How Artificial Intelligence Constrains the Human Experience

ANA VALENZUELA, STEFANO PUNTONI, DONNA HOFFMAN, NOAH CASTELO, JULIAN DE FREITAS, BERKELEY DIETVORST, CHRISTIAN HILDEBRAND, YOUNG EUN HUH, ROBERT MEYER, MIRIAM E. SWEENEY, SANAZ TALAIFAR, GEOFF TOMAINO, AND KLAUS WERTENBROCH

ABSTRACT Artificial intelligence (AI) and related technologies are transforming many consumption activities, powering breakthroughs that expand the human experience by enhancing human capabilities, performance, and creativity. While this explains the consumer enthusiasm and rapid adoption of these technologies, AI systems can also have the opposite effect: reducing and constraining the range of experiences that are available to consumers. This article examines the mechanisms through which AI can constrain the human experience, considering individual, interpersonal, and societal processes. Our analysis uncovers a complex interplay between the advantages of AI and its inadvertent negative repercussions, which potentially restrict human autonomy, self-identity, relational dynamics, and social behavior. In this article, we propose three different mechanisms at the core of these constraining forces: parametric reductionism, agency transference, and regulated expression. Our exploration of these mechanisms highlights the risks connected to system design and points to questions and implications for future researchers and policymakers.

Many consumption decisions and experiences are digitally mediated. As a consequence, consumer behavior is increasingly the joint product of human psychology and ubiquitous algorithms (e.g., Melumad et al. 2020; Sangers et al. 2024). The coming of age of large language models (LLMs) is further accelerating the dissemination and impact of artificial intelligence (AI). AI holds the promise of improving the life of consumers everywhere, in ways small and large. At the same time, the deployment of this technology is not without risks. The societal discourse on the potential risks of AI tends to focus on issues of discrimination and privacy, or on distant "existential" risks (e.g., the possibility of human extinction or an irreversible global catastrophe). Our focus is different. Following the work of consumer researchers who have started to identify psychological tensions in the consumer experience of AI (Puntoni et al. 2021), we contribute to this nascent literature by exploring how AI can reduce the range of people's expression and choices, as well as the opportunities available to them for personal development. In other words,

Ana Valenzuela (corresponding author: ana.valenzuela@baruch.cuny.edu) is a professor at ESADE-Ramon Llul, Barcelona, Spain, and Baruch College, City University of New York (CUNY), New York, NY 10010, USA. Stefano Puntoni (puntoni@wharton.upenn.edu) is a professor at Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA. Donna Hoffman (dlhoffman@email.gwu.edu) is a professor at George Washington University, Washington, DC 20052, USA. Noah Castelo (ncastelo@ualberta.ca) is a professor at University of Alberta, Edmonton AB T6G 2R3, Canada. Julian De Freitas (jdefreias@hbs.edu) is a professor at Harvard Business School, Boston, MA 02163, USA. Berkeley Dietvorst (berkeley.dietvorst@chicagobooth.edu) is a professor at Booth School of Business, University of Chicago, Chicago, IL 60637, USA. Christian Hildebrand (christian.hildebrand@unisg.ch) is a professor at University of St. Gallen, St. Gallen, Switzerland. Young Eun Huh (younghuh@kaist.ac.kr) is a professor at Korea Advanced Institute of Science and Technology, Daejeon, South Korea. Robert Meyer (meyerr@wharton.upenn.edu) is a professor at Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA. Miriam E. Sweeney (mesweeney1@ua.edu) is a professor at University of Alabama, Tuscaloosa, AL 35487, USA. Sanaz Talaifar (s.talaifar@imperial.ac.uk) is a professor at Imperial College London, South Kensington, London SW7 2AZ, UK. Geoff Tomaino (geoffrey.tomaino@ufl.edu) is a professor at University of Florida, Gainesville, FL 32611, USA. Klaus Wertenbroch (klaus.wertenbroch@insead.edu) is a professor at INSEAD, Fontainebleau, France. The authors thank John Deighton for his very clear guidance as Guest Editor and the full review team for their helpful comments. They would also like to thank Ziv Carmon and Paulo Albuquerque, and the whole team at INSEAD, for organizing the 2023 Choice Symposium, which represented the genesis of this article. *JACR* is grateful to John Deighton, who graciously agreed to serve as Guest Editor for this article.

Issue Editors: Stefano Puntoni and Klaus Wertenbroch

Published online May 15, 2024.

Journal of the Association for Consumer Research, volume 9, number 3, July 2024. Published by The University of Chicago Press for the Association for Consumer Research. https://doi.org/10.1086/730709

our intent is to delineate how consumers' possibility spaces are increasingly shaped by algorithms. This approach is aligned with industry calls to understand how "customers face an array of new devices with which to interact with firms, fundamentally altering the purchase experience" (Marketing Science Institute 2018, 4).

The term "artificial intelligence" was coined by John McCarthy and his colleagues in 1955 as "the science and engineering of making intelligent machines," with the emphasis on machines' capabilities to learn, at least, in part, as humans do (McCarthy et al. 2006, 1). Today, AI represents a general-purpose technology (Brynjolfsson and McAfee 2014) with the potential to affect every industry (Agrawal, Gans, and Goldfarb 2018) and transform the economy (Furman and Seamans 2019). Against this backdrop, a rapidly expanding scholarly literature documents the business potential of AI and Big Data (e.g., Agrawal et al. 2018). For consumers, AI holds the promise of helping them make more efficient and effective decisions, save time, and enjoy better products and experiences that more accurately match their preferences. To illustrate, recommender systems like Netflix's and personalized products like Spotify's "made for you" playlists improve the quality of our consumption. Generative AI applications like ChatGPT, which can help people write tedious reports in minutes, help us become more productive and save time (Noy and Zhang 2023). Robo advisers and other decision support systems can help increase the quality of decisions and improve outcomes (Hildebrand and Bergner 2021).

While these gains are real and significant, it is critical to equally consider the potentially pernicious outcomes that could emerge from our increasing reliance on AI for making decisions that affect consumers and citizens, and to assess their potential risks, which seem intrinsic in a powerful technology that remains, in most cases, a "black box" to consumers and researchers alike (De Freitas et al. 2023). In other words, the delegation of a growing number of tasks to AI-systems raises serious concerns about its potential for altering core elements of human agency and motivation in unintended ways. To date, few investigations have focused on the potential for smart technologies like AI to limit, restrict, or reduce human experience (Hoffman and Novak 2018).

Our focus in this article is to examine the potential of AI to constrain human experience. By "human experience," we refer to people's overall perceptions, feelings, and behaviors as they engage in interactions with technology (AI), particularly in their role as consumers and throughout the customer journey (following Brakus, Schmitt, and Zarantonello 2009 and Lemon and Verhoef 2016). We intend to understand how the use of AI in these interactions may impose costs– understood as reductions or limitations of people's experience in consumer contexts. AI can constrain human experience directly, as by feeding users a limited array of options that prevents them from exploring new options and perspectives. These limitations reflect the active, agentic role of AI systems in shaping the nature and extent of human-technology interactions, which may involve removing components, limiting the functionalities of these components, or impeding interactions among them.

In tandem with directly constraining human experience, AI can constrain human experience indirectly, as when people use AI in ways that limit their own engagement with the experience (Hoffman and Novak 2018). The motivation for such self-restriction can vary, but can include concerns over privacy, autonomy, ethical considerations, or accommodating an AI's limitations (e.g., the inability of a voice assistant to understand long, complex expressions). Through self-restriction, people directly limit the scope and functionality of their AI interactions, effectively reducing its potential impact and capabilities. For example, a consumer might decide to turn off the voice recognition feature on her smart speaker due to concerns about eavesdropping or data privacy or may prefer to manually override settings on devices like thermostats rather than rely on AI optimization or personalization features. Self-restriction often empowers consumers to shape their technological environments according to their own taste; but self-restrictions of the human experience are characterized by consumers modifying their behavior, language, or interaction patterns to defend against an AI system or conform to its operational parameters. This adaptation may result in the consumer using a reduced set of their own capacities or altering their interaction styles to accommodate the AI's performance. These reductions can lead to feelings of being diminished or "less than" because the consumer must interact in ways that do not fully represent their potential or preferred modes of interaction. For example, consumers might simplify their language or alter their questions to fit the recognition and response capabilities of voice-activated AI assistants or conform to recommendations or choices presented by AI. Over time, such self-reduction could narrow the consumer's exposure to different options, reducing their agency in decision making.

We propose three different mechanisms that may be at the core of these constraining forces: agency transference, parametric reductionism, and regulated expression. We define them with the intent of mapping consumers' interactions with AI throughout their decision journey: handing over agency to algorithms, which reduce consumers to a limited set of parameters, and potentially constrain how consumers express themselves and communicate. Below, we discuss the key mechanisms that help explain and predict AI's potential to constrain the consumer experience. The article concludes with a set of future research questions and implications for practitioners, including AI developers and policy makers. Aside from consumer experiences, we show how AI may impose constraints on agency, skills, equality, dignity, and diversity. It therefore demands urgent and multi-disciplinary scrutiny to ensure that it empowers, rather than imperils, human experience.

AGENCY TRANSFERENCE

Agency transference relates to AI's ability to limit one's personal agency, as agency is transferred from humans to algorithms (also see De Freitas et al. 2023). Personal agency has been defined as "the sense that I am the one who is causing or generating an action" (Gallagher 2000, 15) or as having the power to influence one's own actions and circumstances (Bandura 2006). By their very deployment, recommendation algorithms leverage past consumer behavior to selectively curate content (André et al. 2018; Wertenbroch et al. 2020). As a result, consumers become less likely to be exposed to options and content that does not correspond to their revealed (first-order) preferences, depriving them of opportunities to change these preferences and choose something else (enacting their second-order preferences; Wertenbroch et al. 2020). Because AI systems can shape our environment—what we see and the opportunities available to us (Grafanaki 2017)—they can subtly manipulate decision-making trajectories in ways that ultimately constrain self-determination (Bhattacharjee et al. 2014). Within this larger theme of agency being "transferred" to AI and thereby constraining human agency, we discuss how AI favors a loss of serendipity in accessing options and enables both cognitive and emotional de-skilling.

The Loss of Serendipity

One of the joys of life is serendipity, making fortunate discoveries by accident. Serendipitous discoveries are ones that are relevant to us, despite being unexpected (Kotkov, Wang, and Veijalainen 2016). For instance, you might pop into a bookstore and discover an obscure book yet find it highly relevant to a paper you have been writing. Or you might talk to a colleague and find that, surprisingly, they are a fountain of knowledge on the local art scene just when you were in search of a great show.

However, recommendation algorithms often limit our experience of serendipity. They do so because most of these systems are set up to feed content based on our past behaviors (André et al. 2018; Wertenbroch et al. 2020). Consequently, they reinforce those behaviors and create inertia that limits exploration and change (Talaifar and Lowery 2023), constraining human agency. For instance, sequentially viewing content causes one to consume more content from the same category, explaining the common experience of "going down a rabbit hole" (Woolley and Sharif 2022).

Moreover, past work finds that recommendation algorithms may also limit aggregate serendipity. Because similar groups of consumers simultaneously receive similar content recommendations, this may encourage similar consumers to homogenize even more (Fleder and Hosanagar 2009; Lee and Hosanagar 2019). Furthermore, as exploration becomes highly correlated among users, this leads to a popularity bias, in which the market share for products that are already popular increases at the expense of others; while recommendation algorithms may increase absolute sales for niche items, these gains may be proportionally small compared to that for more popular items (Lee and Hosanagar 2019).

For these reasons, some have advocated complementing the current "Skinnerian" approach to serendipity in recommendation algorithms with a better psychological understanding of consumers (Morewedge et al. 2023). For instance, just because consumers might click on entertainment videos more often than on educational content does not mean they wish to receive recommendations only for entertainment content. To better balance consumers' wants against their shoulds (Milkman, Rogers, and Bazerman 2009), platforms could employ techniques like analyzing longer consumption time windows or allowing consumers to personalize the degree to which the algorithm recommends previously consumed categories versus serendipitous content (Morewedge et al. 2023). In fact, past research has shown that when recommender systems (e.g., Yelp) suggest novel and serendipitous restaurants and locations, consumers are more likely to follow up on those recommendations (Smets et al. 2022).

De-skilling

As new technologies emerge and mature, they often take over certain tasks or skills that were previously performed by humans. Off-loading tasks or skills to technology can mean that those skills atrophy over time, a phenomenon known as deskilling (Wood 1987). For example, the Industrial Revolution brought about the invention of textile machinery like the spinning jenny, power loom, and cotton gin. These machines replaced artisan weavers and spinners, transforming the textile industry from a skilled craft to a factory-based system where workers operated machinery without needing the skills of the traditional textile artisan (Berg and Hudson 1992).

The emergence of ChatGPT and other forms of generative AI have revived the concern that AI could contribute to a new wave of de-skilling among knowledge workers. For example, generative AI models can now outperform most humans at creative idea generation (Guzik, Byrge, and Gilde 2023; Koivisto and Grassini 2023). Relying on such models can increase the quantity and quality of knowledge workers' output (Noy and Zhang 2023; Dell'Acqua et al. 2024). If workers begin to rely on AI to perform parts of their jobs, such as writing or brainstorming, might they be foregoing opportunities to practice and strengthen those skills? Could this in turn cause those skills to weaken, relative to workers who do not rely on AI assistance? Similar effects have been documented in the context of memory: people seem to offload the need to remember information to the Internet, such that memory becomes worse when relying on the Internet to find information (Ward 2021; Fisher, Smiley, and Grillo 2022).

De-skilling might also be a concern outside of work or task-oriented contexts; by impacting emotional and social skills. For instance, recent research suggested that, as consumers trade off human contact with more reliance on AI-powered technology, there may be an increased variance in social adjustment, impacting emotional intelligence, particularly among our youth (e.g., Beranuy et al. 2009; Ralph and Nunez 2021). There seems to be early support for the possibility that relying on algorithms to manage interactions on social media platforms is connected to a decline in our ability to understand and navigate complex social cues. Furthermore, many people are already interacting extensively with AI-based synthetic companions such as Replika (De Freitas et al. 2024). Due to the "black box" nature of the algorithms, it is impossible to predict in advance how these conversations will unfold. De Freitas et al. (2024) provide evidence that companion AIs often fail to recognize and to respond appropriately to signs of distress, which calls into question the safety of chatbots for individuals with mental health issues. Furthermore, the current regulatory structure is not set up to address these risks (De Freitas and Cohen 2024).

Finally, we posit that there may be additional, broader societal implications of de-skilling due to algorithmic decision making. When a skill is entirely lost in a population, certain experiences may no longer be accessible. For instance, the loss of skill to operate and repair older devices, such as media players for obsolete standards, can make certain information inaccessible, as happened a few years ago in Britain. A massive project of historical documentation carried out in the 80s by the BBC required years of work to be again accessible because the laserdiscs on which the information had been stored could no longer be read (Harford 2023). The danger of automation-caused de-skilling has long been a salient concern in the contexts of complex systems or occupations, such as airline pilots (e.g., Carr 2014). In the same vein, it is possible that if we grow entirely reliant on AI for certain tasks, the resilience of communities could be undermined if those skills suddenly become valuable again (e.g., because of systemic failure of our digital infrastructure). For these reasons, it is essential to maintain a balance between using algorithms to aid decision making and preserving the human capacity to make decisions and carry out tasks independently.

In sum, AI has the potential to constrain human agency by reducing the range of possible choices and actions that consumers might consider, thus limiting opportunities to explore, learn, and change established behavioral patterns. Moreover, reliance on AI may also lead users—consumers and workers alike—to unlearn valuable cognitive and practical skills, not only as individuals but also as societies.

PARAMETRIC REDUCTIONISM

By their very nature, algorithms are reductionists. They need to translate human behavior, identity, preferences, and attributes into a smaller set of independent, computationally readable variables, parameters, and formulae (Hildebrand 2019). In this way, AI systems tend to objectify individuals and communities, reducing or compressing their unique characteristics and cultural contexts. This process may lead to misalignment, that is, the misrepresentation or under-representation of people's actual preferences and interests when they are translated into an algorithmic formula. Objectification and misalignment are the two mechanisms we discuss next.

Objectification

AI functions through parameterization and categorization, reducing the complexities of human beings into a set of quantifiable metrics, classifications, and risk scores to sort, assess, and predict behavior.. Thus, this process is limited in its ability to fully account for the unique characteristics and circumstances of an individual, resulting in the objectification of that individual. While Haslam and Stratemeyer (2016) define dehumanization as "the act of perceiving or treating people as if they are less than fully human," Fiske (2009) specifies that there is one specific form of dehumanization, which might be termed objectification, which views people as automatons (tools, robots, machines). We believe this definition best matches this first layer of the parametric reductionism mechanism defined as "algorithms distilling individuals into data metrics."

In this manner, even AI designed for beneficial purposes can propagate a subtle yet pernicious form of human objectification that persists within technical systems. These systems often overlook how characteristics used to judge individuals will systematically correlate with other aspects of an individual's background. A notable example was Amazon's hiring algorithm, which exhibited a bias against women due to its high positive weighting of characteristics traditionally associated with men (Dastin 2018). Similar patterns of discrimination have been uncovered in AI applications within judicial (Larsson 2019) and educational domains (Engler 2021). Furthermore, there might be biases directly built into the algorithm itself as it aims to represent people's preference structure (Morewedge et al. 2023).

This gap between the real human individual and the AI-coded representation of them can result in unintended harm or consequences, including amplifying systematic inequalities and causing indirect social spillover effects. Objectification can amplify (and obscure) systemic inequalities by reducing individuals to group characteristics. That is, individuals of certain backgrounds are often systematically given different opportunities as a result of algorithmic objectification, affecting crucial outcomes such as loan decisions (Bertrand and Weill 2021) and pricing (Chapdelaine 2020). Consumers thus (rightly) question AI's capacity to appreciate their unique traits and circumstances, and show reluctance to utilize it in important contexts, such as healthcare (e.g., Longoni, Bonezzi, and Morewedge 2019) or financial services (Yalcin et al. 2022).

Additionally, the objectification introduced by AI can have serious indirect social consequences. For instance, objectifying interactions with AIs can lead to spillover effects whereby we objectify other people in real life, such as through treating them more instrumentally (e.g., Onur et al. 2023). Furthermore, consumers view objectifying AI systems as being less capable of assessing interpersonal skills (Castelo, Bos, and Lehmann 2019). As a result, the objectifying nature of AIs can influence perceptions of those selected by AI in hiring processes as well as candidates' own behavior during AI-driven interviews (Cheong, Huh, and Puntoni 2023). As another example, Granulo et al. (2024) find that the use of AI in management tasks can increase feelings of objectification which, in turn, reduces prosocial motivation and behavior. While a great deal of work has been done on these misperceptions, given the rise of commercialized LLMs (e.g., ChatGPT), it is increasingly urgent to further understand the ways in which AIs objectify humans and the consequences thereof.

Misalignment

Secondly, because the algorithms powering today's AI systems are designed to rely on a reductionist representation of user interests, any discrepancy between this simplified representation and the complexity of actual human preferences risks a misalignment between AI recommendations and what users truly want or would choose for themselves. If so, the more decisions are outsourced to AI systems, the more that these systems may yield misaligned choices. One potential opportunity for misalignment arises in determining which outcome should be maximized on behalf of the user. For instance, algorithms curating content on social media platforms typically aim to maximize user engagement (Kim 2017), yet users of these platforms might prefer to optimize for a different outcome. This discord can stem from divergent incentives, such as a company's profit motives versus a consumer's pursuit of their own welfare (Castelo et al. 2023). However, misalignment can occur even with aligned incentives. For instance, people's preferences may be too idiosyncratic, complicated or unobservable for a reductionist algorithm to learn efficiently (e.g., only wanting taco delivery on rainy Tuesdays after a drink). Additionally, people's own behavior, from which algorithms learn, may not align well with their goals (e.g., many drinkers wish to quit drinking). Thus, if consumers receive recommendations only based on their past behavior, they may be shown more temptations, at odds with their goals (Carmon et al. 2019; Morewedge et al. 2023).

Beyond misalignment in outcome preference, discrepancies can also arise in how outcomes are optimized. This is typically governed by an objective function that defines the relative desirability of various outcomes. The choice of objective function is consequential, as different functions lead to different algorithm outputs, potentially misaligning with user interests. For example, in the domain of prediction, algorithms typically use objective functions with increasing (e.g., root mean square error [RMSE]) or constant (e.g., mean absolute error [MAE]) sensitivity to error; however, people often exhibit decreasing sensitivity to prediction error (Dietvorst and Bharti 2020). In practical terms, this could mean that algorithms often prioritize avoiding large errors when making predictions, while users may prefer them to pursue near-perfect predictions even at the risk of large errors (Dietvorst 2023). In the domain of investing, for example, misaligned objective functions could lead robo-advisors to go for more or less risk than a client desires.

The suitability of maximization itself as a decision-making strategy is debatable in some domains. Many algorithms are designed to select the output that maximizes the desired outcome as prescribed by their objective function. However, maximization may not always align with human preferences. For instance, consumers often have ethical concerns about using maximization as a strategy for making morally relevant trade-offs. As a result, they may object to any algorithm implementation based on maximization in these domains, even when developers may have worked to align it with people's preferences as much as possible (Dietvorst and Bartels 2022). Relying solely on maximization could also diminish diversification or willingness to explore new options (Talaifar and Lowery 2023). Finally, if using models built to maximize outcomes encourages people to aim for the best possible options instead of satisficing, research by Schwartz et al. (2002) suggests that this could make consumers less happy with their choices and potentially overall.

In sum, AI systems have the power to limit our experiences by reducing people to rigid functions, parameters, and scores, thus failing to capture nuanced human complexities and possibly perpetuating subtle dehumanization. Such oversimplification risks misrepresenting people's true preferences, potentially leading to misguided decisions. Moreover, this reductionist approach might discourage AI use in sensitive areas or deteriorate human-AI interactions, as individuals feel that AI cannot truly grasp their uniqueness.

REGULATED EXPRESSION

AI systems require large amounts of information from users to operate and, thus, tend to require significant self-disclosure. Such self-disclosure has the potential to be harmful. In this respect, a long-standing result in the study of human-computer interaction is the so-called privacy paradox: while consumers often express worry over the privacy of the information they provide online, they are also often willing to openly share their most intimate thoughts and feelings in social media posts, responses to online surveys, and chatbots when they feel they are receiving something of value in return (e.g., Norberg et al. 2007; Joinson et al. 2010; Acquisti et al. 2015; Tomaino, Wertenbroch, and Walters 2023). Adding to that, the modality of interaction imposed by AI systems can have a profound influence on both how individuals communicate with AI and with other humans. When interacting with algorithms, human expression thus has the potential to become both over- or under-regulated—that is, controlled, rethought, or otherwise altered away from natural expression. Such regulated expression potentially overrides individuals' authentic communication style, expression of self, and the extent to which they self-disclose.

Vulnerability from Self-Disclosure

As noted by Barth and de Jong (2017), among the numerous explanations offered for the privacy paradox, the most parsimonious is that it reflects a bias in the intuitive cost-benefit calculations people undertake when deciding whether to provide information online. Specifically, as real as the potential risks associated with providing personal information online may be, they are also temporally remote and abstract. In contrast, the benefits of providing the information—be it to gain access or make a purchase—are immediate and concrete. As a result, consumers typically give less weight to privacy risks—or ignore them altogether—when making online disclosure decisions (e.g., Acquisti 2004; Flender, Peters, and Müller 2012).

The costs of breaches of privacy to individuals can be substantial, and courts have struggled in their efforts to keep pace with firms' abilities to leverage personal information in ways that can impose harm—an ability that is accelerating with recent advances in AI (Solove and Keats Citron 2018; Keats Citron and Solove 2022). One of the most well-known cases where courts successfully intervened was in the 2016 Facebook-Cambridge Analytica (CA) scandal, where the FTC fined Facebook $5 billion for providing personality profiles of 87 million Facebook users to CA—information that was used by the staffs of Ted Cruz and Donald Trump to advance their 2016 presidential campaigns, a use to which few Facebook users likely consented (Confessore 2018). More commonly, the harm caused by unintentional disclosures ranges from simple intrusiveness (personal tracking information used to propagate unsolicited text messages/ads) to aiding with different forms of financial theft. Between 2019 and 2023, for example, the FTC reported a 59% increase in reports of identity theft, a trend widely attributed to advances in AI that allow work-arounds to privacy software and encryption technologies (Caporal 2024).

Recent work in this domain has attempted to identify the design factors in computer environments that lead to unintended self-disclosure. Much of this work is inspired by the ideas of Clifford Nass (e.g., Nass and Moon 2000), who argued that people often unconsciously respond to computers as if they are people or social actors. As an example, in a series of experimental studies, Nass and Moon (2000) found that people were more likely to disclose information about themselves to a computer if the computer did so first—a reciprocity effect like that observed in human interactions (e.g., Cosby 1973). More recently, the advent of AI-powered chatbots has spawned a growing literature exploring how the linguistic properties of computer speech (both textual and audio) affects the willingness of users to self-disclose (e.g., Cox and Ooi 2022; Choi and Zhou 2023). The core finding of this work is that the efficacy of chatbots in eliciting self-disclosure depends not just on how human-like the interaction is (e.g., Bickmore and Cassell 2001) but also on its suitability for the conversational context.

Finally, the design of the device itself has been found to affect self-disclosure. As an example, in an analysis of almost 20,000 "call to action" web ads administered either on a smartphone or desktop computer, Melumad and Meyer (2020), found that consumers were more willing to volunteer personal information (such as debt and health information) when contacted on their smartphones, and were more self-disclosing in social media posts created on their smartphones.

Constrained Interaction

As discussed, the modality of interaction with AI systems also influences how individuals communicate. For example, the inherent limitations of current AI systems to convert spoken human language into machine-readable streams of data has caused an unwanted increase in so-called failed "intent-matching" by the AI system (i.e., failure to correctly identify the objective of a user request and provide an accurate answer), which makes users switch to more simplified language expressions to interact with these systems (Hildebrand et al. 2020). Qualitative research with Amazon Alexa users revealed that they often engage in a more simplified form of communication compared to a human interlocutor (Ammari et al. 2019). Recent research by Hildebrand, Hoffman, and Novak (2023) provides large-scale evidence using a corpus of actual user-voice assistant interactions that the default modality of users is a more direct, imperative language style compared to human-to-human communication, using fewer personal pronouns (such as "I" or "me"), shorter sentence lengths with fewer words, and an overall less polite language style (not saying "please" when asking the device to perform a task). Therefore, "conversational" AI systems implicitly disincentivize linguistic diversity by requiring users to employ more simplistic, reduced forms of language.

The way these AI systems harvest and process data from their users also exerts a critical and often unwanted influence on human expression. For example, voice interactions seem to elicit a significantly heightened sensitivity around privacy (Pitardi and Marriott 2021; Sweeney and Davis 2021). This sensitivity is well justified. Voice data is unique in capturing paralinguistic features about objective aspects of a person (such as their age, gender, or even country of origin), momentary states (such as classifying a user's experienced emotion), or even the early onset of physical and mental health issues (such as identifying depressive symptoms or COVID-19 from voice samples; Hildebrand et al. 2020; Zierau et al. 2022; Busquet, Efthymiou, and Hildebrand 2024). As users both consciously and subconsciously alter their speech to control the image they present to the AI (Vimalkumar et al. 2021; Melzner, Bonezzi, and Meyvis 2023), these "natural" language interfaces can inadvertently cause a less authentic self-presentation (Moorthy and Vu 2015). Such usage dynamics suggest that the way conversational AI is designed to collect and utilize data of their users not only affects the user's willingness to disclose information but also the quality and authenticity of the communication itself.

Constrained Self-Expression

Finally, research suggests that people feel less able to express their true opinions and unique attributes online than offline. For instance, people report both expressing their personality traits less and engaging in less self-disclosure in online compared to offline contexts (Blumer and Döring 2012; Bunker and Kwan 2021). The extent to which people regulate their self-expression online should depend on features of these digital environments, where AI is an increasingly ubiquitous feature (Talaifar and Lowery 2023). Indeed, people seem to understand that algorithms in these environments are tailored to influence their identities in specific, potentially pernicious, ways (Bhandari and Bimo 2022). For instance, people perceive that AI-systems are meant to amplify some aspects of their identity, while filtering out or suppressing other, more marginalized, parts (Simpson and Semaan 2021). This perception may be in fact accurate (Aral and Walker 2012; Chakraborty et al. 2017; Feldman 2020).

Overall, there are a variety of direct and indirect reasons why AI may constrain self-expression. Indirectly, AI functions by relying on vast quantities of data on users' everyday behaviors. Collecting such large quantities of data could make people feel that their actions and personal communications online are not private or anonymous. In support of this idea, meta-analytic and experimental evidence suggests that perceptions of anonymity and privacy are key predictors of online self-expression (Joinson et al. 2010; Wu and Atkin 2018; Clark-Gordon et al. 2019). More directly, AI may constrain self-expression because, as alluded to above, it may evoke certain aspects of a person's identity to the exclusion of others (Soh, Talaifar, and Harari 2024). For instance, a woman's algorithmically mediated Instagram Explore feed may show beauty and fashion content based on the accounts that she has followed. The saliency of this content may prime aspects of her gender identity rather than aspects of her professional identity, even if it is her professional identity what she is wishing to cultivate and express. In support of these ideas, preliminary experimental research suggests that gender-stereotypical recommendations decrease women's self-reported masculinity, leadership ability, and self-confidence (French 2018).

Constrained self-expression could have several detrimental consequences. Research shows that inauthentic self-expression on social media predicts lower wellbeing and greater mental health symptoms (Bailey et al. 2020; Bunker et al. 2024). Moreover, when people do not express their true opinions, it creates a false perception of actual opinion in the broader public (Noelle-Neumann 1974). Perhaps even more troublingly, self-censoring may undermine people's true opinions themselves, since findings support that people who self-censor their opinions subsequently attribute less importance to them (Talaifar and Ashokkumar 2023).

POLICY IMPLICATIONS AND FUTURE DIRECTIONS FOR RESEARCH

AI is beginning to live up to its promises to deliver enormous benefits to users, from medical applications, to boosting productivity, to amplifying human creativity. At the same time, experts, journalists, and politicians alike warn of potential risks involved in developing and employing AI. Relevant stakeholders include not only individual users (e.g., in their roles as consumers, patients, citizens, etc.) and the private and public organizations that deploy AI, but they also include regulators who are being called upon to protect users from developments of AI. At the time of this writing, the European Union has just adopted the AI Act on March 13, 2024: the world's first legislation to regulate AI. Aside from requiring the use of general-purpose AI (e.g., LLMs such as ChatGPT) be transparently disclosed to consumers, the EU AI Act views certain uses of AI as violating fundamental individual rights and, thus, as unacceptable, including behavioral manipulation, emotion recognition in work or educational contexts, social scoring, and inferences of sensitive personal data such as religious beliefs or sexual orientation. Similarly alerted by the societal debate, both the US Congress and the White House have begun deliberating risks of AI development and deployment.

Aside from these regulatory developments, much is being written about the potential societal risks of AI, from spreading misinformation to manipulating user perceptions, attitudes, and behavior, and from military applications going rogue to artificial general intelligence developing its own objectives and wresting control from humans (e.g. Bao et al. 2022). The autonomous nature of AI and inability to perfectly align advanced systems with human values and context also introduces new risks ranging from privacy violations to embedded biases that could constrain human equality and dignity. For example, AI's tendency to amplify systemic biases threatens to exacerbate inequalities under a veneer of technical objectivity (Qureshi 2023). AI systems can also reinforce problematic social norms and power dynamics through their design. For instance, AI systems may representationally embed harmful class, race, and gender hierarchies in their interface design and/or as part of the user experience (Sweeney 2021). Finally, privacy is not only a necessary condition for individual freedom but also for human dignity (Whitman 2004). Yet privacy cannot be guaranteed in an environment where algorithms are ubiquitous and where online footprints are visible and stored.

In this article, however, we focused not on these societal risks but on personal welfare and psychological consequences of using AI for individual users such as consumers, patients, workers, and so forth. Overdependence on AI could gradually limit exploration of alternative viewpoints or de-skill populations and diminish independent thinking (agency transference). AI may reduce complex human experiences and identities into simplified representations, formulas, or data points in a way that leaves people feeling objectified and unsatisfied (parametric reductionism). The systematic gathering of data necessary for AI infrastructures may create vulnerability by encouraging over disclosure of private information, and—at the same time—privacy concerns may lead people to self-impose limitations on authentic self-presentation and expression (regulated expression). Each of these three major mechanisms and their subpathways may restrict human experience by encouraging self-imposed limits and shrinking the scope of possibilities that are available within an individual's experiential opportunity space. Figure 1 offers a visual summary of these key ideas.

Public Policy Implications

We believe that this experiential perspective is of crucial importance for public policy. The current regulation tends to focus on relatively objective criteria, such as the reinforcement of bias or limits to price competition. We argue that this perspective needs to be complemented with an understanding of how the technology is used in context and of how features of the technology can sometimes interact with human psychology to promote undesirable outcomes. Here we highlight three broad, potential psychological outcomes connected with human-AI interactions that carry major public policy implications.

First, when AI-driven algorithms only reinforce past choices, users' perceived autonomy is curtailed (Wertenbroch, Vosgerau, and Bruyneel 2008; André et al. 2018; Wertenbroch et al. 2020). This psychological outcome, which can arise due to each of the three major mechanisms we have covered, has important public policy implications since the sense of a lack of autonomy has been found to be detrimental to consumer well-being (Langer and Rodin 1976). Interventions exist that could be implemented in the interaction with AI to give back a sense of control to the individual, most of them design-driven (see section below). Furthermore, one might argue that consumers themselves can, and therefore ought to, decide whether and how much of their private data they want to trade in return for the benefits they receive from AI and platform applications, limiting any risk of privacy intrusions by AI. Research has shown that consumers make disclosure choices under great uncertainty about the consequences of disclosure and their own preferences concerning these consequences (see, e.g., Acquisti, Brandimarte, and Loewenstein [2015] on the privacy paradox). Yet none of these findings necessarily undermine the argument that consumers make disclosure decisions rationally, even under conditions of uncertainty. With that said, recent research by Tomaino et al. (2023) presents experimental evidence to suggest that consumers' privacy choices violate transitivity, as they are willing to sell their private data for more when the currency is money than when it is bartering goods. Since transitivity is a fundamental axiom of rational-choice theory, this result suggest that consumers' choices for what to disclose to the AI algorithms that use these data may not be rational, entailing the risk or undesirable, aversive or even harmful outcomes.

Second, when discussing vulnerability from self-disclosure (regulated expression mechanism), we noted the fundamental

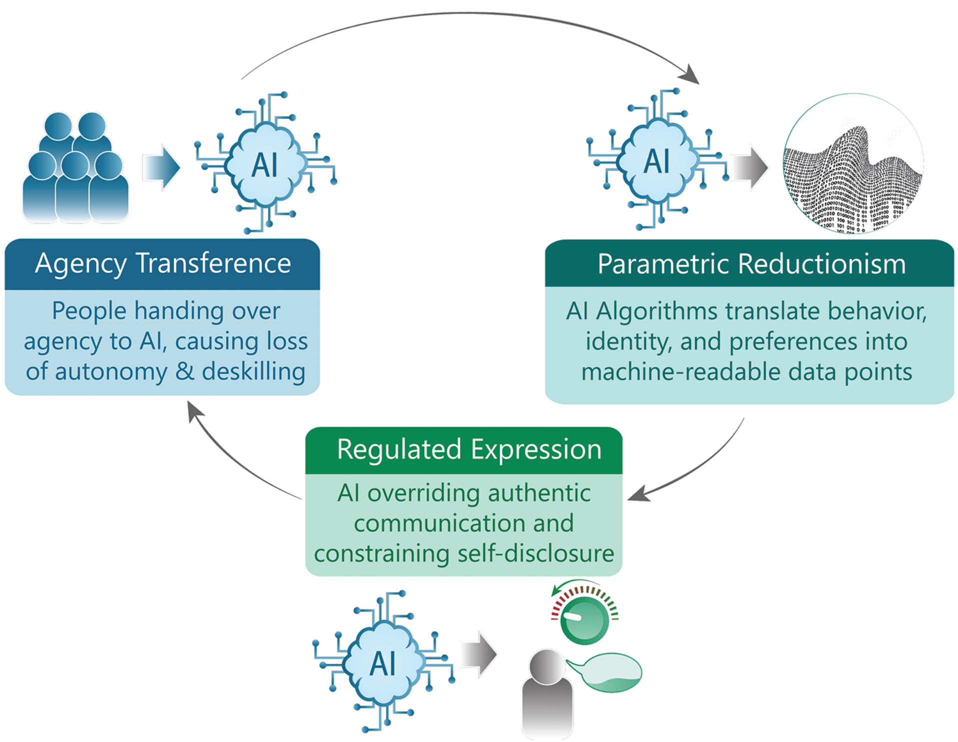

Figure 1. The three mechanisms constraining the human experience.

human tendency to anthropomorphize. For the first time in history, humans are confronted with nonhuman entities of similar or, potentially even greater, intelligence when attempting to accomplish a particular task. Research suggests that consumers will interact with such AI in much the same way they interact with other humans, thus putting them at risk of exhibiting not only the same cognitive and social biases as in human interactions but also of not accounting for AI's superior abilities to extract and exploit information (Nass and Moon 2000; Novak and Hoffman 2019; Puntoni et al. 2021). On top of that, LLMs can now largely pass the Turing test (Turing 1950), such that users cannot distinguish at better than chance levels whether they are dealing with a human or with an AI when the AI is not identified as such. If users react differently to a given piece of communication depending on whether it comes from a human or an AI, then an accurate understanding of whether they are interacting with one or the other is important to avoid deception and fraud. At a minimum, and as required by the EU AI Act, AI must be transparent to avoid misleading users.

Yet, research suggests that even when consumers know that they are interacting with an object or machine, not with a human, they still treat an anthropomorphized object as if it was human—as manifested in consumer behaviors, evaluations, and beliefs. This is because anthropomorphism is easily triggered by humanlike cues such as a voice, smile, spelling errors (Bluvstein et al. 2024), or social responses like politeness and gender stereotypes (e.g., Nass, Steuer, and Tauber 1994). By increasing perceived humanness, such cues can improve marketing-relevant outcomes like product evaluations (Aggarwal and McGill 2007) and trust in the technology (Waytz, Heafner, and Epley 2014). Overall, the most important outcome of this psychological process is that consumers need to be educated on how to approach AI with more deliberative mindsets and to use new metaphors that are not as human-centric—even when anthropomorphizing is intuitive and natural.

Third, related to the major mechanism of parametric reductionism, if individual consumers view AI as misaligned with human interests, then this may also breed broader outrage and discontent, carrying potential downstream implications for societies (De Freitas and Cikara 2021). In a recent study, a representative sample of Americans was surveyed once a year during the last 3 years regarding their hopes and fears for AI—in the two most recent years, the most prominent belief overall was that AI would "make powerful people more powerful" and "allow certain groups to dominate others" (Castelo and De Freitas, forthcoming). This fear was even more prevalent than other widely discussed beliefs like "AI will make it harder for humans to find work" and "AI will make life easier." Regardless of whether AI systems would, in fact, increase economic inequality, the mere belief that it will can have negative psychosocial and political consequences (Willis et al. 2022). Specifically, it may have causal effects on consumers' willingness to rely on AI systems themselves, even at the expense of missing out on meaningful benefits.

AI Design Implications

Apart from highlighting the need for tech-centric education, many of these concerns may be addressed by improving how AI is designed by drawing direct inspiration from the psychological phenomena we have covered here (see also Carmon et al. 2019). We offer three examples of design-based innovations here:

1. *Build-in exploratory behavior*. As discussed, AI's ability to predict consumer preferences from vast amounts of data with ever greater predictive validity and, consequently, to expose consumers to the options they have chosen in the past—be they news content, commercial offers, streaming content, or social media contacts—may prevent consumers from exploring a wider range of alternatives. Designers of AI-based recommender systems are increasingly recognizing how implementing serendipity into their algorithms can help consumers discover novel and relevant products, avoiding consumers' boredom and enhancing their satisfaction (Kotkov et al. 2016). Some platforms such as Spotify already recognize the importance of exploratory behavior by offering content that differs from past consumption (Datta, Knox, and Bronnenberg 2018).
2. *Automatic exposure to alternative stimuli* (even if labeled as such). Relatedly, much research is concerned about echo chambers (Grafanaki 2017; Brugnoli et al. 2019) and the shaping and reinforcement of opinions by recommendation systems (Adomavicius et al. 2013). Design choices driven by a motivation to limit these echo chambers should foster the ability to change one's opinions, which requires exposure to alternative stimuli, perspectives, opinions, choice options, and so forth (e.g., Bakshy et al. 2015).
3. *Normative design personalization*. Consumers often have little ability to express preferences for

Table 1. Future Directions for Research

Mechanism	Research questions
Agency transference:	
Preference and identity exploration	• How can recommender systems increase serendipity without deleterious effects on the accuracy of the recommendation? • Do consumers know whether their preferences around serendipity are aligned with the recommendations they receive? Does explicit preference elicitation help if consumers often struggle to articulate these preferences? • What are the effects of more serendipitous algorithms on customer trust, satisfaction, and retention? • What motivates consumers to engage in behaviors such as erasing their browser history, creating fake social media accounts, or even disconnecting from platforms entirely? • In developing algorithms that increase the chance of serendipity, how can firms account for heterogeneous preferences across consumers or even across a single consumer's lifetime with the firm? What businesses or business models can benefit from the development of more serendipitous algorithms?
De-skilling	• How can interface design enable customers to maintain a balance between using algorithms to aid decision making and preserving the human capacity to make decisions independently? • What elements can be introduced in AI systems to provide users with opportunities to develop and maintain valued skills?
Parametric reductionism:	
Objectification	• What are the "spillover effects" of interacting with AI on human-to-human interactions and evaluations? • What are the consequences of a productivity mindset for human social cognition and behavior? • How can people psychologically bolster themselves against feelings of objectification? • Why do people develop parasocial relationships with AIs and how that impacts their own self-image?
Misalignment with user interests	• How can AI systems be designed to maximize outcomes users care about by using better ways to solicit user input? • How can AI systems allow users more influence over the decision of which tasks they want to control vs. delegate? And what explains their preference for one or the other? • What tensions emerge between customization vs. standardization? How can AI balance customization for each person with the need for some standardization in the name of efficiency? • What are the downstream effects of better aligning objectives for trust, satisfaction, and other outcomes? • Would explicability and control over algorithm objectives improve trust and satisfaction? • What are the unintended consequences of AI deployment for human behavior? And what are the effects on individuals and society?
Regulated expression:	
Disclosure vulnerability	• As our knowledge of how to design computer interfaces that maximize consumer self-disclosure grows, what are the consequences for the "privacy paradox"? • What are the broader consumer safety and welfare risks that emerge with the blurring of the distinction between human-to-human and human-to-computer interactions?
Interaction modality	• To which extent does the simplified language structure required by AI impact how and what we think, such as expressing more complex or creative ideas? • How does the more assertive and direct language consumers use when interacting with chatbots shape our social cognition and attitudes towards authority over time? • How does this change in language use shape our ability to empathize with others? More generally, do regular interactions with conversational AI enhance or impair human emotional intelligence and empathy? • If consumers across the globe are required to communicate with AI systems in simplified conversational styles, what are the ramifications for linguistic diversity and cross-cultural differences? • What are the long-term effects of continued AI interactions on language development at the individual level and even on the evolution of language more broadly?
Identity constraints	• What factors shape individual susceptibility to identity shaping by digital environments? • How does the use of AI systems in one context (e.g. social media) influence behaviors and norms when interacting with humans in other contexts?

Table 1. (*Continued*)

Mechanism	Research questions
	• How do features and affordances of digital environments shape identity development over time? • In what ways can algorithms and predictive personalization be designed to allow for identity exploration rather than constraint? • What forms of self-expression and exploration become more limited vs. expanded through AI systems?
Inequality amplification	• What are the drivers of individuals' beliefs about inequality in society, and what is the perceived role of AI? Do these beliefs influence consumers' willingness to rely on AI? • Beyond lay people beliefs, how does AI affect inequality and how does that shape the range and quality of human experience?
Anthropomorphic biasing	• How do interface design choices in AI systems shape user perceptions and behaviors? • What factors increase or decrease compliance when receiving instructions from an AI system? How does anthropomorphizing AI systems impact compliance compared to instructions from a human or non-human interface? • Can AI systems help mitigate or exacerbate existing discrimination and injustice in society? • Can vocal vulnerability/anthropomorphism be ethically leveraged to increase trust and compliance outcomes when interacting with AI? There are many implications and guidelines regarding emotional manipulation of users.

how they would like the AI system to serve them. Just as an example, Google Maps does not ask users if they want to minimize the likelihood of a badly wrong estimated time of arrival (ETA) or maximize the likelihood of a very accurate ETA. More concerning, TikTok's algorithm does not ask users what skills they wish to develop when serving them an endless stream of video content. In some cases, there may be considerable technical barriers to implementation, but many AI design choices are just that: choices. We should aim to avoid technological determinism and consider how AI systems may embed self-representational preferences in interface design and/or as part of the user experience. Such changes might benefit consumers and firms alike. For instance, one study finds consumers fed news content in line with their ideal preference do not only find the content more helpful but are also more willing to pay for it and use the firm's service again (Khambatta et al. 2023).

Future Research

These examples highlight various ways in which the algorithms underlying AI might not align with people's interests, raising numerous questions about such misalignments and their impacts. Future research (see table 1 for a detailed list of topics, organized by the three mechanisms that we have argued constrain human experiences) could explore optimal methods for gathering user input and balancing this against the desire for streamlined interactions. Similarly, research questions should focus on ways to design more identity-supporting environments, while examining the mechanisms behind spillover effects from AI interactions. There is also a need to examine how to reconcile individual peculiarities with the desire for standardized, efficient algorithmic processes. Finally, further investigation is required to understand the implications of alignment or misalignment of algorithms with user interests, particularly concerning trust, satisfaction, and other significant outcomes.

AI impacts domains that are central to quality of life like buying and consumption behavior, healthcare, transportation, criminal justice, employment, personal growth opportunities, and more. By illuminating where AI serves to empower human self-determination, autonomy and progress, this article attempts to provide directions for both public policy and AI design that could shape whether individuals remain in control and enriched by their own technological creations.

REFERENCES

Acquisti, A. (2004), "Privacy in Electronic Commerce and the Economics of Immediate Gratification," in *Proceedings of the ACM Conference on Electronic Commerce*, ed. J. Breese, New York: Association for Computing Machinery, https://doi.org/10.1145/988772.988777.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015), "Privacy and Human Behavior in the Age of Information," *Science*, 347 (6221), 509–14.

Adomavicius, G., J. C. Bockstedt, S. P. Curley, and J. Zhang (2013), "Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects," *Information Systems Research*, 24 (4), 956–75.

Aggarwal, P., and A. L. McGill (2007), "Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products," *Journal of Consumer Research*, 34 (4), 468–79.

Agrawal, A. K., J. S. Gans, and A. Goldfarb (2018), "Exploring the Impact of Artificial Intelligence: Prediction versus Judgment," Working Paper No. 24626, National Bureau of Economic Research, Cambridge, MA.

Ammari, T., J. Kaye, J. Y. Tsai, and F. Bentley (2019), "Music, Search, and IoT: How People (eRally) Use Voice Assistants," *ACM Transactions on Computer-Human Interaction*, 26 (3), 1–28.

André, Q., Z. Carmon, K. Wertenbroch, A. Crum, D. H. Frank, W. Goldstein, J. C. Huber, L. van Boven, B. Weber, and H. Yang (2018), "Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data," *Customer Needs and Solutions*, 5 (1–2), 28–37.

Aral, S., and D. Walker (2012), "Identifying Influential and Susceptible Members of Social Networks," *Science*, 337 (6092), 337–41.

Bailey, E., S. Matz, W. Youyou, and S. Iyengar (2020), "Authentic Self-Expression on Social Media Is Associated with Greater Subjective Well-Being," *Nature Communications*, 11 (1), 4889, https://doi.org/10.1038/s41467-020-18539-w.

Bakshy, E., S. Messing, and L. Adamic (2015), "Exposure to Ideologically Diverse News and Opinion on Facebook," *Science*, 348 (6239), 1130–32.

Bandura, A. (2006), "Toward a Psychology of Human Agency," *Perspectives on Psychological Science*, 1 (2), 164–80.

Bao, L., N. M. Krause, M. N. Calice, D. A. Scheufele, C. D. Wirz, D. Brossard, T. P. Newman, and M. A. Xenos (2022), "Whose AI? How Different Publics Think about AI and Its Social Impacts," *Computers in Human Behavior*, 130 (May), 107182.

Barth, Susanne, and Menno D. T. de Jong (2017), "The Privacy Paradox—Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior—a Systematic Literature Review," *Telematics and Informatics*, 34 (7), 1038–58.

Beranuy, M., U. Oberst, X. Carbonell, and A. Chamarro (2009), "Problematic Internet and Mobile Phone Use and Clinical Symptoms in College Students: The Role of Emotional Intelligence," *Computers in Human Behavior*, 25 (5), 1182–87.

Berg, M., and P. Hudson (1992), "Rehabilitating the Industrial Revolution," *Economic History Review*, 45 (1), 24–50.

Bertrand, J., and L. Weill (2021), "Do Algorithms Discriminate against African Americans in Lending?" *Economic Modeling*, 104 (November), 105619, https://doi.org/10.1016/j.econmod.2021.105619.

Bhandari, A., and S. Bimo (2022), "Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media," *Social Media and Society*, 8 (1), https://doi.org/10.1177/20563051221086241.

Bhattacharjee, A., J. Berger, and G. Menon (2014), "When Identity Marketing Backfires: Consumer Agency in Identity Expression," *Journal of Consumer Research*, 41 (2), 294–309.

Bickmore, T., and J. Cassell (2001), "Relational Agents: A Model and Implementation of Building User Trust," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ed. J. Jacko and A. Sears, New York: Association for Computing Machinery.

Blumer, T., and N. Döring (2012), "Are We the Same Online? The Expression of the Five Factor Personality Traits on the Computer and the Internet," *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 6 (3), https://doi.org/10.5817/CP2012-3-5.

Bluvstein, S., X. Zhao, A. Barasch, and J. Schroeder (2024), "Imperfectly Human: The Humanizing Potential of (Corrected) Errors in Text-Based Communication," *Journal of the Association for Consumer Research*, 9 (3), forthcoming, https://doi.org/10.1086/728412.

Brakus, J. J., B. H. Schmitt, and L. Zarantonello (2009), "Brand Experience: What Is It? How Do We Measure It? And Does It Affect Loyalty?" *Journal of Marketing*, 73 (3), 52–68.

Brugnoli, E., M. Cinelli, W. Quattrociocchi, and A. Scala (2019), "Recursive Patterns in Online Echo Chambers," *Scientific Reports*, 9 (1), https://doi.org/10.1038/s41598-019-56191-7.

Brynjolfsson, E., and A. McAfee (2014), *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, 1st ed., New York: Norton.

Bunker, C., J. M. Balcerowska, L. M. Precht, J. Margraf, and J. Brailovskaia (2024), "Perceiving the Self as Authentic on Social Media Precedes Fewer Mental Health Symptoms: A Longitudinal Approach," *Computers in Human Behavior*, 152 (March), 108056.

Bunker, C., and V. Kwan (2021), "Do the Offline and Social Media Big Five Have the Same Dimensional Structure, Mean Levels, and Predictive Validity of Social Media Outcomes?" *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 15 (4), https://doi.org/10.5817/CP2021-4-8.

Busquet, F., F. Efthymiou, and C. Hildebrand (2023), "Voice Analytics in the Wild: Validity and Predictive Accuracy of Common Audio-Recording Devices," *Behavior Research Methods*, 56 (May), 2114–34.

Caporal, J. (2024), "Identity Theft and Credit Card Fraud Statistics for 2024," *The Ascent*, February 29.

Carmon, Z., R. Schrift, K. Wertenbroch, and H. Yang (2019), "Designing AI Systems That Customers Won't Hate," *MIT Sloan Management Review*, Reprint Number 61315Z.

Carr, N. (2014), *The Glass Cage: Automation and Us*, New York: Norton.

Castelo, N., J. Boegershausen, C. Hildebrand, and A. Henkel (2023), "Understanding and Improving Consumer Reactions to Service Bots," *Journal of Consumer Research*, 50 (4), 848–63.

Castelo, N., M. Bos, and D. Lehmann (2019), "Task-Dependent Algorithm Aversion," *Journal of Marketing Research*, 56 (5), 809–25.

Castelo, N., and J. De Freitas (forthcoming), "Societal Concerns Contribute to AI Aversion, but Are Outweighed by Incentives," working paper.

Chakraborty, A., J. Messias, F. Benevenuto, S. Ghosh, N. Ganguly, and K. Gummadi (2017), "Who Makes Trends? Understanding Demographic Biases in Crowdsourced Recommendations," in *Proceedings of the International AAAI Conference on Web and Social Media*, ed. D Ruths, Cambridge, MA: MIT Press, 22–31.

Chapdelaine, P. (2020), "Algorithmic Personalized Pricing," *New York University Journal of Law and Business*, 17 (1), 1–47.

Cheong, I., Y. E. Huh, and S. Puntoni (2023), "Consumers' Lay Beliefs about AI Evaluation of Interpersonal Skills," in *Proceedings of the Association for Consumer Research Conference*, ed. L. Chaplin, P. Raghubir, and K. Wilcox, Duluth, MN: Association for Consumer Research.

Choi, S., and J. Zhou (2023), "Inducing Consumers' Self-Disclosure through the Fit between Chatbot's Interaction Styles and Regulatory Focus," *Journal of Business Research*, 166 (November), 114127.

Clark-Gordon, C. V., N. D. Bowman, A. K. Goodboy, and A. Wright (2019), "Anonymity and Online Self-Disclosure: A Meta-Analysis," *Communication Reports*, 32 (2), 98–111.

Confessore, N. (2018), "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," *New York Times*, April 4, https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

Cosby, P. C. (1973), “Self-Disclosure: A Literature Review,” *Psychological Bulletin*, 79 (2), 73-91.

Cox, S. R., and W. T. Ooi (2022), “Does Chatbot Language Formality Affect Users’ Self-Disclosure?” in *Proceedings of the 4th Conference on Conversational User Interfaces*, ed. M. Halvey, M. E. Foster, and J. Dalton, New York: Association for Computing Machinery.

Dastin, J. (2018), “Amazon Scraps Secret AI Recruiting Tool That Showed Bias against Women,” Reuters, October 10, https://www.reuters.com/article/us-amazon-com-jobsautomation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-womenidUSKCN1MK08G.

Datta, H., G. Knox, and B. J. Bronnenberg (2018), “Changing Their Tune: How Consumers’ Adoption of Online Streaming Affects Music Consumption and Discovery,” *Marketing Science*, 37 (1), 5–21.

De Freitas, J., S. Agarwal, B. Schmitt, and N. Haslam (2023), “Psychological Factors Underlying Attitudes toward AI Tools,” *Nature Human Behaviour*, 7 (November), 1845–54.

De Freitas, J., and M. Cikara (2021), “Deliberately Prejudiced Self-Driving Cars Elicit the Most Outrage,” *Cognition*, 208 (March), 104555.

De Freitas, J., and I. G. Cohen (2024), “The Health Risks of Generative AI-Based Wellness Apps,” *Nature Medicine*, forthcoming.

De Freitas, J., A. K. Uğuralp, Z. Oğuz-Uğuralp, and S. Puntoni (2024), “Chatbots and Mental Health: Insights into the Safety of Generative AI,” *Journal of Consumer Psychology*, forthcoming.

Dell’Acqua, F., E. McFowland, E. Mollick, H. Lifshitz-Assaf, K. C. Kellogg, S. Rajendran, L. Krayer F. Candelon, and K. R. Lakhani (2023), “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality,” *Social Science Research Network*, https://doi.org/10.2139/ssrn.4573321.

Dietvorst, B. J. (2023), “Aligning Algorithms with Consumers’ Prediction Preferences,” *ACR North American Advances*, 51, 786–87.

Dietvorst, B. J., and D. M. Bartels (2022), “Consumers Object to Algorithms Making Morally Relevant Trade-Offs because of Algorithms’ Consequentialist Decision Strategies,” *Journal of Consumer Psychology*, 32 (3), 406–24.

Dietvorst, B. J., and B. Soaham (2020), “People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error,” *Psychological Science*, 31 (10), 1302–14.

Engler, P. (2021), “Study Finds Liberal Bias at Canadian, American, and British Universities,” July 17, https://patriciaengler.com/2021/07/17/study-finds-liberal-bias-at-canadian-american-and-british-universities/.

Feldman, B. (2020), “TikTok Is Not the Internet’s Eden,” *Intelligencer*, March 16.

Fisher, M., A. H. Smiley, and T. L. H. Grillo (2022), “Information without Knowledge: The Effects of Internet Search on Learning,” *Memory*, 30 (4), 375–87.

Fiske, S. T. (2009), “From Dehumanization and Objectification to Rehumanization,” *Annals of the New York Academy of Sciences*, 1167 (1), 31–34.

Fleder, D., and K. Hosanagar (2009), “Blockbuster Culture’s Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity,” *Management Science*, 55 (5), 697–712.

Flender, C., M. Peters, and G. Müller (2012), “Measuring Consumer Information Deficits in Transactions of Data-centric Services,” in *CECIIS 2012: Twenty-Third Central European Conference on Information and Intelligent Systems*, Varazdin, Croatia: Central European Conference on Information and Intelligent Systems.

French, M. (2018), “Algorithmic Mirrors: an Examination of How Personalized Recommendations Can Shape Self-perceptions and Reinforce Gender Stereotypes,” unpublished dissertation, Stanford University, Stanford, CA.

Furman, J., and R. Seamans (2019), “AI and the Economy,” *Innovation Policy and the Economy*, 19 (1), 161–91.

Gallagher, S. (2000), “Philosophical Conceptions of the Self: Implications for Cognitive Science,” *Trends in Cognitive Science*, 4 (1), 14–21.

Grafanaki, S. (2017), “Autonomy Challenges in the Age of Big Data,” *Fordham Intellectual Property, Media and Entertainment Law Journal*, 27 (4), 803–68.

Granulo, A. Caprioli, C. Fuchs, and S. Puntoni (2024), “Deployment of Algorithms in Management Tasks Reduces Prosocial Motivation,” *Computers in Human Behavior*, 152 (March), 108094, https://doi.org/10.1016/j.chb.2023.108094.

Guzik, E. E., C. Byrge, and C. Gilde (2023), “The Originality of Machines: AI Takes the Torrance Test,” *Journal of Creativity*, 33 (3), 100065.

Harford, T. (2023), “Cautionary Tales—Laser versus Parchment: Doomsday for the Disc,” https://timharford.com/2023/11/cautionary-tales-laser-versus-parchment-doomsday-for-the-disc/.

Haslam, N., and M. Stratemeyer (2016), “Recent Research on Dehumanization,” *Current Opinion in Psychology*, 11, 25–29.

Hildebrand, C. (2019), “The Machine Age of Marketing: How Artificial Intelligence Changes the Way People Think, Act, and Decide,” *NIM Marketing Intelligence Review*, 11 (2), 10–17.

Hildebrand, C., and A. Bergner (2021), “Conversational Robo Advisors as Surrogates of Trust: Onboarding Experience, Firm Perception, and Consumer Financial Decision Making,” *Journal of the Academy of Marketing Science*, 49 (November), 659–76.

Hildebrand, C., F. Efthymiou, W. Busquet, W. Hampton, D. L. Hoffman, and T. P. Novak (2020), “Voice Analytics in Business Research: Conceptual Foundations, Acoustic Feature Extraction, and Applications,” *Journal of Business Research*, 121 (December), 364–74.

Hildebrand, C., D. Hoffman, and T. Novak (2023), “Your Request Is My Command! How Initiation Modalities Shape Conversational AI Experiences,” Working Paper.

Hoffman, D., and T. Novak (2018), “Consumer and Object Experience in the Internet of Things: An Assemblage Theory Approach,” *Journal of Consumer Research*, 44 (6), 1178–204.

Joinson, A. N., U.-D. Reips, T. Buchanan, and C. B. P. Schofield (2010), “Privacy, Trust, and Self-Disclosure Online,” *Human-Computer Interaction*, 25 (1), 1–24.

Keats Citron, D., and D. Solove (2022), “Privacy Harms,” *Boston University Law Review*, 102, 793–863.

Khambatta, P., S. Mariadassou, J. Morris, and S. Wheeler (2023), “Tailoring Recommendation Algorithms to Ideal Preferences Makes Users Better Off,” *Scientific Reports*, 13 (1), 9325–34.

Kim, S. (2017), “Social Media Algorithms: Why You See What You See,” *Georgian Law Technology Review*, 2 (December), 147–54.

Koivisto, M., and S. Grassini (2023), “Best Humans Still Outperform Artificial Intelligence in a Creative Divergent Thinking Task,” *Scientific Reports*, 13 (1), 13601.

Kotkov, D., S. Wang, and J. Veijalainen (2016), “A Survey of Serendipity in Recommende Systems,” *Knowledge-Based Systems*, 111 (November), 180–92.

Langer, E., and J. Rodin (1976), “The Effects of Choice and Enhanced Personal Responsibility for the Aged: A Field Experiment in an Institutional Setting,” *Journal of Personality and Social Psychology*, 34 (2), 191–98.

Larsson, S. (2019), “The Socio-Legal Relevance of Artificial Intelligence,” *Droit et Société*, 103 (3), 573–93.

Lee, Dokyun, and Kartik Hosanagar (2019), "How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation via Randomized Field Experiment," *Information Systems Research*, 30 (1), 239–59.

Lemon, K. N., and P. C. Verhoef (2016), "Understanding Customer Experience throughout the Customer Journey," *Journal of Marketing*, 80 (6), 69–96.

Longoni, C., A. Bonezzi, and C. Morewedge (2019), "Resistance to Medical Artificial Intelligence," *Journal of Consumer Research*, 46 (4), 629–50.

Marketing Science Institute (2018), "2018–2020 Research Priorities Booklet," MSI, New York.

McCarthy, J. (2007), "What Is Artificial Intelligence?" https://www.formal.stanford.edu/jmc/whatisai/whatisai.html.

Melumad, S., and R. J. Meyer (2020), "Full Disclosure: How Smartphones Enhance Consumer Self-Disclosure," *Journal of Marketing*, 84 (3), 28–45.

Melumad, S., and M. T. Pham (2020), "The Smartphone as a Pacifying Technology," *Journal of Consumer Research*, 47 (2), 237–55.

Melzner, J., A. Bonezzi, and T. Meyvis (2023), "Information Disclosure in the Era of Voice Technology," *Journal of Marketing*, 87 (4), 491–509.

Milkman, K. L., T. Rogers, and M. H. Bazerman (2009), "Highbrow Films Gather Dust: Time-Inconsistent Preferences and Online DVD Rentals," *Management Science*, 55 (6), 1047–59.

Moorthy, A. E., and K. L. Vu (2015), "Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space," *International Journal of Human-Computer Interaction*, 31 (4), 307–35.

Morewedge, C. K., S. Mullainathan, H. F. Naushan, C. R. Sunstein, J. Kleinberg, M. Raghavan, and J. O. Ludwig (2023), "Human Bias in Algorithm Design," *Nature Human Behaviour*, 7 (11), 1822–24.

Nass, C., and Y. Moon (2000), "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues*, 56 (1), 81–103.

Nass, C., J. Steuer, and E. R. Tauber (1994), "Computers Are Social Actors," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ed. B. Adelson, S. Dumais, and J. Olson, New York: Association for Computing Machinery.

Noelle-Neumann, E. (1974), "The Spiral of Silence: A Theory of Public Opinion," *Journal of Communication*, 24 (2), 43–51.

Norberg, P. A., D. R. Horne, and D. A. Horne (2007), "Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors," *Journal of Consumer Affairs*, 41 (1), 100–126.

Novak, T. P., and D. L. Hoffman (2019), "Relationship Journeys in the Internet of Things: A New Framework for Understanding Interactions between Consumers and Smart Objects," *Journal of the Academy of Marketing Science*, 47 (2), 216–37.

Noy, S., and W. Zhang (2023), "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Science*, 381 (6654), 187–92.

Onur, A., A. Seidmann, B. Gu, and N. Mazar (2023), "The Effect of Interpretable AI on Repetitive Managerial Decision-Making under Uncertainty," Research Paper, Boston University Questrom School of Business, Boston, MA, https://doi.org/10.2139/ssrn.4331145.

Pitardi, V., and H. Marriott (2021), "Alexa, She's Not Human But . . . Unveiling the Drivers of Consumers' Trust in Voice-Based Artificial Intelligence," *Psychology and Marketing*, 38 (4), 626–42.

Puntoni, S., R. W. Reczek, M. Giesler, and B. Simona (2021), "Consumers and Artificial Intelligence: An Experiential Perspective," *Journal of Marketing*, 85 (1), 131–51.

Qureshi, S. (2023), "Cycles of Development in Systems of Survival with Artificial Intelligence: A Formative Research Agenda," *Information Technology for Development*, 29 (2–3), 171–83.

Ralph, K., and A. R. Nunez (2021), "Understanding the Link between Social Skills and Phone Use," *Journal of Student Research*, 10 (3), https://doi.org/10.47611/jsr.v10i3.1315.

Sangers, T. E., H. Kittler, A. Blum, R. P. Braun, C. Barata, A. Cartocci, M. Combalia, B. Esdaile, P. Guitera, H. A. Haenssle, and N. Kvorning (2024), "Position Statement of the EADV Artificial Intelligence (AI) Task Force on AI-Assisted Smartphone Apps and Web-Based Services for Skin Disease," *Journal of the European Academy of Dermatology and Venereology*, 38 (1), 22–30.

Schwartz, B., A. Ward, J. Monterosso, S. Lyubomirsky, K. White, and D. R. Lehman (2002), "Maximizing versus Satisficing: Happiness Is a Matter of Choice," *Journal of Personality and Social Psychology*, 83 (5), 1178–97.

Simpson, E., and B. Semaan (2021), "For You, or for 'You'? Everyday LGBTQ+ Encounters with TikTok," in *Proceedings of the ACM Symposium on Human Computer Interaction*, New York: Association for Computing Machinery, https://doi.org/10.1145/3432951.

Smets, A., J. Vannieuwenhuyze, and P. Ballon (2022), "Serendipity in the City: User Evaluations of Urban Recommender Systems," *Journal of the Association for Information Science and Technology*, 73 (1), 19–30.

Soh, S., S. Talaifar, and G. M. Harari (2024), "Identity Development in the Digital Context," *Social and Personality Psychology Compass*, 18 (2), e12940.

Solove, D., and D. Keats Citron (2018), "Risk and Anxiety: A Theory of Privacy Harms," *Texas Law Review*, 96, 737–86.

Sweeney, M. E. (2021), "Digital Assistants," in *Uncertain Archives*, ed. D. Agostinho, C. D'Ignazio, A. Ring, N. B. Thylstrup, and K. Veel, Cambridge, MA: MIT Press.

Sweeney, M. E., and E. Davis (2021), "Alexa, Are You Listening?" *Information Technology and Libraries*, 39 (4), https://doi.org/10.6017/ital.v39i4.12363.

Talaifar, S., and A. Ashokkumar (2023), "Consequences of Self-Censoring Unpopular Opinions," in *Academy of Management Proceedings 2023*, ed. S. Taneja, New York: Academy of Management.

Talaifar, S., and B. S. Lowery (2023), "Freedom and Constraint in Digital Environments: Implications for the Self," *Perspectives on Psychological Science*, 18 (3), 544–75.

Tomaino, G., K. Wertenbroch, and D. J. Walters (2023), "Intransitivity of Consumer Preferences for Privacy," *Journal of Marketing Research*, 60 (3), 489–507.

Turing, A. M. (1950), "I.—Computing Machinery and intelligence," *Mind*, LIX (236), 433–60. https://doi.org/10.1093/mind/lix.236.433.

Vimalkumar, M., S. Sharma, J. Singh, and Y. Dwivedi (2021), "'Okay Google, What about My Privacy?' User's Privacy Perceptions and Acceptance of Voice-Based Digital Assistants," *Computers in Human Behavior*, 120 (July), 106763.

Ward, A. F. (2021), "People Mistake the Internet's Knowledge for Their Own," *Proceedings of the National Academy of Sciences* 118 (43), e2105061118.

Waytz, A., J. Heafner, and N. Epley (2014), "The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle," *Journal of Experimental Social Psychology*, 52 (May), 113–17.

Wertenbroch, K., R. Schrift, J. Alba, A. Barasch, A. Bhattacharjee, M. Giesler, J. Knobe, et al. (2020), "Autonomy in Consumer Choice," *Marketing Letters*, 31 (June), 429–39.

Wertenbroch, K., J. Vosgerau, and S. Bruyneel (2008), "Free Will, Temptation, and Self-Control: We Must Believe in Free Will, We Have No Choice (Isaac B. Singer)," *Journal of Consumer Psychology*, 18 (1), 27–33.

Whitman, J. Q. (2004), "The Two Western Cultures of Privacy: Dignity versus Liberty," *Yale Law Journal*, 113 (6), 1151–221.

Willis, G., E. García-Sánchez, A. Sánchez-Rodríguez, J. D. Garcia-Castro, and R. Rodriguez-Bailon (2022), "The Psychosocial Effects of Economic Inequality Depend on Its Perception," *Nature Review Psychology*, 1 (5), 301–9.

Wood, S. (1987), "The Deskilling Debate, New Technology and Work Organization," *Acta Sociologica*, 30 (1), 3–24.

Woolley, K., and M. A. Sharif (2022), "Down a Rabbit Hole: How Prior Media Consumption Shapes Subsequent Media Consumption," *Journal of Marketing Research*, 59 (3), 453–71.

Wu, T.-Y., and D. J. Atkin (2018). "To Comment or Not to Comment: Examining the Influences of Anonymity and Social Support on One's Willingness to Express in Online News Discussions," *New Media and Society*, 20 (12), 4512–32.

Yalcin, G., S. Lim, S. Puntoni, and S. van Osselaer (2022), "Thumbs Up or Down: Consumer Reactions to Decisions by Algorithms versus Humans," *Journal of Marketing Research*, 59 (4), 696–717.

Zierau, N., H. Christian, A. Bergner, F. Busquet, A. Schmitt, and J. M. Leimeister (2022), "Voice Bots on the Frontline: Voice-Based Interfaces Enhance Flow-Like Consumer Experiences and Boost Service Outcomes," *Journal of the Academy of Marketing Science*, 51 (June), 823–42.