ORIGINAL PAPER

Consumer responses to human-AI collaboration at organizational frontlines: strategies to escape algorithm aversion in content creation

Martin Haupt[1,2] · Jan Freidank[2] · Alexander Haas[1]

Abstract

Although Artificial Intelligence can offer significant business benefits, many consumers have negative perceptions of AI, leading to negative reactions when companies act ethically and disclose its use. Based on the pervasive example of content creation (e.g., via tools like ChatGPT), this research examines the potential for human-AI collaboration to preserve consumers' message credibility judgments and attitudes towards the company. The study compares two distinct forms of human-AI collaboration, namely AI-supported human authorship and human-controlled AI authorship, with traditional human authorship or full automation. Building on the compensatory control theory and the algorithm aversion concept, the study evaluates whether disclosing a high human input share (without explicit control) or human control over AI (with lower human input share) can mitigate negative consumer reactions. Moreover, this paper investigates the moderating role of consumers' perceived morality of companies' AI use. Results from two experiments in different contexts reveal that human-AI collaboration can alleviate negative consumer responses, but only when the collaboration indicates human control over AI. Furthermore, the effects of content authorship depend on consumers' moral acceptance of a company's AI use. AI authorship forms without human control lead to more negative consumer responses in case of low perceived morality (and no effects in case of high morality), whereas messages from AI with human control were not perceived differently to human authorship, irrespective of the morality level. These findings provide guidance for managers on how to effectively integrate human-AI collaboration into consumer-facing applications and advises to take consumers' ethical concerns into account.

Keywords Artificial intelligence · Human-AI collaboration · AI augmentation · AI ethics · Content creation · Algorithm aversion

JEL classification C91 · L86 · M31 · O33

 Springer

1 Introduction

Artificial Intelligence (AI) currently reshapes business and marketing strategies as companies increasingly rely on the use of AI systems (Kanbach et al. 2023). Particularly AI-powered tools such as Chat GPT have seen tremendous interest as they are increasingly able to create compelling content that can barely be distinguished from human-authored texts (Köbis and Mossink 2021; Waddell 2018), and scholars identified content generation as a key application area of AI in marketing, legal, finance and other business fields (Dwivedi et al. 2023; Graefe and Bohlken 2020; Kahnt 2019). Although more and more companies use AI, consumers have a negative perception of AI and indicated rather an unwillingness to trust in AI. A recent survey from Salesforce among 11,000 consumers revealed that nearly three quarters of consumers (74%) are concerned about the unethical use of AI and only half of them are open to use AI to improve their experiences (Salesforce 2023). Scholars acknowledged this phenomenon in various studies and termed this negative perception of AI as *algorithm aversion*, which was observed even when algorithms were objectively outperforming humans (Burton et al. 2020; Castelo et al. 2019; Dietvorst et al. 2015, 2016; Yeomans et al. 2019). Reasons for this are that individuals are frightened that AI will attain too much power and get beyond human control (Alfonseca et al. 2021; Burton et al. 2020; Siau and Wang 2020). These negative perceptions are critical for companies that use AI at organizational frontlines (e.g., content creation for company websites), as transparent AI declaration will become a legal obligation in many countries in the near future (e.g., "EU AI Act"; European Parliament 2023). Thus, companies using AI-tools in consumer-facing applications are increasingly confronted with the question of how to integrate AI transparently without suffering from negative consumer responses.

There is a clear need for further research on this question. Research has started to investigate how to leverage the efficiency of AI while avoiding negative consumer responses resulting from its use (Huang and Rust 2022; Zanzotto 2019). As human-AI collaboration seems particularly fruitful to this end, it has received increasing scholarly attention lately (Hassani et al. 2020; Langer and Landers 2021; Raftopoulos et al. 2023; Zhou et al. 2021). For example, this stream of research found that collaborative work between humans and AI increased trust in AI systems and managers' perceptions of empowerment (i.e., the ability to adapt or change) (Schleith et al. 2022). Despite the growing body of research regarding human-AI collaboration in various fields, empirical studies regarding the use and declaration of human-AI collaboration at organizational frontlines remain scarce, particularly in the fields of management and marketing. This lack of empirical studies is surprising given the high potential to create efficiencies at the organizational frontline and the high performance level of modern text-generating tools such as Chat GPT (Dwivedi et al. 2023) in combination with the challenges due to legislative requirements for AI transparency (European Parliament 2023). In addition, the little research from other fields (Waddell 2019; Wölker and Powell 2018) provides conflicting evidence on consumer responses to human-AI

collaboration. Overall, the question of how to transparently integrate AI at organizational frontlines without suffering from negative consumer responses remains open.

Our study addresses this question and investigates consumer responses to human-AI collaboration at organizational frontlines. Specifically, with a focus on content creation for company homepages, we use the concept of *algorithm aversion* (Burton et al. 2020) and *compensatory control theory* (Landau et al. 2015) to develop a conceptual model in which two forms of human-AI collaboration (i.e., AI-supported human authorship and human-controlled AI authorship; Bailer et al. 2022) relate to consumers' attitude towards the company mediated by message credibility. As the rising use of AI in management and marketing has sparked discussions on corporate responsibility and consumers' perceptions of morality of companies' AI use (Cremer and Kasparov 2021; Hagendorff 2020; Siau and Wang 2020; Wirtz et al. 2022), the conceptual model also includes possible moderator effects of consumer-perceived morality of companies' AI use. We test our conceptual model with data from two experimental studies executed with fictitious scenarios and company profiles on the platforms mTurk and Prolific.

Theoretically, this research contributes to a better understanding of human-AI collaboration effects for consumer-facing applications. Our results reveal that AI use at organizational frontlines does not generate negative consumer responses (relative to human authorship) when content creation by AI is controlled by humans or perceived morality is high. With our findings, we add to the debate whether AI should augment or replace humans in management (Hassani et al. 2020; Huang and Rust 2022) and offer insights for the new field of AI ethics and its links to marketing strategy (Siau and Wang 2020). For managers, this research offers a solution to escape the dilemma between ethical (and upcoming legal) misconduct by hiding AI use and negative consumer reactions to transparent AI use. The findings provide them with a deeper understanding of consumer responses to company's AI use, as well as actionable guidance for the highly relevant question of how to manage human-AI collaboration. By doing this, our research also addresses calls regarding the optimal design of human-AI joint workforces (Huang and Rust 2022; Zhou et al. 2021).

2 Theoretical background and hypotheses development

2.1 Performance and perceptions of AI

Related to consumer research, AI can be defined as "any machine that uses any kind of algorithm or statistical model to perform perceptual, cognitive, and conversational functions typical of the human mind" (Longoni et al. 2019, p. 630). Since its inception in the 1950s, AI has undergone remarkable development. Early years saw symbolic AI approaches, focusing on rule-based systems and expert systems. In the recent years, key technologies such as machine learning or neural networks revolutionized AI applications in areas like natural language processing or image processing (Davenport et al. 2020; Hassani et al. 2020). Parallel to other disciplines, the precision and effectiveness of AI in content creation is rapidly developing (Dwivedi

et al. 2023). Quality and precision of AI are rising drastically and AI content is often not distinguishable from human-written content (Köbis and Mossink 2021). A recent meta-study from Graefe and Bohlken (2020) showed that AI-based texts achieved comparable evaluations to human-written texts in various studies—as long as content authorship was hidden. However, when the use of AI is transparent, consumers were found to react differently.

2.2 Transparent AI triggers algorithm aversion

When companies transparently declare their use of AI, scholars widely observed the phenomenon of *algorithm aversion,* that is consumers' reluctance to use AI (compared to humans) (Castelo et al. 2019; Dietvorst et al. 2015). This phenomenon was found in various instances, including product- and service-recommendations (Longoni and Cian 2022; Wien and Peluso 2021), performance-related forecasts (Dietvorst et al. 2015), and financial advice (Önkal et al. 2009). A systematic literature review of Burton et al. (2020) revealed that algorithm aversion has been consistently documented since the 1950's and can be attributed to several causes: Scholars asserted that humans rated AI generally as less trustworthy, less empathetic, and less competent (Chan-Olmsted 2019; Luo et al. 2019).

Moreover, an AI-driven digital agent (i.e., chatbot) was equally effective as a competent human sales agent in terms of conversion rates—but only as long as the chatbot's identity was hidden. By disclosing the AI identity, the purchase rate dropped by over 75% because consumers perceived the chatbot as less knowledgeable and less empathetic than a human salesperson (Luo et al. 2019). Similarly, Castelo et al. (2019) showed that consumers assume that AI is incapable to successfully complete subjective tasks, leading to lower trust and reliance on AI. However, a recent study of Longoni and Cian (2022) showed that product and service attributes (i.e., hedonic or utilitarian contexts) determine whether people prefer AI or human advice, and thus act as a boundary condition for the algorithm aversion effect. Second, humans seemed to expect more perfect results from an AI than from a human, and seeing AI making a mistake led to lower confidence towards the AI and an AI rejection for further tasks (Dietvorst et al. 2015). Third, many processes of AI, such as machine learning, are hard to explain—even for their creators, and thus are often considered as inherently intransparent or as "black box" (Siau and Wang 2020). This deficit of understanding AI creates information asymmetries and fuels fears and distrust (Puntoni et al. 2021).

In line with these findings, algorithm aversion has also been found related to AI content creation. Individuals often assigned higher ratings regarding credibility, readability, or quality to human—(vs. AI-) generated content when authorship was transparent (Graefe and Bohlken 2020; Waddell 2018). Graefe and Bohlken (2020) showed that these ratings were even made regardless of the actual source. That means, despite an identical text, the assignment of an AI (vs. human) authorship systematically leads to more negative ratings.

Scholars consent that algorithm aversion seems to be mainly driven by a low subjective source credibility of AI rather than a lack of objective AI quality (Graefe and

Bohlken 2020; Luo et al. 2019). Essentially, source credibility could be defined as "qualities of an information source which cause what it says to be believable" (West 1994, p. 159). According to the source credibility theory (Hovland et al. 1953), individuals are more likely to be persuaded when the source is evaluated as credible (i.e., expertful and trustworthy). Manifold studies throughout the last decades support this proposition (for an overview, see Ismagilova et al. (2020)). More (vs. less) credible sources were found to create favorable outcomes including enhanced message evaluations, attitudes, and behavioral intentions. For instance, high source credibility leads to higher brand trust or purchase intentions (Harmon and Coney 1982; Luo et al. 2019; Ohanian 1990; Visentin et al. 2019). Moreover, source credibility significantly increases message credibility perceptions (Ismagilova et al. 2020; Visentin et al. 2019) and thus, even the same content could be perceived differently due to different sources. Essentially, message credibility refers to "an individual's judgment of the veracity of the content of communication" (Appelman and Sundar 2016, p. 63).

While most studies provided evidence for the phenomenon of algorithm aversion (see Graefe and Bohlken 2020), some studies found no effect or even a positive effect of AI authorship on perceived content credibility, e.g., in sports news (Wölker and Powell 2018). Thus, although algorithm aversion dominates human perceptions of AI, the effect was not fully consistent throughout content topics or AI tasks.

2.3 Transparent AI triggers perceived loss of control

In addition to that, many people fear that AI could take over control in several domains or threaten human jobs (Huang and Rust 2022). These feelings are not unjustified. When AI takes over a task, it often replaces human intelligence and inevitably takes away human control and jobs as a long-term consequence (e.g., autonomous cars replace taxi drivers (Frey and Osborne 2017; Huang and Rust 2022; Osburg et al. 2022) or AI agents replace journalists (Yerushalmy 2023)). Scholars consent that already "the mere recognition of AI's capability to act as a substitute for human labor can be psychologically threatening" (Puntoni et al. 2021, p. 140).

The desire for control is an essential human need and refers to people's desire to be able to manage processes and outcomes of events in life (Burton et al. 2020; Chen et al. 2017; Puntoni et al. 2021). Herein, control refers to the ability to influence outcomes in one's environment (Skinner 1996). When this need for control is threatened or remains unmet, people experience negative affect, including discomfort, frustration, demotivation, and helplessness, and respond with negative behavior such as moral outrage or reactance (Chen et al. 2017; Landau et al. 2015; Puntoni et al. 2021). Furthermore, according to the compensatory control theory (Landau et al. 2015), individuals who experience a reduced level of control respond with compensatory strategies to restore their perceived control. As traditional strategy, people bolster their personal agency, which is their belief that they possess the resources needed to perform a specific action (Langer 1975). According to a recent literature review of Cutright and Wu (2023), perceptions of low control shape

consumers' behavior either by motivating them to look for a sense of control and order in their consumption environment; or by motivating them to use consumption as a function to regain control. A growing body of literature examines that product acquisition could satisfy consumers' need for control (Billore and Anisimova 2021; Chen et al. 2017; Cutright and Wu 2023).

When it comes to content marketing, consumers might fear that AI becomes so sophisticated that they will not be able to distinguish an AI from a human author (which would be supported by research results like the study from Köbis and Mossink 2021), resulting in a lack of control over the message provider. This, in turn, creates the fear that companies might use AI and manipulate consumers' activities and perceptions (Jobin et al. 2019). For instance, consumers feel uncertain whether a content is genuine human or not—and who controls it (Graefe and Bohlken 2020).

People do not only rely on themselves but also on other humans to restore control. Therefore, a further strategy mentioned in the compensatory control theory is the so-called secondary control, which is a person's belief to have access to an external agent who possesses a desired or needed ability (Landau et al. 2015). That means, a person or institution outside of one's self can influence personally important outcomes and increase the chances to achieve one's goals (Friesen et al. 2014; Kay et al. 2008; Landau et al. 2015). Scholars showed that when people feel a lack of control, they rely stronger on other entities which provide clear rules and structures and thus satisfy their desire for order and control (Friesen et al. 2014; Kay et al. 2008). For instance, individuals were more supportive of hierarchies in the workplace (and favored hierarchy-enhancing jobs) when their sense of control was threatened (Friesen et al. 2014). Similarly, a study with people from 67 nations showed that lower perceived control is strongly correlated with higher support of governmental control (Kay et al. 2008). We adopted this concept of secondary control to our research design as control is exerted by the external agent—the human author.

2.4 Human-AI collaboration as possible escape to negative consumer responses to AI

One possible, but under-researched, solution to mitigate the negative consequences of AI use at organizational frontlines lies in the collaboration of humans and AI, meaning that "AI systems work jointly with humans like teammates or partners to solve problems" (Lai et al. 2021, p. 390). For various management and marketing tasks, scholars consent that AI and humans could collaborate in manifold ways to use the respective strengths of humans and AI (Huang and Rust 2022; Raftopoulos et al. 2023; Zhou et al. 2021). For instance, human-AI collaboration can support healthcare professionals (Lai et al. 2021), general management (Sowa et al. 2021), or data scientists (Wang et al. 2022). Humans could collaborate with AI in advertising (Vakratsas and Wang 2021), marketing planning and strategy (Ameen et al. 2022), or jointly deliver customer service (Wirtz et al. 2018). AI could also augment salespersons' capabilities in every stage of the sales process (Davenport et al. 2020; Paschen et al. 2020). For example, AI could detect unmentioned complaints with the help of automated customer's voice analysis and a human salesperson could

follow up on this (Davenport et al. 2020); AI could predict leads and personalize content, whereas the human could verify leads and link them to the business context (Paschen et al. 2020); and AI could support user experience evaluations (e.g., by identifying issues in usability test videos) to enhance user engagement and sales (Fan et al. 2022). Table 1 offers an overview of relevant conceptual and empirical research regarding human-AI collaboration.

Human-AI collaboration can be designed in different ways on the continuum between the end-points of a sole human actor and sole AI. Following Huang and Rust's (2022) framework of collaborative AI, human-AI collaboration regularly follows a stepwise pattern: Due to the permanent development of AI, AI starts as augmentation and support for humans, and later could replace humans and fulfill the task autonomously. However, in between support and replacement, several scholars acknowledge that AI might perform the task under the surveillance and control of a human (Longoni et al. 2019; Nyholm 2022; Osburg et al. 2022). Related to content creation, Bailer et al. (2022) distinguish between AI-supported human authorship (i.e., labeled as "AI in the loop of human intelligence ") versus AI task take-over with human control (i.e., labeled as "human in the loop of AI").

Although scholars have acknowledged these different collaboration formats, research currently lacks empirical evidence regarding the impact of these different forms of collaboration between humans and AI on consumer responses. In general, scholars suggest that AI is more effective when it augments (vs. replaces) human marketing managers (Davenport et al. 2020), as the cooperation will lead to higher value and competitive advantage compared to human replacement, e.g., in education, medicine, business, science and others (Paschen et al. 2020; Zhou et al. 2021). Several studies from diverse fields showed that integrating humans into AI tasks is reducing their initial algorithm aversion (Burton et al. 2020; Dietvorst et al. 2016; Tobia et al. 2021). Moreover, empirical evidence showed that engaging in collaborative tasks with an AI-driven robot increased consumers' rapport, cooperation, and engagement levels (Seo et al. 2018).

An analysis of human-AI collaboration effects in content marketing is missing. However, scholars in the related field of journalism and news production have started to evaluate this and labeled it "hybrid" or "tandem" authorship. Several authors draw optimistic scenarios where AI could be integrated into journalistic work, and AI and journalists could reach a state of cooperation instead of cannibalization (Graefe and Bohlken 2020; Graefe et al. 2016; Wölker and Powell 2018). Supporting that, Waddell (2019) asserts pragmatically that many current AI systems in journalism still need some human input anyhow, therefore mentioning both human and AI as cooperative authors is recommended. Empirically, Wölker and Powell (2018) show that a human-AI collaboration for largely standardized sports and finance reports is perceived as an equally credible source as a human author, and the collaboration did not lead to lower news selection. These scholars assume that this might be rooted either in the perception of an AI as a more objective author or in initially low expectations toward AI authorship. In sum, empirical evidence generally supports positive impacts of human-AI collaboration, but specific insights about the effects of different collaboration forms are missing. Therefore, the question of how algorithm aversion can best be escaped when declaring AI use remains unanswered.

Table 1 Selected literature related to Human-AI collaboration

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
Conceptual	Davenport et al. (2020)	Multidimensional framework which integrates AI intelligence levels, task types and appearance	Marketing	–	AI will influence marketing strategies and consumer behavior in manifold ways. Authors suggest that human-AI collaboration is more effective than human replacement, and ethical issues need to be considered cautiously
	Huang and Rust (2022)	Conceptual framework for collaborative AI in marketing	Marketing	–	AI advances from mechanical, to thinking, to feeling intelligence. Human-AI collaboration can be achieved through (1) using the respective strengths of human or AI, (2) using lower-level AI to augment higher-level human intelligence, or (3) using AI to automate lower intelligence processes and humans focus on higher intelligence tasks. Possible boundary conditions such as task (un-)desirability are discussed

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Paschen et al. (2020)	Conceptual framework for human-AI collaboration based on B2B sales funnel	Marketing, Sales	–	AI could add significant value in every stage of the sales funnel. Specific contributions of AI and humans in the collaboration are outlined, and benefits for business are derived
	Raftopoulos and Hamari (2023)	Business potentials for value creation due to human-AI collaboration	IS, general management	–	Human-AI collaboration offers substantial potential to create sustainable business value. Authors posit four main enablers for successful value creation: strategic positioning, human engagement, organizational evolution and technology development (including user acceptance to overcome AI aversion or AI anxiety)

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Sundar (2020)	Framework regarding the psychology of Human–AI interaction	Media, HCI	–	The article proposes a framework to investigate synergies of human agency and machine agency. It presents symbolic and enabling effects of AI-driven media affordances on user perceptions, experiences and engagement with AI medium. The author suggests that human-AI co-creation of media will significantly shape these mentioned outcomes
	Zanzotto et al. (2019)	Human-in-the-loop paradigm as collaborative format in the economy	AI research	–	AI has disruptive potential for economy and labor market due to self-learning characteristics, human-AI collaboration as proposed system helps human knowledge creators to participate in AI-driven revenues and wealth

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Zhou et al. (2021)	Conceptual framework for symbiotic human-AI collaboration	IT, HCI	–	Intelligence Augmentation (IA), meaning that AI is used to amplify human abilities is outlined as recommended and highly likely format of human-AI collaboration. IA is expected to increase business' competitiveness and value
Empirical	Fan et al. (2022)	AI tools as support for user experience evaluations	IT, HCI	–	AI-assisted user experience (UX) assessments increased efficiency. AI was particularly more supportive when the AI work is presented with explanations (i.e., how AI identified UX problems)
	Fügener et al. (2022)	Productive delegation of tasks in human-AI collaboration	IS, HCI, general management	–	Human-AI collaboration could outperform AI, which outperforms humans in classification tasks. However, combined performance is only improved when AI delegates work to humans (and not vice versa), because humans could not assess their own capabilities correctly

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Lai et al. (2021)	Literature review on human-AI collaboration in healthcare	Healthcare, Medical	–	Literature on healthcare-related collaboration discusses different AI areas and technologies, focus on specific diseases, different outcomes (e.g., task performance, completion time, learning, usability, acceptance) and stakeholders involved (e.g., healthcare professionals, patients, clinical researchers)
	Longoni et al. (2019)	Consumer responses to human-AI collaborative medical advice	Healthcare, medical	Uniqueness neglect	Human-AI collaboration in giving medical advice eliminates individuals' resistance to accept AI-advice. Thus, AI acceptance is given when it supports (vs. replaces) a human doctor
	Schleith et al. (2022)	Human-AI collaboration regarding information extraction systems	Information systems	–	Installing specific forms of human-AI collaboration (i.e., manual-rule-based-, black-box-review) offers users more understanding and perceived control, which leads to higher trust and joint task efficiency

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Sowa et al. (2021)	Productivity due to human-AI collaboration in managerial tasks	Management	–	Two studies confirm that human-AI collaboration increases perceived productivity. Moreover, people are generally ready to use AI in managerial professions and that collaborative tools should be personalized. Higher sense of agency and ability to influence tool design is suggested to improve the collaboration
	Waddell (2019)	Human-AI collaboration effects on readers' article credibility evaluations	Journalism	Med.: Perceived bias, Anthropomorphism	Author find two mediation paths: human-AI collaboration for news creation was perceived as less biased (vs. human author), leading to positive effects on article (message) credibility. However, collaborative format was also perceived as less anthropomorphic than a human author, creating negative effects on message credibility. Topic type or story context were not significantly influencing these effects

Table 1 (continued)

Research streams	Source	Topic	Literature field(s)	Mediator(s), Moderator(s)	Main findings
	Wang et al. (2022)	Human-AI collaboration in code documentation as part of data science	Data Science	–	AI-assisted coding and code documentation supports human researcher and leads to improved quality and productivity. In particular, human-in-the-loop (and similar) design principles support these goals
	Wölker and Powell (2018)	Perceived article credibility and news selection of automated journalism	Journalism	Med: message credibility, source credibility, Mod.: Article type	Credibility perceptions were not significantly different between human, AI, or human-AI collaborative authorship formats. Moreover, news selection is not influenced by message or source credibility
	Our study	Consumer perceptions of two forms of Human-AI collaborations on firm evaluation	Marketing, management	Med.: Message credibility, Mod.: perceived morality of companies' AI use	Human-AI collaboration can alleviate negative consumer responses regarding AI use, but only when the collaboration indicates human control over AI (vs. AI supporting a human). When consumers perceive a high morality towards companies' AI use, negative effects of any author type vanish

Mod. Moderator(s); *Med.* Mediator(s); *IS* Information systems; *IT* Information technology; *HCI* Human–computer interaction

2.4.1 The relationship between different forms of human-AI collaboration, message credibility, and attitude towards the company

Scholars acknowledge that human-AI collaboration can be evaluated based on different schemes. As cues for the message credibility evaluation, people could either focus on whether human or AI provided the major part of input, or the level of perceived human authority and control over AI (Burton et al. 2020; Dietvorst et al. 2016). A traditional criterion to evaluate content of two authors is to base the decision on the particular workload or input each author provided. For instance, in academic content with cooperative authorship, the authorship order reflects the level of contribution and input share (Newman and Jones 2006). Given the tendency of people's algorithm aversion (Longoni et al. 2019; Luo et al. 2019), higher input share of a human (AI) author is expected to be perceived as more positive (negative). Thus, a higher level of AI input share is expected to reduce message credibility evaluations because people generally rate AI as a less credible source (Luo et al. 2019).

Next to this, people could also evaluate a human-AI collaboration based on the perceived level of human authority and control over AI in the content creation process. As mentioned above (see 2.2), the use of AI as an autonomous system deprives people's sense of control over processes and outcomes (Huang and Rust 2022; Osburg et al. 2022). To counteract this, humans act as supervisors in many processes where AI is used. For instance, humans supervise AI's (semi-)autonomous steering of a car, or a human doctor controls AI's medical advice (Longoni et al. 2019; Osburg et al. 2022).

Related to AI authorship in content creation, it is practically impossible for the readers to influence who writes the text or to verify the content's truthfulness (i.e., objectivity and honesty) (Waddell 2019). Instead, the reader has to rely on secondary control whenever possible—for instance to trust a human co-author or editor and to hand over the control or verification of the content to them.

In general, the desire to have or restore control over one's environment was found to be an innate human need and a quite strong motivator. For instance, when people's feeling of control is impaired, they react with strongly negative affect including anger, moral outrage, or reactance (Puntoni et al. 2021). Longoni et al. (2019) find that people's resistance to use medical AI could be alleviated when AI supported a human who makes the final decision (i.e., is in control) instead of a sole AI service provision. These results support the effectiveness of the form "human-controlled AI authorship".

In contrast, a high human share of input (as indicated in the form "AI-supported human authorship") is expected to be a less clear and powerful cue for the evaluation of message credibility. Particularly when the human input is not clearly visible and distinguishable from AI input (e.g., as mainly given in human-AI collaborative cases), people perceive a higher level of machine agency compared to human agency, and thus a lack of authority (Sundar 2020). Moreover, without human control, individuals might perceive an increased risk of incorrect information (or action) from AI's input as no hierarchies and control functions are sought to be in place

(Osburg et al. 2022). Thus, human control over AI is thought to have a stronger positive influence on message credibility perceptions than human input share. In particular, when human control is not specified, a high level of human input is not expected to reduce the negative impact of AI authorship (vs. a sole human authored message). However, when human control is stated, the perception of secondary control can mitigate the negative impact of AI authorship even with less human input. We hypothesize:

H1a *AI-supported human authorship (vs. human authorship) leads to lower message credibility evaluations.*

H1b *Human-controlled AI authorship (vs. human authorship) does not lead to different message credibility evaluations.*

Following persuasion research, message credibility affects how people make subsequent judgments about the message-sending institution, such as companies or news agencies (Hovland et al. 1953). In particular, credible messages were found to increase consumers' trust and attitudes towards the message sender, and favorable behavioral intentions, including information adoption or purchase intentions (Ismagilova et al. 2020; Wölker and Powell 2018). Thereby, a positive attitude towards the company refers to a readers' positive impression of the company, its reputation, or image (Darke et al. 2008). We posit:

H2 *Stronger perceptions of message credibility lead to more positive attitudes towards the company.*

2.5 The moderating role of morality of AI use

Due to the increasing popularity of AI technologies, AI has gained a substantial impact on humans and society (Hagendorff 2020). Despite undoubted improvements for service quality and customer experience, AI technologies also pose moral threats, such as issues of fairness, ethical misconduct, or consumer privacy (Puntoni et al. 2021). As a response, the new field of *AI ethics* as part of applied ethics gains relevance and momentum (Hagendorff 2020; Siau and Wang 2020). As overarching goals, AI ethics should promote benefits for humans, foster moral behavior to enhance social good ("beneficence"), and prevent any harmful consequences ("non-maleficence") (Hermann 2022; Jobin et al. 2019). As many consumers were found to have moral concerns and reservations toward AI, discussions about the morality of companies' AI use are ongoing in different domains and consider multiple facets (Siau and Wang 2020). Popular moral concerns are the lack of AI control, non-transparent AI processes ("black box"), discrimination, or low reliability of AI-created information (Jobin et al. 2019; Puntoni et al. 2021; Rai 2020). Furthermore, scholars acknowledged possible morality issues when AI is integrated in consumer-facing applications because it could reduce consumer autonomy (Libai et al. 2020) and might be a highly

manipulative system that could cause or support addictive user behavior (Daza and Ilozumba 2022; Hermann 2022). For example, AI could foster exhaustive social media usage through hyper-personalization and optimization of preferred content and ads, which increases marketing effectiveness but is also detrimental to public health (e.g., causing depression or anxiety) (Daza and Ilozumba 2022). Finally, a recent study warned that the increased use of ChatGPT or related AI-driven technologies is supposed to create immense ethical issues, including a rising level of disinformation due to automated fake news, massive low-quality content creation, and a more indirect communication between stakeholders in the society (Illia et al. 2023).

Nevertheless, the strength of these moral concerns related to AI technologies varies from person to person. In particular, some people were found to have a high technological affinity and are less worried about morality issues or possible downsides of AI use (Parasuraman and Colby 2015; Puntoni et al. 2021). These individuals might mainly focus on the innovativeness of AI and have little concerns about moral violations related to their privacy or freedom in decision-making. In contrast, other consumers perceive a high risk and rather distrust AI. This group is more likely to believe that AI is employed to deceive them or take over control (Burton et al. 2020; Parasuraman and Colby 2015). In general, moral judgements were found to influence consumers' perceptions and behavior (Finkel and Krämer 2022; Schermerhorn 2002; Siau and Wang 2020). Research showed that perception of (non-) ethical behavior of a company is an important factor during the purchase decision process. Individuals rewarded a company's ethical behavior by showing a higher willingness to purchase and by paying higher prices for products (Creyer and Ross 1997). Moreover, a recent study in the related field of humanoid robots revealed that consumers' morality perceptions positively influenced robot credibility attributions (Finkel and Krämer 2022). Similarly, related to video news, positive morality judgments were found to lead to higher message credibility (Nelson and Park 2015).

Building on these results, we expect that moral judgements will influence message credibility perceptions and downstream attitudes and behaviors. In particular, we focus on perceived morality of AI use, which relates to consumers' evaluation of how morally acceptable a company's AI use is to them. When people perceive companies' AI use as immoral (i.e. low morality), the use and declaration of authorship forms with AI involvement (i.e., AI or human-AI collaborative authorships) is sought to harm message credibility perceptions. In contrast, when people perceive companies' AI use as morally acceptable (i.e., high morality), the actual use of AI as sole author or co-author should not be an ethical issue. As these consumers exhibit lower moral objections to this kind of AI use, AI should also be perceived as a credible (co-)author, similar to a traditional human author (Creyer and Ross 1997). Therefore, high morality perceptions are expected to delete the negative effects of authorships on message credibility where AI is involved. Thus, we hypothesize:

H3 *Perceived morality of AI use moderates the relationship between authorship type and message credibility: In case of low perceived morality of AI use, message credibility is lower for authorships where AI is involved than for human authorship, and*
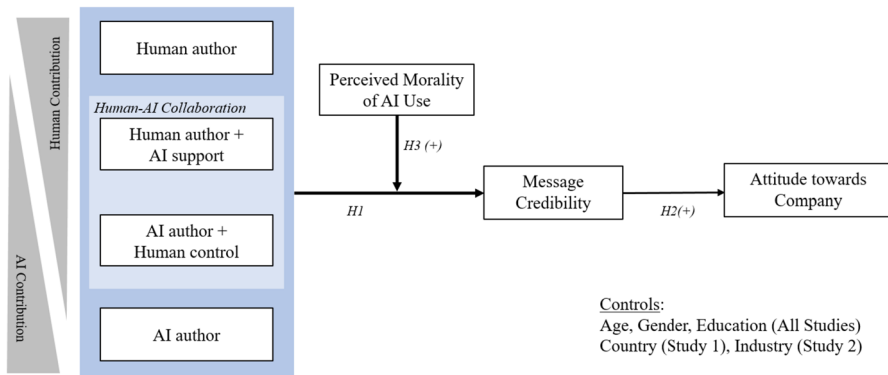
Fig. 1. Conceptual model

there is no difference in message credibility across authorship types when perceived morality of AI use is high.

Figure 1 depicts the conceptual model.

3 Study 1

3.1 Participants and procedure

To examine the proposed causal relationships, we created an experiment and embedded it into an online survey (Hulland et al. 2018). In exchange for a small compensation (\$ 0.75), participants (with a 95% approval rate in former tasks) were recruited from the platform Prolific. Prolific is one of the largest online platforms with over 130,000 participants and widely used in management research to conduct surveys or experiments. These platforms generally reach a more diverse population than traditional sampling methods and allow a quite rapid and inexpensive data collection (Gosling and Mason 2015). In a large comparative study with six major research platforms and panels, Peer et al. (2022) confirmed the data quality of Prolific for academic research. To control for possible effects from a respondent's country of origin, we recruited participants with English as native language from the U.S. and UK. These countries were chosen as many AI-related studies are based on one of these Western countries and the pool of respondents was large enough to ensure a variety of participants (Fig. 1).

After excluding participants who failed the attention check (i.e., "If you read this, please press button 4"), the final sample consisted of 243 participants (54.3% female, $M_{age} = 35$ years, $SD_{age} = 18.29$). As scenario, respondents were exposed to a product information website (i.e., depicting information about a jeans) from a fictitious clothing company (see Fig. 4 in the appendix). We used a simulated company name and website to exclude possibly confounding effects due to prior consumer experiences or attachments with a real brand. Moreover, the jeans scenario was chosen as it represents a common product in the field of consumer goods and does not tend to be a gender-specific product.

To design the scenario content, we have reviewed the design of leading online clothing companies (based on the ranking of the top e-commerce stores in the fashion industry based on revenue in 2022; ECDB 2023). We have included the most-common features of these websites to create a realistic appearance. Moreover, we conducted a pre-test with ten consumers who are experienced in fashion online shopping. They confirmed that the website created resembles those of common clothing companies.

The website was equal across all conditions, except for the author label. Respondents were randomly assigned to one of four experimental conditions, and read one of the following author descriptions: The text was created by (1) a human author (label: "Written by Mary Smith"), (2) AI-supported human authorship (label: "Written by Mary Smith supported by Artificial Intelligence"), (3) a human-controlled AI authorship (label: "Generated by Artificial Intelligence controlled by Mary Smith", (4) an AI author (label: "Generated by Artificial Intelligence"). The author labels were deliberately presented without further details about the form of support or control. A pre-test with seven qualitative interviews with business managers confirmed that managers would label the human-AI collaboration form without any further information. Therefore, the labels used could represent a likely business practice. Moreover, the managers acknowledged that human control refers to a final check of content veracity and indicates human responsibility. In contrast, AI support (for a human) indicates that AI helps with tasks such as text refinement, correct grammar, and spelling. In sum, these results support the theoretical operationalization of the two labels (see chapter 2.2).

After seeing the respective scenario, participants were asked to rate their perceived message credibility (Appelman and Sundar 2016; Obermiller et al. 2005) with four items on a 7-point Likert scale (from 1 = "strongly disagree" to 7 = "strongly agree"). Furthermore, three items were used to assess respondents' attitude towards the company (Darke et al. 2008). Next, we integrated an attention check item and evaluated the case realism with two items from Wagner et al. (2009), namely, "I believe that the described situation could happen in real life" and "I could imagine reading a text like the one presented earlier in real life" ($\alpha = 0.86$; M: 5.26, SD: 1.46). Finally, we asked for participants' age, gender, and education. No significant differences were found between the author groups regarding these three control variables (each $p > 0.1$), suggesting a successful randomization. All psychometric measures were above the recommended levels (see Table 2), indicating construct reliability and validity (Hulland et al. 2018).

As manipulation check, respondents were asked to estimate the share of human versus AI input. Figure 2 illustrates the means, reflecting the expected order. Results of an ANOVA comparing the four author types showed that people perceived that writing shares differ between the author types ($F(3,239) = 89.24$, $p < 0.001$). Post-hoc tests (Bonferroni) showed that all author groups were perceived significantly different from each other (each $p < 0.001$)—except for one. The difference between sole AI authorship and AI controlled by human were not different ($p = 0.13$).

Moreover, to evaluate the effect of authorship types on perceptions of human control over AI, respondents had to indicate "who had the final responsibility for the text", ranging from 1 = AI to 9 = Human) (see Fig. 2). For the ANOVA, the

Table 2 Scale items and statistics

Construct name and items	Standardized loadings	
	Study 1	Study 2
Message Credibility (Study 1/Study 2: α=.88/.91; CR=.88/.86; AVE=.65/.61)		
This text …		
… is generally truthful	0.75	0.75
… leaves one feeling accurately informed	0.79	0.74
… is believable	0.84	0.85
… is authentic	0.83	0.77
Attitude towards the company (Study 1/Study 2: α=.95/.89; CR=.92/.85; AVE=.79/.66)		
This company is a good company	0.88	0.85
This company is a nice company	0.90	0.84
I like the company	0.89	0.74
Morality of AI use (Study 2: α=.93; CR=.93; AVE=.77)		
Companies using artificial intelligence (AI) in marketing texts are…		
Cruel (1) versus Kind-hearted (7)		0.88
Immoral (1) versus Moral (7)		0.90
Uncaring (1) versus Caring (7)		0.83
Unethical (1) versus Ethical (7)		0.89
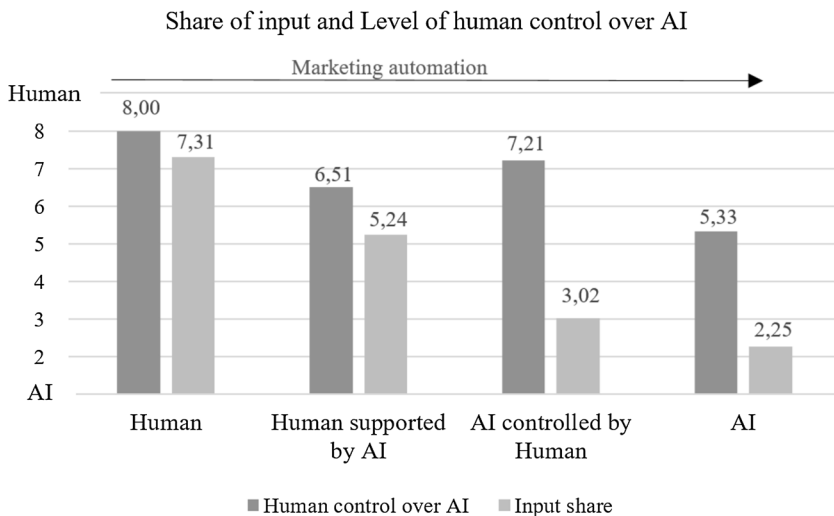
Fig. 2 *Study 1* Consumers' perceptions of Share of Input and Level of Control. Scale ranging from 1=AI to 9=Human

homogeneity of variances was not given (Levene's $F=11.14$, $p<0.001$). To adequately control for this, we used the recommended Welch test and Games-Howell post-hoc tests (Tomarken and Serlin 1986). Results revealed significant differences

between the groups (F_{Welch} (3,131.41) = 13.80, $p < 0.001$). Human control was highest in the case of sole human authorship as no AI was involved, followed by the human-controlled AI authorship and the AI-supported human authorship. Obviously, the lowest level of human control was assigned for sole AI authorship. Post-hoc tests (Games-Howell) showed that human control over AI was significantly higher for human authorship versus AI-supported human authorship or AI (each $p < 0.001$), but not significantly different from the human-controlled AI authorship ($p = 0.25$).

3.2 Results

To test H1 and H2 in one comprehensive model, we ran a mediation model (PROCESS model 4 with 5,000 bootstrapped samples and 95% CI's (Hayes 2018)). The author types were the multicategorical independent variable, message credibility was the mediator, attitude towards the company was the outcome variable, and age, and gender, education, and country of origin were covariates. Related to the author types, the human author was selected as base case to meet the perceptions and attitudes that were given before AI integration. Compared to a human-authored message, respondents perceived an AI author ($b = -0.45$, $p < 0.05$) and an AI-supported human author ($b = -0.71$, $p < 0.005$) as significantly less credible. In contrast, a human-controlled AI author was not perceived significantly different ($p = 0.21$). All covariates had no significant impact on message credibility (each $p > 0.1$). Thus, H1a and H1b could be supported.

In turn, message credibility had a significant impact on attitude towards the company ($b = 0.70$, $p < 0.001$)—supporting H2.

The total effects of authorship types on attitude towards the company were significantly negative for AI authorship ($b = -0.54$, $p < 0.05$) and for the human author supported by AI ($b = -0.57$, $p < 0.05$), but not significant for a human-controlled AI author ($p = 0.15$). Notably, no direct effects of authorship type on attitude towards the company were significant (each $p > 0.1$), indicating a full mediation for the former two author types. Regarding the covariates, no covariate had a total effect on attitude towards the company (each $p > 0.1$).

In sum, both an AI authorship and an AI-supported human authorship have negative effects on readers' attitude towards the company, mediated by lower message credibility perceptions—whereas a human-controlled AI authorship had not such a negative effect (vs. a human author).

4 Study 2

Study 2 aimed to validate the results of Study 1 in another business-related context. In particular, a company's vision statement was chosen as a highly relevant message expressing company values and targets. Furthermore, Study 2 assessed the moderating effects of morality of AI use (H3) on message and company evaluations.

4.1 Participants and procedure

In exchange for a monetary compensation (\$ 0.75), participants from the U.S. were recruited via Amazon mTurk, and randomly assigned to one of the conditions in the 4 (author: human vs. human supported by AI vs. AI controlled by human vs. AI)×2 (industry: kitchen vs. clothing) between-subjects design. We chose mTurk as one of the most prominent online platforms for social science and management research to alter the platform used in study 1 and therefore control for possible confounding effects. Respondents had to surpass 95% completion rate of former tasks and identify English as their native language.

After excluding respondents who failed the attention check or the correct recognition of the author(s), the final dataset consisted of n = 217 respondents (46.5% females, M_{age} = 38 years, SD = 11.37, with an equal or higher than 95% former tasks approval ratio). We altered the industry to control for possible effects due to a more technical or emotional business. Results of two independent samples t-tests showed that the industry type did not influence message credibility (p = 0.31), but the message from the fashion industry was rated marginally more positive than from the kitchen industry ($M_{Fashion}$: 5.59, SD: 1.66, $M_{Kitchen}$: 5.26, SD: 1.38, $t(215)$ = − 1.86, $p < 0.1$).

After accessing the survey, respondents were asked to read a fictitious scenario regarding a company's vision statement that was presented on a website (see Fig. 5 in the appendix). Again, we simulated the stimuli to exclude possible confounding effects (as in study 1). To design this scenario, we compared elements from several large e-commerce companies from the furniture and fashion industry (ECDB Furniture 2023, ECDB Fashion 2023). As in study 1, a pre-test with ten respondents confirmed that the design of the fictitious website is likely to be realistic for a kitchen or fashion company. We used a vision statement as context as it represents a relevant business message and a common online content of many companies. While holding the text equal across the groups, we altered the author types and the industry of the respective company. As measures, participants' perceptions about message credibility, and attitude towards the company were assessed using the same items as in Study 1. Additionally, perceived morality of companies' AI use to create marketing content was evaluated with a 4-item 7-point semantic differential (Olson et al. 2016). Finally, respondents entered their age, gender, and education. All items and factor loadings are shown in Table 2. All psychometric measures were above the recommended levels (see Table 2), suggesting construct reliability and validity (Hulland et al. 2018). Moreover, the experiment groups presented no significant differences regarding the control variables (each $p > 0.1$), suggesting a successful randomization.

As manipulation check, readers of the different author groups had to evaluate the human (vs. AI) share of input. We used the Welch test and Games-Howell post-hoc tests because the assumption of homogeneity of variances was violated. The perceived share of human or AI-input differed significantly across the groups (F_{Welch} $(3,114.63)$ = 180.67, $p < 0.001$). Post-hoc tests (Games-Howell) showed that all groups are significantly different from each other ($p < 0.05$). As expected, people in the human author scenario perceived the highest share of human-input (M: 8.29,

SD: 1.32), followed by the AI-supported human authorship (M: 4.42; SD: 2.06) and the human-controlled AI author (M: 3.39, SD: 1.88), and perceived the lowest share of human authorship in the AI authorship scenario (M: 2.11; SD: 1.39). Regarding human control over AI (i.e., "who had the final responsibility for the text", ranging from $1 = AI$ to $9 = Human$), results were again significantly different between the groups ($F_{Welch}(3,105.34) = 23.92$, $p < 0.001$). Human control was highest in the case of sole human authorship as no AI was involved, followed by human-controlled AI authorship, the AI-supported human authorship, and was least for sole AI authorship. Post-hoc tests (Games-Howell) showed that human control over AI was significantly higher for human authorship vs. AI-supported human authorship or vs. AI (each $p < 0.001$), but not significantly different from a human-controlled AI authorship ($p = 0.62$).

Scenario realism was assessed with two items from Study 1. Again, all scenarios were perceived as realistic ($\alpha = 0.81$; M: 5.97, SD: 1.00), and realism scores did not differ between the author groups ($p > 0.1$). Respondents confirmed that they "want to know about the use of AI" (M: 5.34, SD: 1.48 on a 7-point scale). Furthermore, the call for transparency (European Parliament 2023; Jobin et al. 2019) was also reflected, as respondents agreed that "companies should be obliged to disclose the use of AI" (M: 5.18, SD: 1.59). On average, people seem to perceive companies' AI usage as morally rather acceptable (M: 5.00, SD: 1.36), and this perception did not differ among the authorship groups ($p > 0.1$).

4.2 Results

To assess the hypothesized effects of the authors on message credibility (H1) and subsequently on attitude towards the company (H2), and the moderating effect of morality (H3) in one comprehensive model, we used a moderated mediation analysis with PROCESS (model 8 with 5,000 bootstrapped samples and 95% CIs (Hayes 2018)) based on the same setup as in Study 1. As moderator, we included morality of AI use, and we controlled for age, gender, and industry type. Table 3 illustrates the results.

Respondents rated the text of sole AI authorship as significantly less credible than a (sole) human-authored text ($b = -2.95$, $p < 0.005$). Again, the collaborative authorships were perceived differently: A text from human-controlled AI authorship was not significantly different from a human authorship ($p = 0.65$), but a text from an AI-supported human authorship was rated significantly worse ($b = -1.73$, $p < 0.05$). Thus, although consumers acknowledged that the latter form contains a higher share of human input, this version was rated less credible than a collaboration format with less human input (but human control). The covariates age, gender, industry type, and education had no impact on message credibility ($p > 0.1$).

In turn, message credibility had a significant impact on attitude towards the company ($b = 0.70$, $p < 0.001$). None of the author types had a direct impact on attitude towards the company (each $p > 0.1$, see Table 3), indicating a full mediation via message credibility. Attitudes towards the company were not influenced by age,

Table 3 Study 2. Conditional process model for message credibility as mediator, morality of AI use as moderator, and attitude towards the company as outcome

	Mediator		Outcome	
	Message credibility		Attitude towards the company	
	b	t	b	t
X1: Human supported by AI versus Human	−1.72	−2.07*	−0.48	−0.71[n.s]
X2: AI controlled by Human versus Human	0.39	0.47[n.s]	−0.71	−1.07[n.s]
X3: AI versus Human	−2.93	−3.33**	−0.01	−0.01[n.s]
W: Morality of AI use	0.40	3.22**	−0.03	−0.26[n.s]
M: Message credibility	–	–	0.70	12.48***
X1*W	0.22	1.35[n.s]	0.06	0.44[n.s]
X2*W	−0.12	−0.75[n.s]	0.12	0.90[n.s]
X3*W	0.42	2.38*	−0.01	−0.07[n.s]
COV: Age	0.01	0.80[n.s]	−0.00	−0.50[n.s]
COV: Gender	−0.13	−0.91[n.s]	0.00	0.02[n.s]
COV: Industry type	0.12	0.79[n.s]	0.20	1.70
	Morality	b	Lower	Upper
X1: Human supported by AI versus Human	3.67	−0.64	−1.19	−0.09
	5.02	−0.43	−0.73	−0.15
	6.37	−0.23	−0.53	0.06
X2: AI controlled by Human versus Human	3.67	−0.03	−0.59	0.43
	5.02	−0.15	−0.42	0.11
	6.37	−0.26	−0.54	0.03
X3: AI versus Human	3.67	−0.96	−1.47	−0.47
	5.02	−0.56	−0.86	−0.27
	6.37	−0.16	−0.47	0.17[†]

Conditional indirect effect(s) of X (author types) on Y (attitude towards the company) at values of the moderator (M−1SD, M, M+1SD). Bootstrap 95 percent confidence intervals for conditional indirect effects. †$p < .1$, *$p < .05$, **$p < .01$, ***$p < .001$, *M* mean, *SD* Standard deviation, *n.s.* not significant

gender, or education (each $p > 0.1$), while the fashion industry (vs. kitchen) marginally increased the attitudinal evaluations ($b = 0.21$, $p < 0.1$).

In sum, these results support H1 (a and b) and H2 again. Perceptions of human control over AI were found to be more relevant than share of human input when evaluating message credibility. In particular, human-AI collaboration including explicit human control was found to be equally credible as a sole human authorship, whereas the collaboration with higher human input but without such a human control (i.e., AI-supported human author) was rated as less credible. Thus, in a collaborative setting, people were found to be rather insensitive to human input, but sensitive to human control over AI (H1). In turn, stronger message credibility led to more favorable attitudes towards the company (H2).
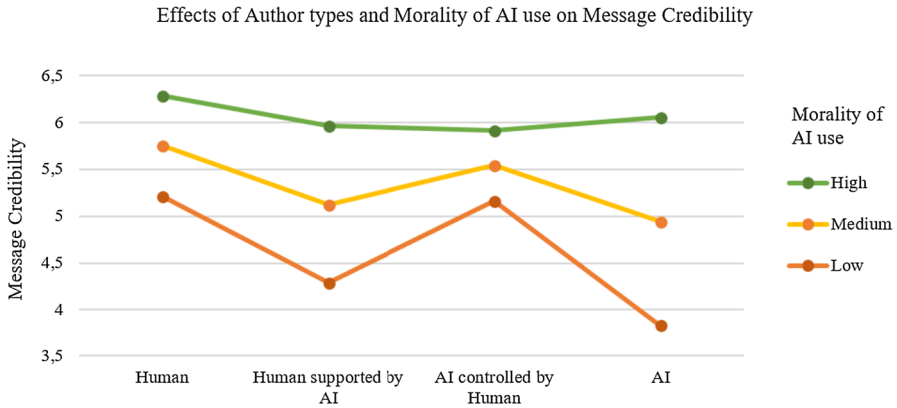
Fig. 3 *Study 2* Influence of morality of AI use on relationship between authors and message credibility. Values represent estimated marginal means, with age, gender and industry as covariates. Morality of AI levels are: low (M−1SD: 3.67), medium (M: 5.02), and high (M+1SD: 6.37)

Regarding the hypothesized moderating effects (H3), the interaction of AI authorship (vs. human)×morality was significant ($b = 0.43$, $p < 0.05$), while interactions of the collaborative author types (vs. human)×morality were not significant (each $p > 0.1$). Yet, the conditional indirect effects of the authorships on message credibility offer a more detailed picture. In case of a low (i.e., M−1SD: 3.67) or medium (M: 5.02) perceived morality of AI use, messages from AI and the human authorship supported by AI were perceived as less credible (low morality: b_{AI}: −1.38, $p < 0.001$, $b_{\text{Human supported by AI}}$: −0.92, $p < 0.005$; medium morality: b_{AI}: −0.81, $p < 0.001$, $b_{\text{Human supported by AI}}$: −0.62, $p < 0.005$). Yet, in case of a high perceived morality (i.e., M+1SD: 6.37), these negative effects on message credibility vanished (high morality: $p_{AI} = 0.53$, $p_{\text{Human supported by AI}} = 0.30$). In contrast, messages from AI controlled by human did not lead to lower message credibility irrespective of the level of perceived morality (each $p > 0.1$). Thus, even for individuals with lower perceptions of morality, the use of AI is not leading to lower credibility perceptions as long as AI is controlled by a human.

Similarly, indirect effects of the author types on attitude towards the company via message credibility were significantly negative for the AI authorship and the human author supported by AI in case of low and medium morality perceptions; and not significant in case of high morality perceptions (see Table 3). These indirect effects were all insignificant for the AI author controlled by human. Despite this clear pattern, the index of moderated mediation was only significant for the AI authorship (index = 0.30 [0.07; 0.53]). Finally, it should be noted that morality had an impact on message credibility ($b = 0.40$, $p < 0.005$), but it had no direct effect on consumers' attitude towards the company ($p = 0.80$). In sum, H3 could be supported for the AI authorship as less credible author type. More generally, perceiving a company's AI use as immoral leads to a stronger credibility devaluation of the author types which lack human control (i.e., AI and Human supported by AI)—whereas this was not the case when AI authorship is controlled by a

human. Furthermore, perceiving a company's AI use as morally acceptable eliminates this effect and consumers accept all author forms of authorships as credible authors (see Fig. 3).

5 Discussion

This study centers on human-AI collaboration as a potential solution to counteract negative consumer responses to AI utilization, given the prevalent skepticism and devaluation of AI in comparison to humans (Dietvorst 2015; Luo et al. 2019). As specific use case, this study examines effects of AI use as content-generating tool, as these applications (such as ChatGPT) have received considerable attention in research and management to enhance business and marketing automation (Bailer et al. 2022; Dwivedi et al. 2023; Kanbach et al. 2023; Puntoni et al. 2021). Across two studies with different contexts related to content on fictitious company websites, we demonstrate that the use and declaration of hybrid authorships could be a solution for the dilemma on how to integrate AI transparently without suffering from negative consumer responses. Thus, human-AI collaborations help exploit the potential for efficiency gains while adhering to the upcoming legislative requirements and circumventing consumers' algorithm aversion.

However, not every form of human-AI collaboration was proven to be effective. In particular, using AI as author with a (final) human control led to comparable message credibility to sole human authorship. In contrast, a human author and AI support led to lower message credibility and reduced consumers' attitudes towards the company—albeit users acknowledged the significantly higher proportion of human input. Thus, consumers were found to care less about the amount of human input, as long as a human had control over AI. Therefore, this study provides a clear recommendation of how to manage human-AI collaboration (and its declaration).

Moreover, the topic of AI use in consumer-facing business applications is also asserted to activate consumers' evaluations of moral behavior of companies (Cremer and Kasparov 2021; Siau and Wang 2020; Wirtz et al. 2022). Therefore, this study evaluates whether consumers' perceptions of morality of a companies' AI use influence their judgment of different author types on message credibility and their attitudes towards the company. Results show that consumers indeed have different levels of moral acceptance of a company's AI use, which lead to an acceptance or rejection of AI as (co-)author. In particular, when individuals do not view the use of AI as highly moral (i.e., low and medium levels of morality), messages from an AI author or an AI-supported human author are perceived as less credible. However, a message from a human-controlled AI author does not decrease credibility, regardless of morality perceptions. When consumers find it morally acceptable that companies use AI for content creation (i.e., high levels of morality), the negative effects of AI use or any collaborative form compared to sole human authorship vanish. Thus, morality perceptions play a substantial role when examining effects of AI use in business applications.

5.1 Theoretical implications

This research offers several relevant theoretical contributions. First, this study extends the emerging literature around human-AI collaboration and AI augmentation (Hassani et al. 2020; Huang and Rust 2022; Zhou et al. 2021) by investigating effects of different human-AI collaboration forms in comparison to human authorship and human replacement (i.e., full AI implementation). More precisely, this paper adds empirical evidence to the sparse literature around human-AI collaboration in management and marketing (Huang and Rust 2022; Zhou et al. 2021). It also enlarges insights into the important field of content creation and content-creating companies such as marketing agencies and news companies (Waddell 2019; Wölker and Powell 2018), which increasingly rely on AI (Yerushalmy 2023). Using the pervasive case of content-generating AI (Chui et al. 2022; Dwivedi et al. 2023; Olson 2022), we also merge this field with business and marketing-related consumer responses. This study therefore enlarges the scope of AI software use to a more emotional and image-related content representing a common marketing-related use case which differs from more "rational " and fact-based content used in journalism. To the best of our knowledge, this study is the first to compare the effects of different human-AI collaboration forms. Scholars have proposed two main forms of collaboration (i.e., AI-supported human authorship vs. human-controlled AI authorship) (Bailer et al. 2022), but research has neglected to evaluate the impact of these forms (e.g., on consumer responses) so far. Furthermore, the results support the ongoing debate about whether AI should augment or replace humans (Hassani et al. 2020; Huang and Rust 2022; Langer and Landers 2021; Shneiderman 2020) by showing that AI could take over the task but human control is desired and human replacement would lead to negative consumer responses. Results reveal that consumers use author labels as evaluation cues to assess the credibility of a message. Thereby, the cue of human control over AI was found to be more effective regarding message credibility and company image than the cue of human input share. Taken together, these results answer scholarly calls how companies should best "distribute work between humans and AI" (Fügener et al. 2022, p. 679) and how "managers can optimize their AI-human intelligence joint workforce" (Huang and Rust 2022, p. 221).

Using the example of text-generating AI, this study also provides insights about how consumers perceive human-AI collaboration related to content creation (as requested by Wölker and Powell 2018). This study further widens the view on AI automation, because the results overcome the one-dimensional view of recent decades that high AI automation is automatically associated with lower human control and safety (Shneiderman 2020). More precisely, scholars asserted that humans have to weigh off between a high level of AI take-over of processes and decisions leading to high AI control and high levels of human control—preventing many AI activities. Instead, using AI with human control offers an escape from this postulated trade-off situation.

Second, this study adds to the literature around individuals' responses to AI. In line with related studies, our results provide evidence for the phenomenon of algorithm aversion (Burton et al. 2020; Dietvorst et al. 2015; Longoni et al. 2019; Luo et al. 2019), leading to a negative impact of AI (vs. human) authorship on message

credibility and company evaluations. However, although consumers were found to have an algorithm aversion, they cared less about the actual amount of human or AI input, but based their evaluation of message credibility on the level of human control over AI. Our findings support the results of Dietvorst et al. (2016) that people need a feeling of control over AI and offer an explanation for the results. Notably, compared to Dietvorst et al. (2016), in our setting, the control over AI was delegated to a human author from the company (as source creator), rather than executed by the consumers themselves. In line with the compensatory control theory (Landau et al. 2015), we show that reducing algorithm aversion can even be achieved via creating perceptions of secondary control. It also supports the notion of Burton et al. (2020) that already the illusion of having control over a decision will alleviate algorithm aversion. Further, our results support Longoni et al.'s (2019) findings that human-AI collaboration (rather than human replacement) could eliminate negative effects of AI which would occur with human replacement. The use of AI in the medical context investigated by Longoni et al. (2019) is very different from the fashion and kitchen industry that we used to settle our study, suggesting generalizability of our findings. Particularly, our results suggest that perceived human control over AI is the underlying mechanism of this consequence. Thereby, scholars acknowledge that consumers could interpret human control in different ways (see Nyholm (2022) for an overview), particularly when further explanations regarding the form of control implementation or details of its execution are missing as it was deliberately the case in our scenario. Interestingly, the positive effect of human control over AI was already present without such further information.

Third, we extend findings to the new field of AI ethics and link it to strategic management and marketing strategy (Cremer and Kasparov 2021; Hagendorff 2020; Siau and Wang 2020). Scholars acknowledge that the implementation of AI in management and consumer-related tasks creates ethical issues for the organization and marketing management, including possible discrimination, loss of consumer autonomy and privacy problems (Puntoni et al. 2021; Siau and Wang 2020). Recently, scholars have increasingly investigated how to handle these issues. For instance, Wirtz et al. (2022) discuss that management should support structures and human personnel as governance mechanism to enable corporate digital responsibility. Our results support this notion and underline the necessity for a human control function to "provide human oversight of AI and refinement of data capture and technologies" (Wirtz et al. 2022, 9).

Finally, the integration of consumers' perception of morality of AI use by companies shows that the moral judgment of AI tools in management and marketing is diverse in society (supporting the research of Parasuraman and Colby 2015 and Puntoni et al. 2021) and influences acceptance or rejection of consumer-facing AI tools and downstream impacts on company evaluations. Therefore, consumers' evaluations about morality of AI use is an important dimension to be considered in AI-related business and marketing research.

5.2 Managerial implications

AI-driven tools offer managers manifold opportunities to raise efficiencies and profitability, but research has also shown that AI use could alienate customers and harm business (Luo et al. 2019; Puntoni et al. 2021). For instance, companies could use generative AI for content creation in various business fields (see Chui et al. (2022) for details). To use the advantages of AI, insights about proper implementation of AI in business strategy and consumer-facing processes are crucial for the company image and customer retention (for a review regarding required competencies, see Santana and Díaz-Fernández 2023). Today, most companies are users of AI technology, and not creators of it. Due to its high complexity, time and cost, the design of AI needs specialized labor—which means that managers have no or only limited possibilities to build or customize AI tools (Kozinets and Gretzel 2021). However, managers could decide about whether and how AI and humans should work together in their processes, which makes guidelines for human-AI collaboration highly relevant for them.

First, results of this study suggest that the use of human-AI collaboration is an effective option to use advantages of AI-driven process automation and, at the same time, to protect the company image. In view of the ethical and upcoming legal obligation to disclose AI use (e.g., European AI Act; European Parliament 2023), this research offers managerial guidance to optimize teamwork of AI and humans (Huang and Rust 2022; Rust 2020). By comparing two main forms of human-AI collaboration (i.e., (1) AI augmentation and (2) AI takeover with human control) (Huang and Rust 2022; Longoni et al. 2019; Osburg et al. 2022), this study showed that the latter form is more beneficial as it did not harm message credibility perceptions and company image. Moreover, in contrast to full AI use or AI augmentation, AI takeover with human control was also perceived similarly to sole human authorship irrespective of readers' perceptions of morality of AI use. This means, that managers could use almost the full potential of AI automation—as long as they install a final human responsibility as well. This is a win–win situation for managers, as they can be ethical and transparently declare AI use and at the same time use AI automation to a high level. This also corresponds to the result of the recent Salesforce report, where 81% of consumers want a human to be in the loop to review or validate AI-created output (Salesforce 2023). Moreover, using a human lead author and AI augmentation was found to have negative impacts on message credibility perceptions and attitudes towards the company. Thus, managers should emphasize the human control function and not the input level. By using a human control function, companies also follow suggestions of scholars to harness higher levels of AI autonomy (Osburg et al. 2022; Santoni de Sio and van den Hoven 2018). Such integration of a human control function enables companies to assume their digital responsibility (Wirtz et al. 2022).

Second, managers and software designers need to be aware that consumers have different judgments of how morally acceptable AI use is. Interestingly, respondents in Study 2 rated the use of AI by companies as quite morally acceptable on average. However, the individual moral judgments differed and influenced the evaluation

of the credibility of the message and the company image. In order to support the moral acceptance of AI use, managers might integrate a message next to the author description explaining the reasons of AI use. For instance, consumers are supposed to understand and accept the need for AI to generate highly personalized content based on individual preferences and past behavior (Puntoni et al. 2021). Moreover, several scholars assert that using explainable AI (i.e., provide information how AI makes decisions and performs actions) could lead to more favorable consumer reactions such as trust in AI or fairness perceptions (Rai 2020).

5.3 Limitations and further research

This research has some limitations, which direct to interesting future research opportunities.

First, both studies relied on standardized and one-way communication with company-generated content in fully simulated scenarios where consumers could not have prior experience or relationships to the companies. We used a fictitious company and website to exclude possibly confounding effects due to such prior experiences or attachments with the brand. Future research could therefore add validity to our findings by replicating our studies for real companies and contexts, or using traditional survey data rather than crowdsourcing platforms. In addition, one of the strengths of AI is the ability to build personalized content based on big data and past consumer behavior (Puntoni et al. 2021). Future studies could examine user behavior in real life context with real companies. This would also allow for evaluations of AI use for personalized content, and incorporate, for example, consumers' trade-offs between appreciating more appropriate information and privacy concerns.

Second, this research focused on consumers' credibility assessment and attitudes towards the company as sender of the message. Future research could investigate other outcomes, for instance actual behavior such as adherence to product recommendations or click rates on web links in the message.

Third, as the use and transparent declaration of AI is touching the field of AI ethics, we integrated the moderator 'morality of AI use' (Cremer and Kasparov 2021; Hagendorff 2020). Although this variable was found to differentiate consumers' author evaluations, the judgment whether companies' AI use is (un-)ethical could vary drastically depending on context, cultural environment, and personality, among others (Zhou et al. 2021). Therefore, future studies might assess the effects of further variations, for instance different settings including morally critical products and services (such as messages related to weapons or politics), varying consumer-company relationships, different countries with divergent ethical norms, or individual-related factors. For instance, consumers' topic involvement might influence their evaluation of the message credibility depending on the author. For highly relevant personal or sensitive topics, people may be less willing to accept AI and maybe even feel devalued being served by a machine. In contrast, for technology-related topics or high-tech companies, people might even admire AI-created content or human-AI collaboration as expressions of an innovative and future-oriented business.

Fourth, in our study, the disclosure of human control over AI in the scenarios does deliberately not include the form of control implementation or details of its execution. However, according to Nyholm (2022), different forms of control exist and might thus be evaluated differently. Future studies could evaluate the impact of different control forms or control framings on consumers' perceptions and company assessments.

Finally, this study uses a cross-sectional design and represents a current snapshot on this dynamic topic. As AI is continuously and rapidly evolving, future research might investigate long-term effects, for instance whether familiarization with AI-generated content leads to more favorable AI evaluations. Parallel to the growth of AI tools, research from different disciplines should orchestrate efforts to explore further effects of human-AI collaborations and the human control function over AI, to achieve an ethical and beneficial use of AI.

Appendix

See Figs. 4 and 5.
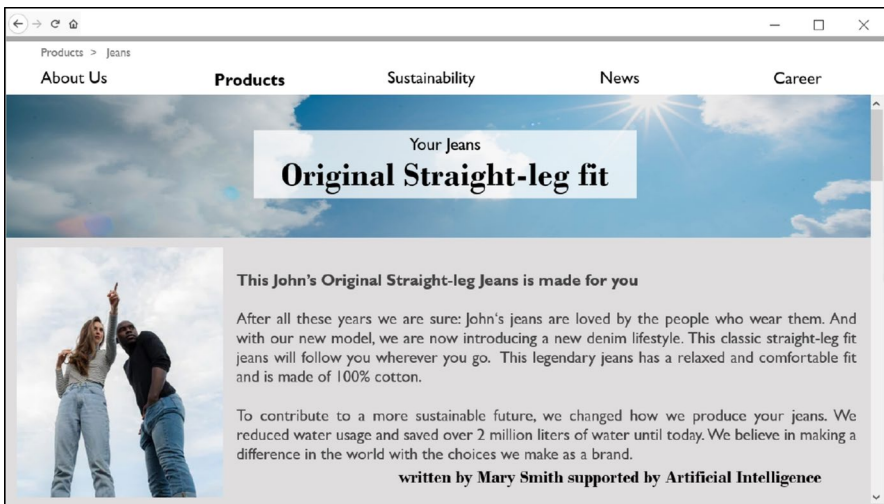
Fig. 4 *Study 1*. Exemplary scenario
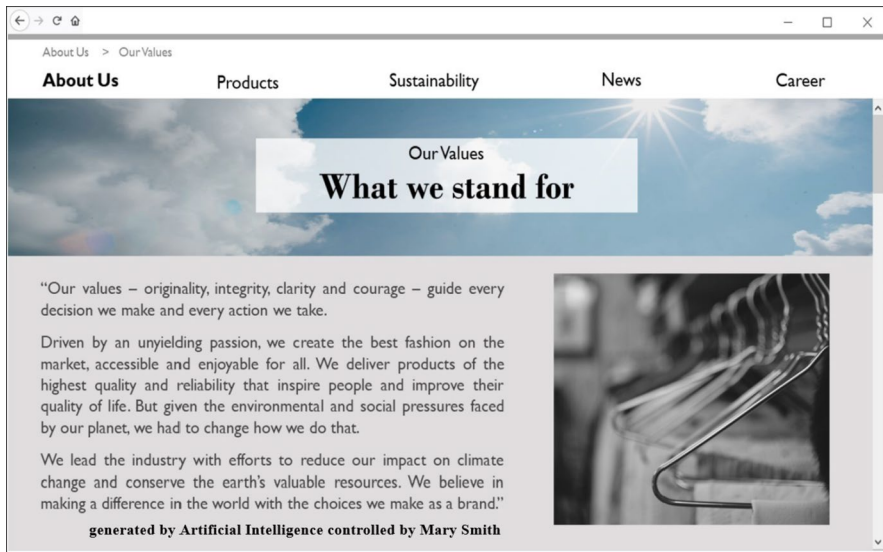
Fig. 5 *Study 2*. Exemplary scenario

Data availability The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest The authors did not receive support from any organization for the submitted work. The authors have no competing interests to declare that are relevant to the content of this article.

References

Alfonseca M, Cebrian M, Fernandez Anta A, Coviello L, Abeliuk A, Rahwan I (2021) Superintelligence cannot be contained: lessons from computability theory. J Artif Intell Res 70:65–76. https://doi.org/10.1613/jair.1.12202

Ameen N, Sharma GD, Tarba S, Rao A, Chopra R (2022) Toward advancing theory on creativity in marketing and artificial intelligence. Psychol Mark 39(9):1802–1825. https://doi.org/10.1002/mar.21699

Appelman A, Sundar SS (2016) Measuring message credibility. Journal Mass Commun Q 93(1):59–79. https://doi.org/10.1177/1077699015606057

Bailer W, Thallinger G, Krawarik V, Schell K, Ertelthalner V (2022) AI for the media industry application potential and automation levels. In: Huet B, ÞórJónsson B, Gurrin C, Tran MT, Dang-Nguyen DT, Hu AMC, Huynh Thi Thanh B (eds) Lecture notes in computer science. multimedia modeling. Springer International Publishing, Cham, pp 109–118

Billore S, Anisimova T (2021) Panic buying research: a systematic literature review and future research agenda. Int J Consum Stud 45(4):777–804. https://doi.org/10.1111/ijcs.12669

Burton JW, Stein M-K, Jensen TB (2020) A systematic review of algorithm aversion in augmented decision making. J Behav Decis Mak 33(2):220–239. https://doi.org/10.1002/bdm.2155

Castelo N, Bos MW, Lehmann DR (2019) Task-dependent algorithm aversion. J Mark Res 56(5):809–825. https://doi.org/10.1177/0022243719851788

Chan-Olmsted SM (2019) A review of artificial intelligence adoptions in the media industry. Int J Media Manag 21(3–4):193–215. https://doi.org/10.1080/14241277.2019.1695619

Chen CY, Lee L, Yap AJ (2017) Control deprivation motivates acquisition of utilitarian products. J Consum Res 43:1031–1047. https://doi.org/10.1093/jcr/ucw068

Chui M, Roberts R, Yee L (2022) Generative AI is here: how tools like ChatGpt could change your business. www.mckinsey.com/capabilities/quantumblack/our-insights/generative-ai-is-here-how-tools-like-chatgpt-could-change-your-business/

Creyer EH, Ross WT (1997) The influence of firm behavior on purchase intention: do consumers really care about business ethics? J Consum Mark 14(6):421–432. https://doi.org/10.1108/07363676971 0185999

Cutright KM, Wu EC (2023) In and out of control: personal control and consumer behavior. Consum Psychol Rev 6(1):33–51. https://doi.org/10.1002/arcp.1083

Darke PR, Ashworth L, Ritchie RJ (2008) Damage from corrective advertising: causes and cures. J Mark 72(6):81–97

Davenport T, Guha A, Grewal D, Bressgott T (2020) How artificial intelligence will change the future of marketing. J Acad Mark Sci 48(1):24–42. https://doi.org/10.1007/s11747-019-00696-0

Daza MT, Ilozumba UJ (2022) A survey of AI ethics in business literature: maps and trends between 2000 and 2021. Front Psychol 13:1042661. https://doi.org/10.3389/fpsyg.2022.1042661

de Cremer D, Kasparov G (2021) The ethical AI—paradox: why better technology needs more and not less human responsibility. AI Ethics. https://doi.org/10.1007/s43681-021-00075-y

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: people erroneously avoid algorithms after seeing them err. J Exp Psychol Gen 144(1):114–126. https://doi.org/10.1037/xge0000033

Dietvorst BJ, Simmons JP, Massey C (2016) Overcoming algorithm aversion: people will use imperfect algorithms if they can (Even slightly) modify them. Manage Sci 64(3):1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Dwivedi YK, Kshetri N, Hughes L, Slade EL, Jeyaraj A, Kar AK, Baabdullah AM, Koohang A, Raghavan V, Ahuja M, Albanna H, Albashrawi MA, Al-Busaidi AS, Balakrishnan J, Barlette Y, Basu S, Bose I, Brooks L, Buhalis D, Wright R (2023) "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int J Inf Manage 71:102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fan M, Yang X, Yu T, Liao QV, Zhao J (2022) Human-AI collaboration for UX evaluation: effects of explanation and synchronization. Proc ACM Hum-Comput Interact 6(CSCW1):1–32. https://doi.org/10.1145/3512943

ECDB Fashion (2023) eCommerceDB. Store Ranking. Top eCommerce stores in the Fashion market, Global revenue 2022 in USD.https://ecommercedb.com/ranking/stores/ww/fashion

Finkel M, Krämer NC (2022) Humanoid robots – artificial. Human-like. Credible? empirical comparisons of source credibility attributions between humans, humanoid robots, and non-human-like devices. Int J Soc Robot 14(6):1397–1411. https://doi.org/10.1007/s12369-022-00879-w

Frey CB, Osborne MA (2017) The future of employment: how susceptible are jobs to computerisation? Technol Forecast Soc Chang 114:254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Friesen JP, Kay AC, Eibach RP, Galinsky AD (2014) Seeking structure in social organization: compensatory control and the psychological advantages of hierarchy. J Pers Soc Psychol 106(4):590–609. https://doi.org/10.1037/a0035620

Fügener A, Grahl J, Gupta A, Ketter W (2022) Cognitive challenges in human-artificial intelligence collaboration: investigating the path toward productive delegation. Inf Syst Res 33(2):678–696. https://doi.org/10.1287/isre.2021.1079

ECDB Furniture (2023) eCommerceDB. Store Ranking. Top eCommerce stores in the Furniture market, Global revenue 2022 in USD. https://ecommercedb.com/ranking/stores/ww/furniture

Gosling SD, Mason W (2015) Internet research in psychology. Annu Rev Psychol 66:877–902. https://doi.org/10.1146/annurev-psych-010814-015321

Graefe A, Bohlken N (2020) Automated journalism: a meta-analysis of readers' perceptions of human-written in comparison to automated news. Media Commun 8(3):50–59. https://doi.org/10.17645/mac.v8i3.3019

Graefe A, Haim M, Haarmann B, Brosius H-B (2016) Readers' perception of computer-generated news: credibility, expertise, and readability. Journalism 19(5):595–610. https://doi.org/10.1177/1464884916641269

Hagendorff T (2020) The ethics of AI ethics: an evaluation of guidelines. Mind Mach 30(1):99–120. https://doi.org/10.1007/s11023-020-09517-8

Harmon RR, Coney KA (1982) The persuasive effects of source credibility in buy and lease situations. J Mark Res 19(2):255–260. https://doi.org/10.1177/002224378201900209

Hassani H, Silva ES, Unger S, Taj Mazinani M, Mac Feely S (2020) Artificial intelligence (AI) or intelligence augmentation (IA): what is the future? AI 1(2):143–155. https://doi.org/10.3390/ai1020008

Hayes AF (2018) Introduction to mediation, moderation and conditional process analysis. A regression-based approach, 2nd edn. The Guilford Press, New York

Hermann E (2022) Leveraging artificial intelligence in marketing for social good-an ethical perspective. J Bus Ethics JBE 179(1):43–61. https://doi.org/10.1007/s10551-021-04843-y

Hovland CI, Janis IL, Kelley HL (1953) Communication and persuasion; psychological studies of opinion change. Yale University Press, New Haven

Huang M-H, Rust RT (2022) A framework for collaborative artificial intelligence in marketing. J Retail 98(2):209–223. https://doi.org/10.1016/j.jretai.2021.03.001

Hulland J, Baumgartner H, Smith KM (2018) Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. J Acad Mark Sci 46(1):92–108. https://doi.org/10.1007/s11747-017-0532-y

Illia L, Colleoni E, Zyglidopoulos S (2023) Ethical implications of text generation in the age of artificial intelligence. Bus Ethics Environ Responsib 32(1):201–210. https://doi.org/10.1111/beer.12479

Ismagilova E, Slade E, Rana NP, Dwivedi YK (2020) The effect of characteristics of source credibility on consumer behaviour: a meta-analysis. J Retail Consum Serv 53:101736. https://doi.org/10.1016/j.jretconser.2019.01.005

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1(9):389–399. https://doi.org/10.1038/s42256-019-0088-2

Kahnt I (2019) Künstliche intelligenz im content marketing. In: Wesselmann M (ed) Content Gekonnt: strategie, organisation, umsetzung, ROI-messung und fallbeispiele aus der praxis. Springer Gabler, Cham, pp 211–225

Kanbach DK, Heiduk L, Blueher G, Schreiter M, Lahmann A (2023) The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. Rev Managerial Sci. https://doi.org/10.1007/s11846-023-00696-z

Kay AC, Gaucher D, Napier JL, Callan MJ, Laurin K (2008) God and the government: testing a compensatory control mechanism for the support of external systems. J Pers Soc Psychol 95(1):18–35. https://doi.org/10.1037/0022-3514.95.1.18

Köbis N, Mossink LD (2021) Artificial intelligence versus Maya Angelou: experimental evidence that people cannot differentiate AI-generated from human-written poetry. Comput Hum Behav 114:106553. https://doi.org/10.1016/j.chb.2020.106553

Kozinets RV, Gretzel U (2021) Commentary: artificial intelligence: The marketer's dilemma. J Mark 85(1):156–159. https://doi.org/10.1177/0022242920972933

Lai Y, Kankanhalli A, Ong D (2021) Human-AI collaboration in healthcare: a review and research Agenda. In: Proceedings of the annual Hawaii international conference on system sciences. Hawaii International Conference on System Sciences, pp 390–399. https://doi.org/10.24251/hicss.2021.046

Landau MJ, Kay AC, Whitson JA (2015) Compensatory control and the appeal of a structured world. Psychol Bull 141(3):694–722. https://doi.org/10.1037/a0038703

Langer EJ (1975) The illusion of control. J Pers Soc Psychol 32(2):311–328. https://doi.org/10.1037/0022-3514.32.2.311

Langer M, Landers RN (2021) The future of artificial intelligence at work: a review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. Comput Hum Behav 123:106878. https://doi.org/10.1016/j.chb.2021.106878

Libai B, Bart Y, Gensler S, Hofacker CF, Kaplan A, Kötterheinrich K, Kroll EB (2020) Brave new world? On AI and the management of customer relationships. J Interact Mark 51:44–56. https://doi.org/10.1016/j.intmar.2020.04.002

Longoni C, Cian L (2022) Artificial intelligence in utilitarian vs. hedonic contexts: the "Word-of-machine" effect. J Mark 86(1):91–108. https://doi.org/10.1177/0022242920957347

Longoni C, Bonezzi A, Morewedge CK (2019) Resistance to medical artificial intelligence. J Consum Res 46(4):629–650. https://doi.org/10.1093/jcr/ucz013

Luo X, Tong S, Fang Z, Qu Z (2019) Frontiers: machines vs. humans: the impact of artificial intelligence Chatbot disclosure on customer purchases. Mark Sci. https://doi.org/10.1287/mksc.2019.1192

Nelson MR, Park J (2015) Publicity as covert marketing? The role of persuasion knowledge and ethical perceptions on beliefs and credibility in a video news release story. J Bus Ethics 130(2):327–341. https://doi.org/10.1007/s10551-014-2227-3

Newman A, Jones R (2006) Authorship of research papers: ethical and professional issues for short-term researchers. J Med Ethics 32(7):420–423. https://doi.org/10.1136/jme.2005.012757

Nyholm S (2022) A new control problem? Humanoid robots, artificial intelligence, and the value of control. AI Ethics. https://doi.org/10.1007/s43681-022-00231-y

Obermiller C, Spangenberg E, MacLachlan DL (2005) Ad skepticism: the consequences of disbelief. J Advertising 34(3):7–17

Ohanian R (1990) Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. J Advert 19(3):39–52. https://doi.org/10.1080/00913367.1990.10673191

Olson JG, McFerran B, Morales AC, Dahl DW (2016) Wealth and welfare: divergent moral reactions to ethical consumer choices. J Consum Res 42(6):879–896. https://doi.org/10.1093/jcr/ucv096

Olson P (2022) Google faces a serious threat from ChatGPT. https://www.washingtonpost.com/business/energy/google-faces-a-serious-threat-from-chatgpt/2022/12/07/363d2440-75f5-11ed-a199-927b334b939f_story.html

Önkal D, Goodwin P, Thomson M, Gönül S, Pollock A (2009) The relative influence of advice from human experts and statistical methods on forecast adjustments. J Behav Decis Mak 22(4):390–409. https://doi.org/10.1002/bdm.637

Osburg VS, Yoganathan V, Kunz WH, Tarba S (2022) Can (A)I give you a ride? Development and validation of the CRUISE framework for autonomous vehicle services. J Serv Res. https://doi.org/10.1177/10946705221118233

Parasuraman A, Colby CL (2015) An updated and streamlined technology readiness index. J Serv Res 18(1):59–74. https://doi.org/10.1177/1094670514539730

European Parliament (2023) EU AI Act: first regulation on artificial intelligence. European Parliament. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Paschen J, Wilson M, Ferreira JJ (2020) Collaborative intelligence: how human and artificial intelligence create value along the B2B sales funnel. Bus Horizons 63(3):403–414. https://doi.org/10.1016/j.bushor.2020.01.003

Peer E, Rothschild D, Gordon A, Evernden Z, Damer E (2022) Data quality of platforms and panels for online behavioral research. Behav Res Methods 54(4):1643–1662. https://doi.org/10.3758/2Fs13428-021-01694-3

Peer E, Rothschild D, Gordon A, Evernden Z, Damer E (2022) Data quality of platforms and panels for online behavioral research. Behav Res Methods 54:1643–1662. https://doi.org/10.3758/s13428-021-01694-3

Puntoni S, Reczek RW, Giesler M, Botti S (2021) Consumers and artificial intelligence: an experiential perspective. J Mark 85(1):131–151. https://doi.org/10.1177/0022242920953847

Raftopoulos M, Hamari J (2023) Human-AI collaboration in organisations: A literature review on enabling value creation. In: European conference on information systems, p 31

Rai A (2020) Explainable AI: from black box to glass box. J Acad Mark Sci 48(1):137–141. https://doi.org/10.1007/s11747-019-00710-5

Rust RT (2020) The future of marketing. Int J Res Mark 37(1):15–26. https://doi.org/10.1016/j.ijresmar.2019.08.002

Salesforce (2023) Artificial intelligence. Businesses Adopting AI Risk a 'Trust Gap' with Customers – Salesforce Report. 28.08.2023. https://www.salesforce.com/news/stories/customer-engagement-research-2023/

Santana M, Díaz-Fernández M (2023) Competencies for the artificial intelligence age: visualisation of the state of the art and future perspectives. RMS 17(6):1971–2004. https://doi.org/10.1007/s11846-022-00613-w

Santoni de Sio F, Van den Hoven J (2018) Meaningful human control over autonomous systems: a philosophical account. Front Robot AI 5:15. https://doi.org/10.3389/frobt.2018.00015

Schermerhorn JR (2002) Management, 7th edn. Wiley, Hoboken

Schleith J, Hoffmann H, Norkute M, Cechmanek B (2022) Human in the loop information extraction increases efficiency and trust. https://doi.org/10.18420/muc2022-mci-ws12-249

Seo SH, Griffin K, Young JE, Bunt A, Prentice S, Loureiro-Rodríguez V (2018) Investigating people's rapport building and hindering behaviors when working with a collaborative robot. Int J Soc Robot 10(1):147–161. https://doi.org/10.1007/s12369-017-0441-8

Shneiderman B (2020) Human-centered artificial intelligence: reliable, safe & trustworthy. Int J Hum-Comput Interact 36(6):495–504. https://doi.org/10.1080/10447318.2020.1741118

Siau K, Wang W (2020) Artificial intelligence (AI) ethics. J Database Manage 31(2):74–87. https://doi.org/10.4018/JDM.2020040105

Skinner EA (1996) A guide to constructs of control. J Pers Soc Psychol 71(3):549–570. https://doi.org/10.1037/0022-3514.71.3.549

Sowa K, Przegalinska A, Ciechanowski L (2021) Cobots in knowledge work. J Bus Res 125:135–142. https://doi.org/10.1016/j.jbusres.2020.11.038

Sundar SS (2020) Rise of machine agency: a framework for studying the psychology of human–AI interaction (HAII). J Comput-Mediat Commun 25(1):74–88. https://doi.org/10.1093/jcmc/zmz026

Tobia K, Nielsen A, Stremitzer A (2021) When does physician use of AI increase liability? J Nucl Med Off Publ Soc Nucl Med 62(1):17–21. https://doi.org/10.2967/jnumed.120.256032

Tomarken AJ, Serlin RC (1986) Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. Psychol Bull 99(1):90–99. https://doi.org/10.1037/0033-2909.99.1.90

Vakratsas D, Wang X (2021) Artificial intelligence in advertising creativity. J Advert 50(1):39–51. https://doi.org/10.1080/00913367.2020.1843090

Visentin M, Pizzi G, Pichierri M (2019) Fake news, real problems for brands: the impact of content truthfulness and source credibility on consumers' behavioral intentions toward the advertised brands. J Interact Mark 45:99–112. https://doi.org/10.1016/j.intmar.2018.09.001

Waddell TF (2018) A robot wrote this? Digit Journal 6(2):236–255. https://doi.org/10.1080/21670811.2017.1384319

Waddell TF (2019) Can an algorithm reduce the perceived bias of news? Testing the effect of machine attribution on news readers' evaluations of bias, anthropomorphism, and credibility. Journal Mass Commun Q 96(1):82–100. https://doi.org/10.1177/1077699018815891

Wagner T, Hennig-Thurau T, Rudolph T (2009) Does customer demotion jeopardize loyalty? J Mark 73(3):69–85. https://doi.org/10.1509/jmkg.73.3.69

Wang AY, Wang D, Drozdal J, Muller M, Park S, Weisz JD, Liu X, Wu L, Dugan C (2022) Documentation matters: human-centered AI system to assist data science code documentation in computational notebooks. ACM Trans Comput-Hum Interact 29(2):1–33. https://doi.org/10.1145/3489465

West MD (1994) Validating a scale for the measurement of credibility: a covariance structure modeling approach. Journal Q 71(1):159–168. https://doi.org/10.1177/107769909407100115

Wien AH, Peluso AM (2021) Influence of human versus AI recommenders: the roles of product type and cognitive processes. J Bus Res 137:13–27. https://doi.org/10.1016/j.jbusres.2021.08.016

Wirtz J, Patterson PG, Kunz WH, Gruber T, Lu VN, Paluch S, Martins A (2018) Brave new world: service robots in the frontline. J Serv Manag 29(5):907–931. https://doi.org/10.1108/JOSM-04-2018-0119

Wirtz J, Kunz WH, Hartley N, Tarbit J (2022) Corporate digital responsibility in service firms and their ecosystems. J Serv Res. https://doi.org/10.1177/10946705221130467

Wölker A, Powell TE (2018) Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. Journalism. https://doi.org/10.1177/1464884918757072

Yeomans M, Shah A, Mullainathan S, Kleinberg J (2019) Making sense of recommendations. J Behav Decis Mak 32(4):403–414. https://doi.org/10.1002/bdm.2118

Yerushalmy J (2023) German publisher Axel Springer says Journalists could be replaced by AI. https://www.theguardian.com/technology/2023/mar/01/german-publisher-axel-springer-says-journalists-could-be-replaced-by-ai

Zanzotto FM (2019) Viewpoint: human-in-the-loop artificial intelligence. J Artif Intell Res 64:243–252. https://doi.org/10.1613/jair.1.11345

Zhou L, Paul S, Demirkan H, Yuan L, Spohrer J, Zhou M, Basu J (2021) Intelligence augmentation: towards building human- machine symbiotic relationship. AIS Trans Hum-Comput Interact 13(2):243–264. https://doi.org/10.17705/1thci.00149

Authors and Affiliations

Martin Haupt[1,2] · Jan Freidank[2] · Alexander Haas[1]

✉ Martin Haupt
martin.haupt@w.thm.de

Jan Freidank
jan.freidank@w.thm.de

Alexander Haas
alexander.haas@wirtschaft.uni-giessen.de

[1] Justus-Liebig University, Marketing and Sales Management, Licher Strasse 66, 35394 Giessen, Germany

[2] Technische Hochschule Mittelhessen, THM Business School, Wiesenstrasse 14, 35390 Giessen, Germany