspect person.[147] While the average consumer test is not a statistical test,[148] the CJEU did, in the *Teekanne* judgment, acknowledge that the average consumer of EU law may be prone to ignore, or misunderstand, important product information.[149] Just like *Purely Creative*, *Teekanne* has therefore been interpreted as a knowing nod to the literature on cognitive biases.[150]

However, biases and emotions abound, and are part of many decision processes. Hence, as discussed in the foundations part, their presence should be deemed a material distortion only if it can generally not be expected that the influence is noted and, at least potentially, countered.[151] This may be assumed, for example, if the average consumer is caught completely off-guard, e.g. because of the timing and place of the approach (cf. Art. 9(a) UCPD). Generally speaking, though, if the average consumer is the yardstick it will be difficult to find that rational choice was impaired so significantly by emotional or cognitive appeal that it must be outlawed. However, that assessment may differ in cases in which vulnerable consumers are precisely targeted, for example by Emotional AI, to which we now turn.

3.2.2.2 | *Vulnerable consumers*

As discussed in the foundations section (2.1.2), targeting generally denotes directing offers to actors with specific characteristics. What is special about algorithmic manipulation is that, with the help of data analyses and AI, offers may be targeted toward specific consumers that are particularly susceptible to cognitive biases and emotional weaknesses. This is at odds with the notion that the reference actor should be an average consumer. Rather, it is submitted that the very act of targeting cognitive or emotional weaknesses implies that the yardstick should be shifted toward a more specific analysis of the effects on those vulnerable consumers, pursuant to Article 5(2)(b) and (3) UCPD.

3.2.2.2.1 | *Offers directed to vulnerable consumers, Article 5(2)(b) UCPD*

According to Article 5 UCPD, there are two ways to achieve a “personalisation” of the average consumer. First, if the practice is addressed toward a specific group of consumers, members of this group become the relevant reference actors according to the second half-sentence of Article 5(2)(b) UCPD. This covers cases in which the targeting is intentionally directed toward a specific group,[152] for example toward insecure teenagers in the Facebook Australia case. More generally, cases of Emotional AI explicitly targeting certain emotions must be taken to be specifically “directed” toward that subgroup in the sense of Article 5(2)(b) UCPD.

Specific conditions for manipulation

However, emotions underlie every decision-making process;[153] similarly, many everyday decisions are ridden with bias[154]—hence, appealing to them cannot be generally illegal under the UCPD, even if directed toward certain groups. Nowadays, hyperbolic statements and puffery, which are also supposed to trigger emotions, are rather

[146] See, e.g. Incardona and Poncibo (n. 143), 36; Trzaskowski (n. 98), 391; Sibony (n. 143).

[147] See, in particular, Case C-210/96, *Gut Springenheide and Tusky*, ECLI:EU:C:1998:369, para. 31.

[148] See, e.g. Case C-220/98, *Estée Lauder (Lifting)*, ECLI:EU:C:2000:8, para. 28 et seq.

[149] Case C-195/14, *Teekanne*, ECLI:EU:C:2015:361, para. 36–41.

[150] H. Schebesta and K. Purnhagen, ‘The behaviour of the average consumer’ (2016) 41 *European Law Review* 590, 596; see also Trzaskowski (n. 92), 324.

[151] See P. Hacker, ‘Personalised law and the behavioral sciences’, in C. Busch and A. de Franceschi (eds), *Data Economy and Algorithmic Regulation* (Beck/Hart/Nomos, 2020) 241, para. 42.

[152] C. Alexander, ‘Grundfragen des neuen §3 UWG’, (2016) *Wettbewerb in Recht und Praxis*, 411, 415.

[153] Lerner (n. 96), 816; G. Loewenstein et al., ‘Risk as feelings’ (2001) 127 *Psychological Bulletin*, 267.

[154] See n. 27.

accepted.[155] Therefore, specific conditions concerning the type of weakness or the circumstances must be obtained for a significant deviation from rational decision-making (Art. 2(e) UCPD) to be plausible. In particular, this may be assumed in any one of the three cases introduced in the foundations part (2.1.1): targeting of strong emotions or biases (intensity variety); targeting of mutually reinforcing weaknesses, for example the violation of the long-term preferences of the consumer (combination variety); or targeting vulnerable consumers in complex or unanticipated environments (complexity variety). Hence, if consumers vulnerable along any one of these dimensions are targeted with offers containing features negatively deviating from the general offers of the trader and which are linked to the specific vulnerability, material distortion can be assumed. For instance, this might be a targeted price increase reflecting an inflated willingness to pay due to strong emotions or biases.[156]

In addition, let us consider an example of material distortion in which the targeting of weaknesses runs against the likely long-term preferences of the consumer.[157] For instance, one may imagine that Emotional AI, like eyeQ's or Microsoft's technology,[158] might be installed in airports. It might, inter alia, measure if travellers are afraid of flying. Such information might subsequently be used by rental car companies to target specifically those travellers exhibiting strong fears with suggestions to use a car instead of an airplane for the next trip. These consumers may not regret choosing a car, but the targeting could be considered to influence decision-making in a way that counteracts the long-term safety preferences of the consumers, which in turn are discounted via present bias (combination variety). There is empirical evidence showing that, while flying is statistically safer than driving, fear of flying displaces rational safety decisions and leads travellers to choose a car over a plane.[159]

Shifting the burden

All of the three named conditions were already discussed above in the context of the concept of manipulation. What remains to be asked is if a violation of Article 5(2) UCPD can be found automatically in all of these cases. Conceptually, non-hidden manipulation requires the significant impairment of rational choice.[160] Translated into legal terms, the concrete impairment is often difficult to prove, and therefore becomes a question of evidentiary standards. In my view, in the named three instances a significant impairment of rational choice is sufficiently likely to constitute a prima facie case of Article 5(2) UCPD violation which, again, shifts the burden of argumentation and proof to the trader.[161] Hence, the breach is not automatic in cases of manipulation by way of the intensity, combination or complexity variety, but only presumed, incentivising the trader to disclose otherwise hidden information about the functioning of the model in the real world.[162]

In practice, courts will have to make plausible assumptions to determine whether these conditions are fulfilled. If any of them is, this should be taken as a strong indication that the decision was materially distorted, shifting the burden of proof with respect to material distortion. If, however, none of the three mentioned conditions is met, a trader may, in principle, engage in cognitive or emotional targeting under the general clause of the UCPD, unless there is clear additional evidence of a material distortion. By allowing traders to present even cognitively and emotionally targeted offers if they are unlikely to manipulate consumers, the named conditions install an important backstop against throwing out the baby with the bathwater and facilitate preference matching potentially beneficial to consumers, too.

3.2.2.2.2 | Offers affecting vulnerable consumers, Article 5(3) UCPD

Article 5(3) UCPD covers practices relating to explicitly enumerated vulnerable groups. Notably, Article 5(1)(b) AIA now contains a similar provision sanctioning physical or psychological harm resulting from the exploitation, by AI

[155] Trzaskowski (n. 98), 388.

[156] See nn. 34–35 and accompanying text.

[157] See K. Eliaz and R. Spiegler, 'Contracting with diversely naive agents' (2006) 73 *Review of Economic Studies*, 689, 690.

[158] See nn. 58 and 66.

[159] G. Gigerenzer, 'Dread risk, September 11, and fatal traffic accidents' (2004) 15 *Psychological Science*, 286.

[160] See above, Section 2.1.

[161] See Helberger et al. (n 15), Part I., para. 184 et seq.

[162] See I. Ayres and R. Gertner, 'Filling gaps in incomplete contracts' (1989) 99.1 *Yale Law Journal*, 87, 91 et seq. (on penalty defaults).

systems, of vulnerabilities related to 'age, physical or mental disability' (see Section 5.2). In the UCPD, Article 5 (3) changes the benchmark to vulnerable consumers if those consumers are exclusively affected and the vulnerability stems from 'mental or physical infirmity, age or credulity in a way which the trader could reasonably be expected to foresee'. The practice, therefore, need *not* be intentionally directed toward that group. Hence, Article 5(3) UCPD could cover unintentional cognitive or emotional targeting. If it does, a material distortion of that group should be found under the same three conditions as in Article 5(2)(b) UCPD. However, there are several significant difficulties in qualifying unintentional manipulative targeting as a case of Article 5(3) UCPD.

First, cognitive biases or emotional weaknesses would have to count as "mental infirmity" or "credulity". Indeed, an expansive reading of these elements would include any significant deviations from rational decision-making, including even cognitive biases and severe emotional weaknesses ("hot states"). The Commission, in its guidelines, did mention "behavioural characteristics"[163] for the interpretation of Article 5(3) UCPD, and noted that it covers 'a wide range of situations'.[164] On the other hand, a number of scholars suggest limiting the content of the mentioned categories to exceptional cases.[165] Similarly, the CJEU tends to interpret exceptions, like the present deviation from the average consumer standard, restrictively. Therefore, it is quite doubtful whether cognitive biases and emotional weaknesses are covered under this provision, unless they assume the severity of a recognised medical condition.[166]

Second, the practice must, according to the wording, *exclusively* affect members of the vulnerable group ("only").[167] Offers digitally targeted to emotionally or cognitively weak consumer subgroups are problematic with respect to this element, too. As is well known, machine learning operates on a probabilistic basis, and models will therefore always make false positive predictions as well. Hence, an ad designed for addressees ranking high in fear will almost inevitably reach some consumers who are not fearful at all. In my view, however, this modelling error should not exonerate the trader. Rather, it should be dispositive that, from an internal perspective, the practice exclusively affects consumers who were *predicted* to be vulnerable, irrespective of whether they exhibit these traits in reality. Again, this corresponds to an equitable attribution of the advantages and disadvantages of machine learning to the person employing it, i.e. the trader.

Finally, the targeting of vulnerable consumers must be *foreseeable* for the trader, an element demanded by the principle of proportionality (Recital 18 UCPD). In my view, this element links back to the question of professional diligence, which also necessitates a proportionality analysis. Therefore, if a proper algorithmic audit would have uncovered the potential to target specific vulnerable audiences, this should not only imply that professional diligence was breached, but also that the effect on that audience was foreseeable in the sense of Article 5(3) UCPD. In essence, therefore, the restrictions of Article 5(3) UCPD can be overcome if that provision is interpreted rather extensively. While it remains to be seen whether the CJEU will follow this route, it would present an opportunity to effectively counter the growing personalisation of commercial practices with a more personalised legal yardstick.

3.2.2.2.3 | Mapping technical distinctions onto the law

From an economic perspective, such a focus on more homogeneous, vulnerable subgroups via Article 5(2)(b) or (3) UCPD would imply that the effects of cognitive and emotional targeting are easier to determine than in the case of the more heterogeneous average consumer pool. For example, teaser rates, when specifically offered to financially unsophisticated, highly boundedly rational consumers will, with a high likelihood, affect their decision-making in a way that significantly deviates from rational decision-making in a long-term perspective.[168]

[163] European Commission, 'Guidance on the Implementation/Application of Directive 2005/29/EC on Unfair Commercial Practices', SWD (2016) 163 final, 43.

[164] European Commission (n. 163), 28.

[165] M. Friant-Perrot, 'The vulnerable consumer in the UCPD and other provisions of EU Law' in W. van Boom et al. (eds.), *The European Unfair Commercial Practices Directive* (Ashgate, 2014) 89, 100; see also Micklitz (n. 132), 112 et seq.

[166] Hacker (n. 151), para. 47.

[167] This implies that even the rare case in which *only* vulnerable consumers 'fall for' general offers can be analysed under this heading.

[168] See also M. Apetz, *Das Verbot aggresiver Geschäftspraktiken* (Heymanns, 2011), 319; H. Ru and A. Schoar, 'Do credit card companies screen for behavioral biases?', Working Paper No. w22360. National Bureau of Economic Research, 2016; Bar-Gill (n. 137).

Importantly, within the framework offered here, the technical distinction between intentional and unintentional exploitation can be neatly mapped onto the difference between the personalisation of reference actors according to Article 5(2)(b) UCPD ("directed" = intentional targeting) vs. Article 5(3) UCPD ('materially distort … only' = also unintentional targeting). Hence, if the UCPD is recalibrated toward vulnerable consumers in this way, it may adequately capture algorithmic manipulation through both dimensions of cognitive and emotional targeting. As the three conditions (intensity; combination; complexity variety) show, it will often—but not always—violate the general clause.

4 | ALGORITHMIC MANIPULATION AND DATA PROTECTION LAW

Algorithmic manipulation presents veritable and novel challenges for the UCPD. Additionally, however, it may also implicate data protection law. While one could surmise that digital manipulation, particularly when considered unfair under the UCPD, will also often be unlawful under the GDPR,[169] matters are more complicated upon closer scrutiny. Not only is the analysis of digital influence under the GDPR complex, but it is even doubtful whether, if the addressees of a commercial practice are not immediately identified, personal data is being processed at all. Under current law, that is a prerequisite for the GDPR to apply, Articles 2(1), 4(1) GDPR. For example, when an Emotional AI (like the ones sold by eyeQ, Microsoft or Amazon) just analyses affective states in otherwise unidentified passers-by, and triggers a pertinent in-store advertising, some scholars argue that the GDPR does not apply,[170] and opinions by the Article 29 Working Party suggest as much.[171] While this claim would necessitate a more comprehensive analysis,[172] the GDPR quite clearly covers targeting by means of email or advertisements on social media profiles.[173] In the following sections, the applicability of the GDPR to algorithmic manipulation is therefore assumed.

Within the framework of the GDPR, algorithmic manipulation is intimately related to the protection of vulnerable data subjects.[174] For reasons of scope, however, specific GDPR instruments for their protection (e.g. Art. 8, 9, 22, and 35 GDPR) cannot be analysed in detail here. It should be noted, however, that the duty to perform a data protection impact assessment according to Article 35 GDPR not only mirrors the exigencies for algorithmic auditing under the UCPD; but that Recital 75 GDPR, in this context, explicitly mentions risks for vulnerable subjects. The GDPR, therefore, is not blind to their needs. On this basis, the remaining sections will focus on the influence of the UCPD on two specific GDPR norms: the balancing test (Section 4.1), and the principle of fair data processing (Section 4.2). These provisions are influenced, in various ways, by the analysis of algorithmic manipulation under the UCPD.

4.1 | The balancing test, Article 6(1)(f) GDPR

The balancing test of Article 6(1)(f) GDPR is particularly relevant in the case of algorithmic manipulation as targeted advertisements and offers are often based on the secondary use of personal data for which no (valid) consent exists.[175] Hence, the question arises whether the fairness or unfairness of a commercial practice, according to unfair commercial practice law, may tip the balance in the assessment pursuant to Article 6(1)(f) GDPR.

[169] Clifford (n. 70), para. 302; for an analysis of the IoT in the GDPR framework more generally, see S. Wachter, 'Normative challenges of identification in the internet of things' (2018) 34 *Computer Law & Security Review*, 436.

[170] See the interviews with legal scholars in McStay (n. 65), 6–8; McStay (n. 3), 4.

[171] Article 29 Working Party, 'Opinion 3/2012 on developments in biometric technologies', WP 193, 16 (soft biometrics); id., 'Opinion 02/2012 on Facial Recognition in Online and Mobile Services' WP 192, 4 (categorisation system).

[172] See for the parallel problem of applying the GDPR to AI training data M. Oostveen, 'Identifiability and the applicability of data protection to big data' (2016) 6 *International Data Privacy Law* 299, 307; P. Hacker, 'A legal framework for AI training data', *Law, Innovation and Technology* (forthcoming), https://ssrn.com/abstract=3556598, 6 et seq.; see also Clifford (n. 70), para. 309–311.

[173] F. Zuiderveen Borgesius, 'Singling out people without knowing their names' (2016) 32 *Computer Law & Security Review*, 256, 262–265.

[174] See also G. Malgieri and J. Niklas, 'Vulnerable data subjects' (2020) 37 *Computer Law & Security Review*, Art. 105, 415.

[175] F. Zuiderveen Borgesius, 'Personal data processing for behavioural targeting: which legal basis?' (2015) 5 *International Data Privacy Law*, 163, 167 et seq.

4.1.1 | Relevance of the unfairness of a commercial practice

If data processing is intended to specifically identify data subject vulnerabilities in order to target them with unfair commercial practices, it seems close to impossible that the processing can be based on the balancing provision. Data protection law does not, and should not, facilitate processing for illegal means, including violations of the UCPD; in this respect, the interests of the data controller, that is the trader, are not worthy of protection. This is also highlighted by a guidance document published by the German Data Protection Authorities.[176] Hence, in cases in which emotional or cognitive targeting violates the UCPD, there is a very strong indication that the processing for such targeting will lack a legal basis under the GDPR, too, unless valid consent is obtained (in which case the UCPD is typically not breached, either, see above, Section 3.2.1.1.2).

4.1.2 | Relevance of the fairness of a commercial practice

Conversely, one may ask how the legality of a practice, under unfair commercial practice law, impacts the balancing test of the GDPR. It seems safe to say that an assessment of the fairness of the commercial practice under the UCPD does not exhaustively cover all the elements relevant for the balancing test under the GDPR. For example, the UCPD analysis typically does not directly address the type of data processing (e.g. profiling), risks of re-identification, discrimination and other data protection risks mentioned in Recital 75 GDPR. Therefore, the legality of the commercial practice under the UCPD can only be one element to be considered under the balancing test of the GDPR; it cannot conclusively determine its outcome.[177]

Similarly, the fact that a direct marketing campaign complies with Article 13 ePD should be taken as an indication, but not as a conclusive determination, that the balancing test will be resolved in favour of the trader.[178] Emotional and cognitive targeting, however, entails significant data protection risks not addressed under Article 13 ePD; in fact, it comes close to the processing of sensitive health data (Art. 9(1) GDPR).[179] Hence, even if it meets the requirements of Article 13(2) ePD, it cannot necessarily be condoned under Article 6(1)(f) GDPR.[180]

4.2 | The principle of fair data processing, Article 5(1)(a) GDPR

Beyond the balancing test, the unfairness of commercial practices pursuant to the UCPD may also implicate the principle of fair data processing, Article 5(1)(a) GDPR. This is particularly important because the breach of said principle constitutes a serious violation of the GDPR (Art. 85(5)(a) GDPR), even if the data processing itself may be covered by a legal basis under Article 6 GDPR.[181]

The jurisprudence of the CJEU on the principle of consistency of EU law, Article 7 TFEU, suggests an interconnected interpretation of different fairness provisions that, simultaneously, respects the idiosyncrasies of each legal field.[182] Hence, it is submitted that when data processing concerns a practice which violates the UCPD, this should trigger a rebuttable presumption that the principle of fair data processing enshrined in Article 5(1)(a) GDPR was breached, too. To return to the example mentioned above, if data processing is used to identify vulnerable

[176] DSK, 'Orientierungshilfe der Aufsichtsbehörden für Anbieter von Telemedien', 2019, 20.

[177] See CJEU, Case C-109/17, *Bankia*, ECLI:EU:C:2018:735, para. 49.

[178] See also the advice by the German DPAs: Düsseldorfer Kreis, 'Anwendungshinweise der Datenschutzaufsichtsbehörden zur Erhebung, Verarbeitung und Nutzung von personenbezogenen Daten für werbliche Zwecke' (2014) 4; Art. 29 Working Party (n. 112).

[179] See T. Mulder, 'The protection of data concerning health in Europe' (2019) 5 *European Data Protection Law Review*, 209, 216 et seq.

[180] See also the discussion in Section 3.2.1.1.2.

[181] See D. Clifford and J. Ausloos, 'Data protection and the role of fairness' (2018) 37 *Yearbook of European Law*, 130, 134.

[182] See particularly Case C-694/17, *Pillar Securitisation*, ECLI:EU:C:2019:345, paras. 34–35; Case C-45/13, *Kainz*, ECLI:EU:C:2014:7, para. 20; see also CJEU, Case C-508/12, *Vapenik*, ECLI:EU:C:2013:790, para. 25; AG Trstenjak, Opinion, Case C-453/10, *Pereničová and Perenič*, ECLI:EU:C:2011:788, para. 90.

subjects in order to target them with unfair commercial practices, it would seem contradictory to argue that the processing still adheres to a fairness standard relevant to data protection.

This not only corresponds to a growing understanding of said principle as a provision that links data protection law to other areas of EU law dealing with the fairness of transactions;[183] it is also in line with CJEU jurisprudence on the concrete interaction between different fairness provisions. The Court held in *Pereničová and Perenič*, and recently affirmed in *Bankia*, that a contracting practice violating the UCPD does not automatically count as unfair under the Unfair Contractual Terms Directive (UCTD).[184] However, UCPD unfairness must be considered as one of the elements of the UCTD analysis.[185] The motive behind this pointed, yet limited interaction is that there may be circumstances relevant for the UCTD assessment that were not included in the UCPD analysis.

This reasoning can be transposed to the relationship between the UCPD and the GDPR. The CJEU jurisprudence may be interpreted as a risk-specific interaction between different fairness elements in EU law. While different fairness assessments may diverge in their outcomes to the extent that the relevant risks differ, these outcomes must be aligned to the extent that the contemplated risks are similar. This is, indeed, the *ratio* behind Article 3(4) UCPD as well.[186] Hence, UCPD unfairness ought to be considered as a serious factor in the GDPR unfairness assessment. There may, however, be cases where the data protection risks mentioned in Recital 75 GDPR are addressed, by the trader/data controller, in ways relevant for the GDPR but less so for the UCPD, for example if data are strongly secured, pseudonymised, data protection by design is implemented, and the data is not transmitted to third parties. In this case, processing leading to a UCPD violation may still adhere to data protection fairness. In most circumstances, though, a UCPD breach will indicate a violation of Article 5(a) GDPR.

Finally, it should be noted that, in principle, this interaction runs both ways, so that a violation of data protection fairness may also indicate, but does not conclusively determine, the unfairness of the pertinent commercial practice.[187] None of these two fairness assessments commands a logical priority. However, from a methodological perspective, it seems convincing to start with the more concrete standard, in this case the UCPD analysis which contains much more detailed textual guidance for the assessment than the rather vague GDPR clause. Interpreted in this way, different fairness provisions in EU law, while not identical, may complement one another without eliding the peculiarities of their respective legal fields.

5 | MITIGATING ALGORITHMIC MANIPULATION IN THE FUTURE

As the preceding sections have shown, the UCPD and the GDPR provide a flexible and interlocking framework to deal with algorithmic manipulation in market contexts. However, important deficiencies remain. In the context of the UCPD, there is significant legal uncertainty as to its interpretation with respect to mind-reading technologies. Furthermore, opaque micro-targeting processes present important obstacles to effective enforcement.[188] The GDPR, in turn, is not even applicable if the technologies refrain from identifying, or rendering identifiable, individual persons.

Therefore, this last section of the paper proceeds in three steps. First, it discusses to what extent the recent DSA proposal may contribute to overcoming these deficits and reining in algorithmic manipulation (Section 5.1). Second, it provides a critical account of the AIA in this respect (Section 5.2). In a third step, building on this discussion but going beyond the DSA and the AIA, it submits three concrete proposals to mitigate algorithmic manipulation from a forward-looking policy perspective (Section 5.3).

[183] W. Maxwell, 'Principles-based regulation of personal data' (2015) 5 *International Data Privacy Law*, 205, 210; G. Malgieri, 'The concept of fairness in the GDPR', *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 154, 156.

[184] Case C-453/10, *Pereničová and Perenič*, ECLI:EU:C:2012:144, paras. 43–44; Case C-109/17, *Bankia*, ECLI:EU:C:2018:735, para. 49.

[185] Id.

[186] B. Keirsbilck, 'Interaction between consumer protection rules on unfair contract terms and unfair commercial practices' (2013) 50 *Common Market Law Review*, 247, 256; see also AG Trstenjak, Opinion, Case C-453/10, *Pereničová and Perenič*, ECLI:EU:C:2011:788, para. 89–90.

[187] Keirsbilck (n. 186) 257, 260.

[188] Willis (n. 8), Part II.C.

5.1 | The proposed Digital Services Act

The DSA is an ambitious proposal that, if enacted, will significantly reshape the legal environment for online services, particularly when provided by platforms. According to Article 1(5)(h) DSA, the proposal is without prejudice to the UCPD and the GDPR, whose framework and obligations therefore remain untouched. However, the DSA contains two sets of provisions that are of clear relevance for the mitigation of algorithmic manipulation and that, conceptually, can be associated with the two varieties of manipulation identified at the start of this paper. First, transparency provisions concerning advertising may counter manipulation as unawareness (Section 5.1.1). Second, risk management provisions might help to rein in manipulation as significant impairment of rational choice (Section 5.1.2).

5.1.1 | Transparency provisions

The DSA contains two forms of transparency provisions relevant in our context: for advertising (Art. 24 and 30 DSA) and for recommender systems (Article 29 DSA). Concerning the former, Article 24(c) DSA stipulates that customers of online platforms must be provided, in real time, with 'meaningful information about the main parameters used to determine the recipient to whom the advertisement is displayed'. Structurally, this mirrors the heavily contested meaningful information requirement in Art. 13(2)(f) and 15(1)(h) GDPR for automated decision-making,[189] but applies it to the main features responsible for targeted advertisements. Article 30 DSA raises the transparency bar even higher for very large online platforms (>45 million active monthly EU users) who have to create a repository, accessible for one year after the last display of the advertisement, which contains information about the main parameters used to single out particular groups of recipients. Similarly, according to Article 29(1) DSA, the main parameters governing the functioning of recommender systems must be disclosed transparently by very large online platforms. Taken together, these proposals aim to unlock the black box behind recommendations and advertisements, and may divulge hitherto hidden influence parameters that could counter manipulation as unawareness of influence. For example, advertising based on Emotional AI would arguably have to disclose which emotions were responsible for the selection of the advertisement.

However, three significant problems remain. First, the empirical literature on other online-related disclosures, such as privacy policies, suggests that the overwhelming majority of consumers will flatly ignore additional information about targeting or recommendations.[190] Nevertheless, the information provided might be used, primarily, by consumer watchdogs or public authorities to screen for potentially manipulative features, such as cognitive traits or emotions.[191] While it would be illusory to assume that customers would, all of a sudden, become avid and proficient consumers of information, the true added value might lie in the processing of these disclosures by the mentioned information intermediaries and enforcement agencies. Second, however, the main parameters will comprise weaknesses like emotions or bias only if algorithmic models explicitly build on and measure these features (e.g. Emotional AI), but not if the weakness constitutes a latent variable the model unintentionally optimised on. This harks back to the distinction made in the computer science foundations: the disclosure will be useful only in cases of intentional targeting of weaknesses, not of unintentional side effects of the model because, in these cases, the "main parameters" will likely be variables that serve as, often non-obvious, proxies for weaknesses. It should be noted, though, that the potentially negative reputation effects of disclosure in cases of intentional targeting could slow the uptake of explicit emotion-based advertisement.

Third, and perhaps most importantly, the disclosure obligations only apply to (very large) online platforms. Service providers other than online platforms are supposed to be included in codes of conduct on online advertising

[189] A. Selbst and J. Powles, 'Meaningful information and the right to explanation' (2017) 7.4 *International Data Privacy La*, 233.

[190] See n. 74.

[191] See also European Commission (n. 12), 5.

(Art. 36 DSA); however, their enactment and compliance with them are entirely voluntary. This constitutes a genuine loophole: Manipulative targeting can arise in direct marketing by smaller traders, as evidenced by the possibility to embed emotion recognition in any apps (Microsoft, Amazon), or even in the context of brick-and-mortar stores (eyeQ). The restricted scope of the DSA fails to cover these important extensions.

5.1.2 | Risk management provisions

The second set of relevant provisions is constituted by the rules on risk management (Art. 26–28 DSA), which may set incentives for very large online platforms to tackle manipulation by impairment of rational choice. These platforms must conduct an assessment of specific systemic risks (Art. 26 DSA), implement suitable, risk-based mitigation strategies (Art. 27 DSA) and open themselves to an annual independent audit (Art. 28 DSA). These provisions clearly mirror the risk-based rules on data protection by design and impact assessments in Articles 25 and 35–36 GDPR.

The proposals should be welcomed as risk-based compliance tools which unfold an effect even if individual customers do not exercise their rights or ignore the disclosures. For the context of manipulation proper, however, it seems problematic that risk management is geared exclusively toward strictly defined systemic risks, listed in Article 26(1) DSA: dissemination of illegal content; negative effects on certain fundamental rights; and intentional manipulation of the service. It seems unlikely that the targeting of cognitive or emotional weaknesses, or of vulnerable consumers more generally, is covered by these points.

First, intentional manipulation (Art. 26(1)(c) DSA) concerns manipulation *of* the platform (e.g. by hackers), not manipulation of consumers *by* the platform (see Recital 57 DSA). Second, protection from manipulation in markets does not form part of the fundamental rights listed in Article 26(1)(b) DSA, unless it coincides with discrimination based on protected attributes in the sense of Article 21 of the Charter. Third, Article 26(1)(a) DSA arguably comes closest to a monitoring and auditing obligation with respect to manipulation. It concerns the dissemination of illegal content through the digital services. Illegal content, in turn, comprises inter alia activities related to the infringement of consumer protection law (Recital 12, Art. 2(g) DSA), and hence also of the UCPD. However, the assessment obligation only covers the *dissemination* of illegal content, which denotes the act of hosting products or pieces of information that in turn violate EU law (including consumer protection law), but not to the violation of consumer protection law *through the services themselves* (see Recital 57, Art. 2(i) DSA).[192] This restriction likely follows from the focus of the DSA on platforms conceived as largely neutral intermediaries who merely host (Art. 2(h) DSA) and thereby make available misinformation or counterfeited products (Recital 57).

Therefore, in my view, Article 26(1) DSA should be amended to include, for example as Article 26(1)(d), an assessment of the extent to which the services of the platforms themselves repeatedly and systematically, in ways intentional or unintentional, violate consumer protection law. Arguably, this may also take on a dimension of "systemic risk" because of the wide spread of harm across consumers in cases of repeated infringements. Mirroring Article 35 GDPR, this would be tantamount to a 'consumer protection impact assessment'. It would be in keeping with the pronounced position afforded to consumer protection in the Charter (Art. 38) and the specific value of consumer autonomy for free and fair market decisions.[193] On the downside, even such an addition would be restricted to very large online platforms, giving rise to the same limited scope noted with respect to the transparency regime of the DSA.

[192] Martini et al. (n. 38), 73.

[193] See also below, n. 209 and accompanying text.

5.2 | The proposed Artificial Intelligence Act

Most recently, the Commission in April 2021 unveiled the AIA, which also seeks to tackle the problem of algorithmic manipulation.[194] Recital 15 AIA acknowledges that AI systems may exploit vulnerabilities of actors they are dealing with. Importantly, however, the AIA defers largely to the UCPD by making only limited additions to the existing rules on non-manipulation. These novel proposals fall into two categories: rules on prohibited AI practices (Art. 5(1) AIA), and transparency provisions for, inter alia, emotion recognition systems (Art. 52 AIA).

5.2.1 | Unfair AI practices

The relevant prohibited actions fill important gaps left by the UCPD: they cover cases in which an AI system inflicts physical or psychological harm, irrespective of the presence of a "commercial practice". Such an activity related to the 'promotion, sale or supply of the product to consumers' is a prerequisite for the application of the UCPD (Art. 5, 2(d) UCPD)—but not of the AIA. That Act thus covers cases in which the AI system "misbehaves" during its deployment, unrelated to an offer or sale. While the protection of transactional decisions, resulting in purely economic losses, is therefore left to the UCPD, the AIA complements the UCPD by expanding its applicability from unfair commercial practices to "unfair AI practices" more generally,[195] to the extent that they cause physical or psychological harm.

With respect to its concrete prerequisites, however, the two prohibitions relevant for manipulation in Article 5 (1)(a) and (b) AIA operate on familiar territory: not only do they both require a material distortion of decision-making, they also quite exactly match the two dimensions of manipulation discussed in this paper. Article 5(1)(a) AIA outlaws subliminal influence of AI systems resulting in physical or psychological harm, epitomising the hidden-influence prong of manipulation. A hypothetical example could be subconscious cues doctored into videos by an AI, spurring violent behaviour. Article 5(1)(b) AIA, in turn, guards against the exploitation of vulnerabilities of a specific group of persons due to their age, physical or mental disability, if that exploitation, via a material distortion, leads to physical or psychological harm. That provision represents the case of significant impairment of rational choice, closely mirroring Article 5(3) UCPD. An example invoked by the Commission might be 'toys using voice assistance encouraging dangerous behaviour of minors'.[196] Indeed, such actions would likely not constitute a commercial practice in the sense of the UCPD, but nevertheless could target the reduced decision-making capacities of children or disabled persons.

While the rules therefore represent a meaningful addition to the scope of protection afforded by the UCPD, they also have significant shortcomings and partially inherit the interpretational difficulties of the UCPD concerning manipulation. First, the wording "in order to" in Article 5(1)(a) and (b) AIA and Recital 16 AIA suggests an intentionality requirement concerning material distortion, which would be even stronger than foreseeability demanded by Article 5(3) UCPD. As discussed above,[197] intentionality will often be difficult to prove in AI settings, particularly if the behaviour was acquired independently of explicit human code specifications. Second, while the mentioning of vulnerabilities due to age, physical or mental disability does point to specific reductions of rational decision-making capacity, the strict enumeration of the protected groups again leaves the question of bounded rationality, or vulnerabilities stemming from yet other trait combinations, unresolved. With "credulity" not transplanted from Article 5 (3) UCPD, the only potentially fitting category in Article 5(1)(b) AIA for significant bounded rationality would be "mental disability". However, according to the CJEU jurisprudence in non-discrimination law, disability implies 'physical, mental or psychological impairments', which must generally be clinically relevant and long-term, and which

[194] See European Commission, Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence, Press Release (21 April 2021), https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682, under heading 'unacceptable risk'.

[195] See J.-P. Schneider and Ch. Wendehorst, 'Response to the public consultation on the White Paper: On Artificial Intelligence', Working Paper 2020, 9.

[196] See European Commission (n. 194).

[197] See the end of Section 3.2.1.1.2.

generally hinder 'the participation of the person concerned in professional life'.[198] Virtually no cognitive biases would seem to reach that threshold. Hence, given the narrowly defined groups in the AIA and the term "disability", it seems unlikely that the mere presence of even strong biases could qualify as a relevant vulnerability under Article 5(1) (b) AIA.

Moreover, the analysis of the structure of the personalised yardstick under Article 5 UCPD has shown that the specific subgroups mentioned in Article 5(2) UCPD play an important role in covering cases of vulnerable persons beyond the strictly enumerated groups of Article 5(3) UCPD. However, while the AIA mirrors the latter provision, it fails to provide an analogon to the former. This opens a significant gap as the detection and potential exploitation, via AI-based analysis, of new vulnerabilities unrelated to age, physical or mental disability (e.g. cognitive or emotional weaknesses) is arguably a key risk in the deployment of AI systems, as the preceding discussion has sought to show.

5.2.2 | Transparency for Emotional AI

The AIA does include, in Article 52(2), a transparency provision obliging users of AI systems to disclose the use of certain "emotion recognition systems" operating through the processing of biometric data (Art. 3(34) and (33) AIA) to those exposed to them. Arguably, this seeks to empower consumers to avoid such tools, and to guard against hidden influence by emotional analysis. However, two major shortcomings must be noted here, too. First, the applicability of that provision again turns on the contested question of the processing of *personal* data (see Art. 3(34) in conjunction with Art. 3(33) AIA). Second, reputational effects aside, transparency alone is rarely an effective remedy, as a host of empirical studies show (see Section 5.3.1).

The AIA therefore contains meaningful rules complementing the UCPD to confront manipulation by AI systems, but fails to overcome some of the gaps and shortcomings identified in the existing legal framework. More encompassing rules may be needed—a discussion to which the final part of the paper now turns.

5.3 | Beyond the DSA and the AIA: Three concrete proposals

As the preceding sections have shown, the transparency, counter-manipulation and risk management provisions in the DSA and AIA represent steps in the right direction. Nevertheless, looking beyond the DSA is warranted first and foremost because of its restricted focus on platforms, but also because of the substantive shortcomings just discussed. Similarly, the proposals in the AIA do neatly complement the UCPD in cases of mere physical or psychological harm; they are also limited, however, to very specific cases (subliminal influence), harms to narrowly defined groups (e.g. children, disabled people) and specific remedies (transparency for Emotional AI based on personal data). Given this restricted scope and the conceptual difficulties noted, both proposals make valuable contributions, but do not fully resolve the threat of algorithmic manipulation. Therefore, the final part of the paper makes three proposals toward a further mitigation of that risk.

Following the conceptual distinction between manipulation as hidden influence and manipulation as the impairment of rational choice, it seems sensible to create, on the one hand, possibilities for the reflection of influence (5.3.1), but also to consider stricter regulation and mandates mitigating the prevention of rational choice ex ante (5.3.2). Finally, beyond rational choice, mind-reading technologies in particular implicate privacy norms in novel ways, which raises the possibility of privacy infringements even in the absence of identifiability in the sense of the GDPR (5.3.3). In spirit, these suggestions complement the transparency and risk management proposals of the DSA and AIA, but modify and expand them significantly beyond platforms and high-risk AI systems while focusing on the issue of manipulation.

[198] CJEU, Case C-13/05, *Chacón Navas*, ECLI:EU:C:2006:456, para. 43.

5.3.1 | Creating spaces for reflection: Information and withdrawal rights

The first strategy would seek to better empower consumers and endow them with enhanced possibilities to reflect on choices made under algorithmic influence. This may counteract manipulation in the sense of hidden influence (manipulation as unawareness). In legal terms, such avenues for reflection may particularly take the form of transparency obligations, as in the DSA and AIA, or withdrawal rights. For example, the transparency obligations concerning the main parameters of advertising could be expanded, beyond online platforms covered in the DSA, to traders in general if advertisements are targeted to specific subgroups. Furthermore, the reference to "personal data" should urgently be eliminated from the definition of "emotion recognition systems" in the AIA.[199] Salient information on the use of Emotional AI should be provided irrespective of whether the analysed data relates to an identifiable data subject or not. As mentioned, however, it seems doubtful whether more information alone will be conducive to better consumer decision-making, given rampant information overload and rational ignorance documented by a host of empirical studies.[200]

There is, however, a case for withdrawal rights if manipulation leads to *temporarily limited* distorted mental states. Such rights already exist in many cases of manipulative targeting if it is employed in the context of a distance or off-premise contract, Article 9 Consumer Rights Directive (CRD). Beyond this, withdrawal rights should be granted in case of emotional or cognitive targeting for on-premises contracts in retail (brick-and-mortar) stores, too.[201] In this sense, Article 9 CRD would have to be expanded. Such rights offer a space of reflection which can be helpful if the underlying bias or weakness leading to the transaction is transitory. This is often the case with hot emotional states.[202] However, it seems much less likely concerning entrenched cognitive biases.[203] Barring extensive feedback and learning,[204] such biases and preference changes will often not magically disappear through the passage of time.

5.3.2 | Shaping technology: Non-manipulation by design

When algorithmic manipulation operates not through hidden influence, but (also) through the impairment of rational choice, withdrawal rights will likely not suffice. Vulnerable consumers may not necessarily exercise these rights, for various cognitive reasons (e.g. endowment effect; status quo bias; procrastination; or simple convenience).[205] Hence, it seems worthwhile to discuss strategies for mitigating such influence ex ante, rather than attempting to correct distorted outcomes ex post via traditional UCPD enforcement or withdrawal rights.[206] In this vein, an important strategy is what will here be termed "non-manipulation by design".[207] While the GDPR, in its Article 25, now features a duty of data protection by design, an analogous provision is lacking in the UCPD. While such a duty could be included in the Digital Services Act, particularly in an extension to the risk assessment in Article 26(1) DSA (see Section 5.1.2), an amendment to the UCPD would be preferable as it is not restricted to large online platforms. Similarly, non-manipulation by design could form part of the risk management owed under Article 9 AIA. Here again, however, the provision is limited in scope, in this case to high-risk AI systems, which systems generating targeted offers do not, at the moment, belong to.[208] Clearly, though, the proposals in Article 26 DSA and Article 9 AIA do show that there is a window of opportunity for similarly spirited, but more broadly applicable rules in the UCPD.

[199] See Art. 3(34) in conjunction with Art. 3(33) AIA.

[200] See the references in n. 74.

[201] See also G. Wagner and H. Eidenmüller, 'Down by algorithms: siphoning rents, exploiting biases, and shaping preferences' (2019) 86 *University of Chicago Law Review*, 581, 597.

[202] Lerner (n. 96), 811.

[203] R. Larrick, 'Debiasing', in D. Koehler and N. Harvey (eds.), *Handbook on Judgment and Decision Making* (Blackwell, 2004) 316, 318.

[204] Larrick (n. 203).

[205] See, e.g. J. Luzak, 'To withdraw or not to withdraw?' (2014) 37 *Journal of Consumer Policy*, 91, 100 et seq.

[206] See also A. Renda, 'Making the digital economy "fit for Europe"' (2021) *European Law Journal* (forthcoming).

[207] See also, in a similar vein, Willis (n. 15), Part C.III. ('fair marketing by design').

[208] Except in cases of remote biometric identification, see Annex III to the AIA.

Including such a provision in the UCPD, restricted in its applicability to the digital sector (algorithmic models facilitating commercial practices), seems important for two main reasons. First, Article 25 GDPR does not exhaustively cover cases of manipulative targeting. It is geared toward data protection, not non-manipulation, by design; and it does not even apply to cases not involving the processing of personal data, such as potentially the emotional or cognitive targeting of otherwise unidentified passers-by. The same holds true for Article 35 GDPR. Second, however, non-manipulation by design does seem as important as data protection by design. While data protection is grounded in the fundamental rights enshrined in Article 8 of the Charter and Article 16 TFEU, non-manipulation is rooted in the protection of individual autonomy and freedom of choice. They form the basis of an undistorted internal market and sovereign participation in economic activity, thus connecting to primary law provisions on the internal market (Art. 26 TFEU), consumer protection (Art. 38 of the Charter) and, ultimately, human dignity (Art. 1 of the Charter).[209] In fact, as mentioned, the ultimate goal of data protection on many accounts is precisely the safeguarding of individual autonomy.[210]

Admittedly, the interpretation of Article 5 UCPD offered above already sets incentives to audit algorithmic models for manipulation. This interpretation of the general clause remains, however, contested, and should be bolstered by an explicit duty. Hence, it is suggested that a novel obligation should be included in the UCPD which compels traders who employ algorithmic models (processing personal or non-personal data) to implement strategies of non-manipulation by design in a way that is adequate for the risk level of the respective model, data, and application. In this sense, it would mirror, and complement, the risk-based approach in Articles 25 and 35 GDPR. Irrespective of the applicability of the GDPR and the interpretation of the general clause of the UCPD, such a duty would compel the trader to proactively audit the model and correct its functioning in case it differentiates between different consumer groups in ways that are likely to infringe the UCPD. This implies that the often vague provisions of the UCPD need to be rendered more concrete, for example by the delineation of cases typically violating the UCPD.

While technical tools will likely not perfectly match context-oriented UCPD analysis,[211] they may at least single out clear violations and thereby mitigate particularly worrisome forms of manipulation. In technical terms, such an auditing obligation could build on a vast literature in computer science on algorithmic fairness, which has developed a range of tools to screen and correct for discrimination by algorithmic models, that is, for differential impact of the models on groups protected by anti-discrimination law.[212] These very same techniques could be employed to detect and mitigate the differential treatment of consumers because of different degrees of rationality or emotions.[213] Importantly, these techniques contain flexible tools which allow traders to fine-tune the degree toward which the distribution of outcomes is approximated between different groups.[214]

If, for example, traders are concerned about unintentional exploitation of negative emotional states, they could force the model to treat consumers exhibiting positive and negative emotions the same, on average. A similar rule could be implemented concerning high and low degrees of rationality. This would mean that individual differences between members of the respective emotional or cognitive groups are still allowed, but that, for example, boundedly rational consumers are not on average offered worse deals than highly rational ones. If, on the other hand, traders feel that there is a legitimate reason for differentiating between different mental states, for example because certain products are helpful only for sad and not for happy consumers, they could adjust the correction parameter so that a

[209] See Micklitz (n. 132); A. von Bogdandy et al., 'Reverse Solange' (2012) 49.2 *Common Market Law Review*, 489, 495; E. Eriksen, 'Why a Charter of Fundamental Human Rights in the EU?' (2003) 16.3 *Ratio Juris*, 352, 359, 364; O. Schachter, 'Human dignity as a normative concept' (1983) 77.4 *American Journal of International Law*, 848, 851.

[210] See n. 115.

[211] See, e.g. C. Benzmüller, D. Fuenmayor and B. Lomfeld, 'Encoding legal balancing: automating an abstract ethico-legal value ontology in preference logic', arXiv preprint arXiv:2006.12789 (2020).

[212] J. Dunkelau and M. Leuschel, 'Fairness-aware machine learning', Working Paper, 2019; D. Pessach and E. Shmueli, 'Algorithmic Fairness', *arXiv preprint arXiv:2001.09784* (2020).

[213] See P. Thomas et al., 'Preventing undesirable behavior of intelligent machines' (2019) 366.6468 *Science*, 999.

[214] See, e.g. M. Zehlike et al., 'Matching code and law: achieving algorithmic fairness with optimal transport' (2020) 34 *Data Mining and Knowledge Discovery*, 163.

differential treatment of these groups is technically allowed again.[215] Of course, these technical decisions would not prejudice the finding in law that, ultimately, such differential treatment constitutes a breach of the UCPD. It does, however, offer a way for the trader to adjust the technical strategy to different situations in which different degrees of non-manipulation by design might be appropriate. As seen, the UCPD analysis is highly context-sensitive, and does allow for certain types of emotional and cognitive targeting in specific situations, particularly in the case of consent.

The key problem, however, remains the allocation of targeted consumers to the different groups defined by vulnerability (emotions, degrees of rationality). While automated mental state detection increasingly facilitates such an analysis from a technical perspective, it raises data protection concerns even if such analysis is undertaken not to exploit consumers but rather to mitigate manipulation. Such data could constitute sensitive personal data in the sense of Article 9 GDPR.[216] This raises a true conflict between the UCPD and the GDPR, as the former demands the prevention of manipulation by traits which the latter forbids to be analysed.[217] A possible solution could be based on Article 9(2)(g) GDPR, according to which the processing of sensitive data is allowed if necessitated by substantial public interest and sufficient safeguards are installed. One could argue that the prevention of manipulation, as defined in the UCPD, constitutes such a significant public goal and that, therefore, the processing is legal under Article 9 GDPR if it is strictly limited to the auditing and internal correction of the model. Given these tensions, the envisioned novel exception to Article 9 GDPR contained in Article 10(5) AIA, covering precisely the use of sensitive data for non-discrimination purposes in AI development, must be explicitly welcomed. It should, however, be expanded to include data management for the sake of non-manipulation, too, under the safeguards already contained in Article 10(5) AIA.

Until the AIA is enacted, future work will have to delineate more clearly where exactly the boundaries of the processing of sensitive data in the context of audits in the public interest are currently located. Meanwhile, as this section has shown, non-manipulation by design seems a worthwhile strategy and is technically increasingly possible, but needs to be complemented by a rigorous data protection framework and, possibly, supervision by data protection authorities. At the very least, if codes of conduct for online advertising are drawn up pursuant to Article 36 DSA, they should contain a commitment to non-manipulation by design. Mirroring an instrument from the GDPR toolbox (Art. 40 GDPR), they could become an additional vehicle for raising industry awareness of and compliance with this important issue.

5.3.3 | Privacy beyond identifiability: Toward a subjective right to invisibility

With respect to mind-reading technologies more particularly, it should be noted that they are bound to become mainstream, with Microsoft and Amazon offering simple plug-ins for any kinds of apps.[218] Therefore, it seems crucial to recognise that the development and deployment of such technologies, such as automated mental state analysis and Emotional AI, raise important regulatory challenges even if: (i) market transactions are not distorted in the sense of the UCPD; and (ii) the GDPR as well as Article 52(2) AIA do not apply because individuals are not sufficiently identifiable. Even in these cases, for example concerning the analysis and non-manipulative emotional targeting of passers-by in public space, there is a case for privacy protection beyond the tools of EU law discussed above.[219] In fact, mental analysis on otherwise unidentifiable subjects presents a rare occasion where data protection and privacy

[215] Id., at 188 et seq.

[216] See C. Jasserand, 'Legal nature of biometric data: From "generic" personal data to sensitive data', (2016) 2 *European Data Protection Law Review*, 297.

[217] See also M. Veale and R. Binns, 'Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data' (2017) 4(2) *Big Data & Society*, Art. 2,053,951,717,743,530.

[218] See n. 58–59.

[219] McStay (n. 3).

law part ways.[220] Concerning the comparable case of body scanners in airports, scholars have argued that, even in the absence of data protection, the fundamental right to privacy applies.[221]

Similarly, in the case of automated emotion or cognitive analysis, privacy protection suggests considering a strict regulation for the use of such technologies. There are two reasons why privacy is implicated in important ways even if individuals are not identified. First, it should be recognised that mental states, such as emotions or cognitive styles, are part of the *forum internum* of human beings. They constitute essentially intimate information that has remained, for the most part, inherently private so far.[222] While privacy law traditionally recognises three different spheres defined by context and location—spheres of intimacy, of private, and of public space[223]—automated mental state analysis arguably adds a fourth dimension to this triad by making available, in a systematic, automated and effective way, information on the *forum internum*: the mind. People may find such analysis highly intrusive even if they are not identified, and even—or particularly so—if it occurs in public space. Second, such analysis clearly and notably changes the relationship of analysed subjects to public and private spaces in which such technology is deployed.[224] Like video cameras, automated mental state recognition re-codes spaces of relative anonymity as precincts of potential surveillance and intrusion. Such observation is bound to change behaviour, as the literature on chilling effects shows.[225]

Therefore, it seems important to go beyond transparency (envisioned in Article 52 AIA), as it may be difficult to justify the use of automated mental state analysis from a privacy perspective even in the presence of disclosure, at least for purposes such as marketing. Admittedly, as the empirical study already mentioned shows, around 10% of participants did support the personalisation of ads and offers by way of Emotional AI.[226] While a complete ban may therefore be controversial, at the very least a right to a privacy-preserving option should be granted, both online and in public spaces. This would mean that traders may only engage in automated mental state analysis if privacy subjects have validly consented to such practices—like in the case of cookies, Article 5(3) ePD. Importantly, in the absence of such consent, it could not be justified by a "privacy balancing test" analogous to Article 6(1)(f) GDPR, or by mere disclosure. Consent being a problematic concept, it will likely have to be primarily exercised via (semi-)automatic privacy assistants able to digest larger amounts of information than humans.[227] Normatively, such technologically mediated consent may build on the provision concerning browser settings in the proposal for the ePrivacy Regulation.[228] Conversely, making such consent a necessary condition would allow those with high privacy preferences to "become invisible", with respect to automated mental state analysis, by withholding consent. This seems essential in our increasingly networked surroundings, in which participation in social life should not depend on finding ways to physically avoid emotional analysis. It would also tie in with the proposal in Article 29(1) DSA, according to which customers must have the option to use recommender systems without profiling—which would create a similar right to non-profiled recommendations.

All of this highlights that there is a need for new regulation, at the EU level, on sensing technology operating on the human body without necessarily identifying individuals, going beyond Article 52 AIA. As the current ePrivacy Directive already covers both personal and non-personal data, the new ePrivacy Regulation could offer an appropriate forum for this demarche—if it is eventually enacted. However, the latest compromise proposal does not even include a provision on tracking walls anymore.[229] It therefore

[220] Clifford (n. 70), para. 312.

[221] R. Gellert and S. Gutwirth, 'The legal construction of privacy and data protection', (2013) 29 *Computer Law & Security Review*, 522, 527; see also O. Mironenko, 'Body scanners versus privacy and data protection' (2011) 27 *Computer Law & Security Review*, 232, 237 et seq., 240.

[222] Gellert and Gutwirth (n. 221), 527.

[223] See, e.g. H. Nissenbaum, 'Protecting privacy in an information age: The problem of privacy in public' (1988) *Law and Philosophy*, 559, 567 et seq.

[224] McStay (n. 65), 6.

[225] See, e.g. Y. Hermstrüwer and S. Dickert, 'Sharing is daring: An experiment on consent, chilling effects and a salient privacy nudge', (2017) 51 *International Review of Law and Economics*, 38.

[226] McStay (n. 3), 9.

[227] See, e.g. A. Das et al., 'Personalised privacy assistants for the internet of things: providing users with notice and choice' (2018) 17 *IEEE Pervasive Computing* 35; see also S. Wachter, 'Normative challenges of identification in the internet of things' (2018) 34 *Computer Law & Security Review*, 436, 445.

[228] Council of the EU, Presidency, Interinstitutional File: 2017/0003(COD), Nov. 4, 2020, Art. 4a(2).

[229] Id., Art. 10.

seems unlikely that comprehensive rules on tracking, recognition of body parts and mental states will form part of the final version. Hence, an academic and societal debate is all the more crucial. Ultimately, the EU may need to enact, beyond the GDPR, a General *Privacy* Protection Regulation, to deal with technologies that, while not identifying individuals, still threaten to redefine privacy as we know it in public and private spaces alike.

6 | CONCLUSION

This paper has sought to both provide an analysis of mind-reading technologies such as Emotional AI under current and future EU law, and to discuss, based on the broader issue of algorithmic manipulation, interactions between unfair commercial practice, data protection and privacy law. In doing so, it makes four contributions to the existing literature. First, it clarifies the concept of manipulation in digital contexts by differentiating between manipulation as unawareness and manipulation as the significant impairment of rational choice. For the latter case, three particularly important consumer weaknesses are discussed that are normatively and legally relevant: strong emotions or biases (intensity variety); mutually reinforcing emotions or biases (combination variety); and vulnerabilities in complex or unanticipated decision-making environments (complexity variety). Furthermore, the paper harnesses insights from computer science and economics to differentiate between two types of manipulative targeting: intentional targeting of said weaknesses through mind-reading technologies, and unintentional targeting as a result of other types of machine learning optimisation.

Second, building on these distinctions, it shows that the three mentioned, manipulative varieties of cognitive and emotional targeting may run afoul of the general clause in Article 5 UCPD, particularly if personalised targeting simultaneously triggers a personalised yardstick of consumer behaviour which focuses on vulnerable groups. In this context, privacy and data protection norms suggest that the *intentional* use of mind-reading technologies will often violate professional diligence under the UCPD, particularly when the GDPR is breached. However, *unintentional* targeting arguably only violates professional diligence if the trader fails to pro-actively audit the model for manipulative features. Third, vulnerable data subjects are addressed in data protection law, too. In the context of algorithmic manipulation, core norms of the GDPR exhibit a pronounced, yet risk-specific interaction with the UCPD. More precisely, a violation of the UCPD indicates, but does not conclusively determine, a breach of the principle of fair data processing (Article 5(1)(a) GDPR) and a preponderance of data subjects' interests in the balancing test (Article 6(1)(f) GDPR).

Finally, while the interlocking frameworks of the UCPD and the GDPR may go some way toward mitigating algorithmic manipulation, important gaps remain, particularly when the unidentifiability of analysed subjects renders the GDPR moot. The DSA and the AIA represent important steps into the right direction with the proposed rules on advertising and emotion recognition transparency, systemic risk management, and unfair AI practices. However, both Acts also exhibit various limitations of scope and conceptual shortcomings. Hence, an update of EU law should not be restricted to the DSA and the AIA, but must go beyond them. To alleviate the challenges raised by consumer manipulation and mind-reading technologies, three concrete proposals are made: first, a withdrawal right should be created in the CRD for on-premises contracts concluded in the wake of cognitive or emotional targeting. Second, a duty of "non-manipulation by design" ought to be included in the UCPD. Fairness tools developed in computer science may facilitate its implementation. Third, privacy norms should be activated to guard against mind-reading technologies even if they do not identify subjects in the sense of the GDPR or of Article 52(2) AIA. Ultimately, this suggests that, in our increasingly networked environments, the GDPR may need to be complemented by a General Privacy Protection Regulation to safeguard autonomy and privacy beyond the sometimes overly narrow precinct of "personal data".

ORCID

Philipp Hacker https://orcid.org/0000-0002-3006-895X

How to cite this article: Philipp Hacker. Manipulation by algorithms. Exploring the triangle of unfair commercial practice, data protection, and privacy law. *Eur Law J*. 2023;29(1-2):142–175. https://doi.org/10.1111/eulj.12389