Archives of Sexual Behavior
https://doi.org/10.1007/s10508-025-03318-9

SPECIAL SECTION: ARTIFICIAL INTELLIGENCE AND SEXUALITY

Anti-AI Bias Toward Couple Images and Couple Counseling: Findings from Two Experiments

Nicola Döring[1] · **M. Rohangis Mohseni**[1]

Received: 18 August 2025 / Revised: 11 September 2025 / Accepted: 12 September 2025

Abstract
Generative artificial intelligence (AI) systems can produce text, images, videos, and audio in response to prompts. They are increasingly applied across various domains, including intimacy and sexuality—ranging from AI-generated pornography to sexual counseling via AI chatbots. While AI-generated content holds significant potential, it is also met with skepticism. Anti-AI bias is defined as a systematic tendency to evaluate AI-produced outputs more negatively than equivalent human-created content, regardless of actual quality. Following the experimental labeling paradigm, this study examined whether identical couple images (H_{1a}) and couple counseling excerpts (H_{2a}) were evaluated less favorably when labeled as AI-generated rather than human-created, and whether AI attitudes and AI literacy moderated these effects for images (H_{1b}) and counseling dialogues (H_{2b}). Two consecutive online experiments were conducted in 2024 with a national sample of adults in Germany ($N = 2{,}658$). In Experiment 1, identical romantic couple images received less positive evaluations when labeled as AI-generated images versus as human-generated photographs ($d = .21$; H_{1a}). In Experiment 2, identical sexuality-related couple counseling excerpts labeled as involving an AI counselor were rated less favorably than those labeled as involving a human counselor ($d = .23$; H_{2a}). AI attitudes and AI literacy combined moderated the labeling effect for images ($\eta^2 = .01$; H_{1b}) but not for counseling dialogues ($\eta^2 = .003$; H_{2b}). These findings extend the literature on anti-AI bias into intimate contexts. They also underscore the importance of considering user dispositions toward AI when designing and implementing generative AI systems in intimacy- and sexuality-related domains.

Keywords Artificial intelligence (AI) · AI attitudes · AI-supported sexual activities (AISA) · AI-generated content · Online experiment

Introduction

Since late 2022, a new kind of artificial intelligence (AI) has entered everyday life: generative AI—systems that can produce text (e.g., ChatGPT, Gemini), images (e.g., DALL-E, Midjourney), videos (e.g., Sora, Veo), and audio (e.g., Suno, Udio) in response to prompts (Dwivedi et al., 2023; He et al., 2025). These systems create new content that closely resembles human-made work (Sengar et al., 2025). Their use has expanded rapidly across domains such as work, education, entertainment, health, politics, social, and intimate lives (Storey et al., 2025).

As a result, people now encounter AI-generated content nearly everywhere: in news articles, advertising campaigns and product descriptions, on book covers, in curated music playlists, on social media, in search engine results, and even in personal emails and chats (Zagalo & Keller, 2026). Generative AI has also moved into more intimate and sensitive spheres, such as sexuality. Online sexual activities (OSA) are now complemented by AI-supported sexual activities (AISA), which include, among others, AI-generated erotic and pornographic imagery as well as AI-based sexual and relationship counseling (Döring et al., 2025a).

The rise of AI-generated content has sparked intense debate. While some are excited by its creative potential, quality, and democratizing effects, others express deep concerns about misinformation, exploitation, deception, and a flood of low-quality outputs—what critics call AI "slop" (Smith & Southerton, 2025). This divide mirrors a more general tension between so-called AI optimism (i.e., the belief that

✉ Nicola Döring
nicola.doering@tu-ilmenau.de

[1] Department of Economic Sciences and Media, Technische Universität Ilmenau, Ehrenbergstr. 29, 98693 Ilmenau, Germany

Published online: 08 October 2025

AI will bring major benefits to society) versus AI pessimism (i.e., the belief that AI is overhyped, unlikely to deliver on big promises, and will bring major threats to society; Gondlach & Regneri, 2023; Guingrich & Graziano, 2025; Montag et al., 2025).

The aim of this brief report is to test the hypothesis that people evaluate intimate content more negatively when they assume it is AI-generated rather than human-created, regardless of its actual quality—an expression of so-called anti-AI bias. Building on this premise, the study addresses a notable research gap: while anti-AI bias has been examined in domains such as art, health, and journalism, its operation in relation to intimate and sexual content remains largely unexplored. By focusing on this under-investigated area, the study broadens the scope of anti-AI bias research and provides an initial empirical basis for examining how assumptions about content origin may shape perceptions of mediated intimacy. The findings aim to stimulate further scholarly inquiry into the intersection of generative AI, sexual representation, and user evaluation.

Anti-AI Bias

In the context of generative AI, two broad types of biases are typically investigated. The first concerns bias in AI outputs (short: AI bias)—that is, stereotypes or discriminatory patterns embedded in AI-generated content due to unbalanced training data, flawed model architectures, or inequitable algorithmic decision-making processes (Mehrabi et al., 2022). The second concerns bias in human evaluations of AI, which—in line with AI pessimism—can manifest as anti-AI bias—a systematic tendency to evaluate AI-produced outputs (e.g., an image, a song, a newspaper article) more negatively compared to human-created content, regardless of its actual quality (Ansani et al., 2025; Bellaiche et al., 2023; Reis et al., 2024).

Anti-AI bias should be distinguished from *algorithm aversion*, which refers to people's reluctance to use, trust, or rely on algorithms after observing them make mistakes—even when the algorithm still outperforms human decision-makers on average (Dietvorst et al., 2015). While algorithm aversion is typically triggered by the observation of errors and concerns about predictive accuracy, anti-AI bias can occur without any such observation and is tied instead to perceptions of source credibility (Hovland et al., 1953), authenticity (Kernis & Goldman, 2006), and human agency (Bandura, 2018). For example, in the case of artistic or intimate content, people might perceive AI-generated output as inherently machine-like and lacking humanness, and thus regard it as "soulless," "cheap," "banal," or "inauthentic," and ultimately inferior to human-created content (so-called negative machine heuristic; Molina & Sundar, 2024). This anti-AI bias emerges both in empirical studies (e.g., Grassini & Koivisto, 2024) and in theoretical papers in which experts reflect on AI-generated content—such as synthetic pornography—often emphasizing a perceived lack of "humanness," "authenticity," "truth," and "aura" in such works (e.g., Alilunas, 2024; Easterbrook-Smith, 2025).

In line with AI optimism, there is also reason to assume a pro-AI bias—a systematic tendency to evaluate AI-produced outputs more positively than human-created content in certain contexts. For example, in the case of highly debated political issues, human-created journalistic content may be perceived as biased, opinion-driven, and partisan, whereas AI-generated output—precisely because of its lack of humanness and its machine-like character—may be regarded as more "neutral," "fact-based," "objective," and "credible" (the so-called positive machine heuristic; Molina & Sundar, 2024; Sundar & Kim, 2019; Waddell, 2019). Some experts even extend the positive machine heuristic to art, suggesting that machine advantages (e.g., speed, volume, pattern recognition) need not be equated solely with soulless efficiency but could also open up new creative possibilities (Chatterjee, 2022). Although theoretical conceptualizations around a positive machine heuristic and some empirical findings (e.g., Ovsyannikova et al., 2025) point to a pro-AI bias in certain contexts, public and academic debates have so far focused more on anti-AI bias.

State of Research

Different methodological approaches are used to examine individuals' perceptions and evaluations of AI-generated content, including potential negative bias toward it.

- Survey-Based Assessments: Some studies ask laypersons and/or experts for their opinions on certain types of AI-generated content. For example, one survey asked patients (Nadarzynski et al., 2020) and another asked health professionals (Nadarzynski et al., 2023) about the acceptability of AI chatbots in sexual and reproductive health. Such studies typically find evidence of anti-AI bias, reflected in reluctance or skepticism toward AI—particularly among respondents unfamiliar with AI technology who perceive its involvement as inferior to human engagement. However, a negative stance toward AI reported in surveys can reflect either a heuristic-based anti-AI bias or a rational critique based on actual characteristics of AI output.
- Direct Content Comparisons: Another approach involves showing participants actual AI-generated and/or human-generated content and asking them to evaluate or directly compare the two. These studies often explore participants' ability to detect AI-generated content when it is unlabeled, yielding mixed results regarding detectability. Findings frequently indicate an anti-AI bias, manifested as a preference for content perceived as human-made—whether that

perception is correct or not—over content perceived to be AI-generated. In an image evaluation study, participants rated appeal and realism highest for unlabeled human-made photographs and lower for AI-generated images, with significant differences among AI image models (Göring et al., 2023). In another study, participants rated realism and sexual arousal elicited by unlabeled human-made photographs of underwear or swimwear models; their sexual arousal ratings were positively correlated with their perception of realism, i.e., their assumption of human-production (Study 1 in Marini et al., 2024). Such studies can point to anti-AI bias, however, the negative evaluations could also reflect actual content characteristics.

- Experimental Labeling Paradigm: This approach presents participants with identical content, labeled either as AI-generated or human-generated, and compares their evaluations across dimensions such as quality, aesthetic appeal, or credibility. For example, Study 2 in Marini et al. (2024) found that participants rated identical images as less sexually arousing when, by experimental variation, a label led them to believe the images were AI-generated versus human-generated. As labeling experiments use identical content, any differences in evaluation can be ascribed solely to source attribution rather than actual content differences—thereby isolating evaluator bias.

Among these three methodological approaches, the experimental labeling paradigm provides the highest internal validity for examining anti-AI bias and is therefore the focus of the following state-of-research summary. A literature search in Google Scholar, Web of Science, and Scopus identified 15 peer-reviewed journal publications (2020–2025) reporting 27 experiments using the labeling paradigm to test for anti-AI or pro-AI bias across different content domains (see Table 1).

Across content domains and publications, previous research stresses the context dependence of bias toward AI-generated content and therefore uses variations of stimuli to operationalize the independent variable (e.g., art images with different motifs, news articles with different topics; see Stimuli column in Table 1). Most often, experimental studies compare content labeled as AI-generated versus human-generated with a two-level between-subjects or within-subjects factor. But some studies also test for collaborative scenarios where a human and an AI allegedly co-created the presented content and, hence, use higher-level factors (see Label Manipulation column in Table 1).

The selection of dependent variables varies across content domains and publications, including both single item measures and psychometric scales, with the typical answer format being rating scales (see Dependent Variables column in Table 1). Some evaluation dimensions are context-independent, e.g., when participants are asked to rate how much they "like" the content or how they evaluate its "quality". Other evaluation dimensions are content-dependent, e.g., when participants rate the "beauty" of artistic content or the "accuracy" of journalistic content.

Several previous labeling experiments included context-specific moderator and/or mediator variables, such as attitudes toward creativity in art-related studies or political opinions in news-related studies. The moderator variable most often included across content domains was AI attitudes, i.e., the general tendency of individuals to evaluate AI negatively or positively overall (Grassini, 2023). Typically, negative AI attitudes are associated with anti-AI bias, as shown in 6 of the 15 identified publications (Ansani et al., 2025; Bellaiche et al., 2023; Horton et al., 2023; Marini et al. 2024; Lim & Schmälzle, 2024; Wischneski & Krämer, 2024).

Of the 15 journal articles, 11 found evidence for an anti-AI-bias, 2 for a pro-AI bias and 2 reported no consistent effects (see Main Effect column in Table 1). Regarding standardized effect size measures, 6 papers reported Cohen's *d* values, which mainly indicated small- to medium-sized anti-AI bias.

The Current Study

While experimental research on anti-AI bias has expanded across domains such as art, health, and journalism, intimacy-related contexts remain underexplored. To date, only one published labeling experiment has examined sexualized imagery, finding that AI-labeled swimwear or underwear model images were rated as less sexually arousing than identical images labeled as human-made photographs with a medium effect size (Experiment 2 in Marini et al., 2024; see Table 1). Romantic and sexual content is particularly sensitive to perceptions of humanness, empathy, and trust, suggesting that AI involvement may evoke stronger negative machine heuristic and anti-AI bias than in other domains (Liu et al., 2025; Nass & Moon, 2000; Rubin et al., 2025).

To extend the experimental labeling paradigm into the domain of intimacy, we selected two types of content: romantic couple images and excerpts from sexuality-related couple counseling. These two stimuli types were selected based on several criteria such as relevance, realism, and acceptance: In the realm of AI-supported sexual activities, AI-generated sexually explicit and erotic material plays an important role (Döring et al., 2025a; Lapointe et al., 2025). At the same time, using AI chatbots as personal coaches and counselors for different issues including sexual and romantic ones has also been documented as a popular and promising activity (Döring et al., 2025a; Hatch et al., 2025; Vowels et al., 2024). To ensure high acceptability of our study for adult participants aged 18–75 years from the general population in Germany, we did not use sexually explicit images and dialogues. Instead, we chose romantic couple image (i.e.,

Table 1 Overview of 15 peer-reviewed publications (2020–2025) reporting on 27 experiments using the experimental labeling paradigm to test for anti-AI or pro-AI bias across four content domains (source: compiled by authors)

Publication Number	Citation	Sample Size(s)	Stimuli	Label Manipulation	Dependent Variable(s)	Main Effect	Effect Size(s) *d*
Art and Creativity							
1	Ansani et al. (2025)	$N = 120$ (1 E)	3 classic piano music performance videos (about 1 min.)	"AI" vs. "Pianist"	aesthetic judgments (e.g., liking, quality) [visual analogue scale ratings 1–100]	Anti-AI	$d = 0.68$ (liking) $d = 0.81$ (quality)
2	Bara et al. (2025)	$N_1 = 100$ (3 E)	40 impressionist AI art images of landscapes and people imitating the style of Spanish artist Joaquín Sorolla	„AI" plus information about how generative AI works vs. "AI" without further information	aesthetic and moral judgments (e.g., aesthetic appreciation, moral acceptance) [ratings 1–5]	Anti-AI (when informed about how generative AI works)	/
3	Bellaiche et al. (2023)	$N_1 = 149$ $N_2 = 148$ (2 E)	30 AI art images in representational and abstract painting styles	"AI-created" vs. "Human-created"	aesthetic judgments (e.g., liking, beauty, profundity) [ratings 1–5]	Anti-AI	$d_1 = 0.17$ (liking) $d_1 = 0.22$ (beauty) $d_1 = 0.47$ (profundity)
4	Horton et al. (2023)	$N_1 = 119$ $N_2 = 415$ $N_3 = 405$ $N_4 = 789$ $N_5 = 709$ $N_6 = 527$ (6 E)	28 images of paintings in different artistic styles	"AI-made" vs. "Human-made"	aesthetic judgments (e.g., liking, skillful, inspiring) [ratings 1–7]	Anti-AI	$d_1 = 0.25$ (liking) $d_1 = 0.22$ (inspiring) $d_1 = 0.48$ (skillful)
5	Marini et al. (2024)	$N_2 = 108$ (2 E)	120 photos of male (60) and female (60) models in swimwear/underwear	"images generated by AI" vs. "photos taken by professional photographers"	perceived level of sexual arousal [ratings 1–6]	Anti-AI	$d_2 = 0.43$ (sexual arousal)
6	Ragot et al. (2020)	$N = 565$ (1 E)	40 impressionist-style paintings of landscapes and people	"created by AI" vs. "created by artists"	aesthetic judgments (e.g., liking, beauty, meaning) [ratings 1–7]	Anti-AI	$d = 0.06$ (liking) $d = 0.05$ (beauty) $d = 0.13$ (meaning)
7	Shank et al. (2023)	$N_{2a} = 399$ $N_{2b} = 136$ $N_{3a} = 393$ $N_{3b} = 136$ (5 E)	8 musical clips (15 s.) of electronic music and classical music	"created by an AI composing software/e.g. MelodyBot" vs. "created by various composers/e.g. Morgan Walker" vs. no source information [control condition]	Aesthetic judgments (e.g., liking, musical quality, arousal) [ratings 1–5 and 1–7]	No consistent effects	/

Table 1 (continued)

Publication Number	Citation	Sample Size(s)	Stimuli	Label Manipulation	Dependent Variable(s)	Main Effect	Effect Size(s) *d*
Health and Science							
8	Karinshak et al. (2023)	$N_3 = 1{,}496$ (3 E)	2 COVID-19 pro-vaccination messages	„AI " vs. „CDC [Centers for Disease Control]" vs. "Doctor" vs. No Label	perceived message quality (e.g., argument strength, message effectiveness) [ratings 1–5]	Anti-AI	/
9	Lim and Schmälzle (2024)	$N_1 = 142$ $N_2 = 183$ (2 E)	2 nicotine vaping prevention messages	"AI-generated message" vs. "Human-generated tweet"	perceived message effectiveness [ratings 1–5]	Anti-AI	/
10	Reis et al. (2024)	$N_1 = 1{,}050$ $N_2 = 1{,}230$ (2 E)	4 scenarios with medical advice for different patient cases (smoking cessation, colonoscopy, agoraphobia, reflux disease)	"AI" vs. "human physician" vs. "human + AI"	Perceived medical advice quality (e.g., empathy, reliability) [ratings 1–5 and 1–7]	Anti-AI	$d_1 = 0.27$ (empathy) $d_1 = 0.28$ (reliability)
News and Information							
11	Cloudy et al. (2022)	$N_1 = 235$ $N_2 = 279$ (2 E)	2 news articles on polarizing topics (abortion, COVID-19 vaccination)	"AI journalist: Newsbot, CNN" vs. "Human: Quinn Smith, CNN"	machine heuristic scale (e.g., news production as objective, unemotional, reliable) [ratings 1–7]	Pro-AI	/
12	Jia et al. (2024)	$N = 269$ (1 E)	1 news article on sweetener Aspartame as possible cancer cause	„staff writer " vs. „staff writer with AI tool" vs. "staff writer + AI assistant" vs. "staff writer + AI collaborator" vs. "AI author"	message credibility scale (e.g., accurate, authentic, believable) [ratings 1–7]	Anti-AI	/
13	Waddell (2024)	$N = 612$ (1 E)	2 news articles (Trump dispute with Khan; Paris climate accord)	"Automated Insights [AI]" vs. "Kelly Richards, reporter [Human]" vs. "Kelly Richards, reporter, and Automated Insights" [Human-AI-Collaboration]	message credibility scale (e.g., accurate, authentic, believable) [ratings 1–7]	Pro-AI	/
14	Wischneski & Krämer (2024)	$N = 428$ (1 E)	2 news articles pro versus against gender-neutral language in Germany	"generated by an algorithm named AlgoInfo" vs. "written by a human staff reporter, Kim Haas"	message credibility scale (e.g., accurate, authentic, believable) [ratings 1–7]	No Effect	/

Table 1 (continued)

Publication Number	Citation	Sample Size(s)	Stimuli	Label Manipulation	Dependent Variable(s)	Main Effect	Effect Size(s) *d*
Social and Moral Action							
15	Monahan et al. (2025)	$N_3 = 307$ (3 E)	2 donation appeal messages for an animal shelter with a dog image	"AI-generated" vs. "no information [= assumed human-generated photo]"	donation likelihood [ratings 1–5] anticipated donation amount [US$ 0-US$ 500]	Anti-AI	/

The table is structured alphabetically by content domains and by first authors within each domain. The number of experiments per paper is reported in the Sample Size(s) column, ranging from 1 to 6 experiments (1 E-6 E). Sample sizes (e.g., N_2 = sample size of experiment 2) and study characteristics are only reported for those experiments from the respective publications that followed the labeling paradigm and used label manipulation as the experimental variation. Where reported in the publications, Cohen's *d* effect size measures are included in the *Main Effect* column, conventionally classified as small (0.20), medium (0.50), or large (0.80). The experiment to which an effect size *d* refers is indicated in subscript (e.g., d_2 = effect size from experiment 2). In Marini et al. (2024), the results in Table 2 (p. 356) are mistakenly presented in the wrong order: the AI results in Study 2 are the lower values for male and female participants, as confirmed in personal communication with authors and consistent with the figure on the same page of the publication

representations of hugging, clothed couples) and sexuality-related couple counseling excerpts (i.e., dialogues addressing reduced sexual desire in long-term relationships).

As the study was conducted in Germany, cultural contextualization is relevant: Public engagement with AI in Germany is characterized by relatively high usage but often comparatively low trust—a broader tendency often referred to as *German* "Angst"—that fosters cautious and risk-aware evaluations of new technologies. This contrasts with more optimistic technology adoption cultures, such as those in the U.S. and China, where perceived benefits of AI are more pronounced (Brauner et al., 2024; Richter et al., 2025).

In line with previous research and its aforementioned most commonly used domain-independent moderator variable, we included AI attitudes (i.e., the general tendency of individuals to evaluate AI negatively or positively overall; Grassini, 2023) in our study. Prior studies summarized above show that people with negative AI attitudes are more likely to display an anti-AI bias. In addition, we included AI literacy as a moderator variable. AI literacy captures individuals' knowledge, skills, and critical awareness regarding AI systems (Carolus et al., 2023). Higher AI literacy enables people to understand how AI functions, which should reduce reliance on stereotypes or fears about AI. Conversely, low literacy may leave individuals more susceptible to heuristic or biased AI judgments.

Consequently, the current study tests the following hypotheses in two consecutive experimental tasks with a shared participant sample:

H_{1a}: Romantic couple images labeled as AI-generated are evaluated more negatively than those labeled as human-generated.

H_{1b}: The negative effect of AI-generated labeling on the evaluation of romantic couple images is moderated by AI attitudes and AI literacy, with more negative AI attitudes and lower AI literacy being associated with greater anti-AI bias.

H_{2a}: Excerpts from couple counseling dialogues labeled as involving an AI counselor are evaluated more negatively than those labeled as involving a human counselor.

H_{2b}: The negative effect of AI counselor labeling on the evaluation of excerpts of couple counseling dialogues is moderated by AI attitudes and AI literacy, with more negative AI attitudes and lower AI literacy being associated with greater anti-AI bias.

Method

In two related online experiments, participants evaluated romantic couple images (Experiment 1) and sexuality-related couple counseling excerpts (Experiment 2). To ensure transparency and reproducibility, the instrument (in German and in English translation), the data file, the R analysis script,

and a supplementary table are publicly available at https://osf.io/mz4gp/.

Participants

Participants aged 18–75 residing in Germany were recruited in November and December 2024 via an incentivized online panel managed by Bilendi, a certified provider of market and social research services. Panel members voluntarily join through a double opt-in process and are invited to online studies in exchange for small monetary rewards (typically €0.50–1.00). Bilendi complies with ISO 20252:2019 standards, the European Union General Data Protection Regulation, and German data protection laws. The company applies strict quality controls throughout the panel lifecycle, including multi-source recruitment, behavioral monitoring, and regular updates, as outlined on their website.[1]

To approximate the internet-using population aged 18–75 in Germany, Bilendi applied an uncrossed quota sampling approach based on age, gender, education, marital status, and federal state. A total of 85,136 panel members were invited to take part in our study, of whom 4,780 (5.6%) accessed the study during the two-week fieldwork period. Of those, 253 were screened out due to exclusion criteria (e.g., under the age of 18), 150 due to quality criteria, and 1,085 due to the quota being full. A further 524 did not complete the survey, leaving 2,768 participants. From those, we filtered 110 participants with a quality indicator lower than 0.14 that indicates implausibly short completion times. This cleaning step had no impact on the original quota structure. The final sample comprised 2,658 participants, with an average age of 48.7 years ($SD = 15.4$); 49.7% self-identified as women. Selected sociodemographic sample characteristics (gender, age, education, and marital status) as well as lifetime self-reported use of AI in professional and private life are displayed in Table 2. The size and sociodemographic composition of the sample were determined by the requirements of the survey component rather than the experimental component of the study; the sample composition reflects the distribution of the online population in Germany (Döring et al., 2025b).

Measures

Measures are presented according to their role in the two experiments as independent, dependent, and moderator variables.

[1] https://www.bilendi.com/, last accessed: June 16, 2025.

Table 2 Sociodemographic characteristics of online survey participants in Germany (N = 2,658), absolute and relative frequencies

Characteristic	Participants	
	n	%
Gender[a]		
Women	1,320	49.7
Men	1,338	50.3
Age (in years)		
18–39	891	33.5
40–59	961	36.2
> 60	806	30.3
Education		
Low	843	31.7
Moderate	794	29.9
High	1,021	38.4
Marital status		
Unmarried	1,202	45.2
Married	1,456	54.8
AI use		
AI use in job and vocational training	702	26.4
AI use in private life	1,119	45.1

[a]At the time of data collection, the panel provider could only supply representative quotas for participants identifying as women or men. Consequently, gender-diverse individuals were not systematically recruited, although they were not intentionally excluded. This methodological constraint is acknowledged in the limitations section

Independent Variables and Stimulus Material

The two experiments followed the experimental labeling paradigm in anti-AI bias research: To test for an anti-AI bias, participants evaluated identical content that was either labeled as AI-generated or human-generated (two-level between-subjects factor). Both the romantic couple images and the sexuality-related couple counseling excerpts were designed to be realistic, non-explicit, and broadly acceptable as confirmed in a pretest with five participants.

To avoid bias from individual images, two similar romantic couple images were used as stimuli and assigned to participants randomly with random assignment of the AI-generated or human-generated label in the first experiment. The images showed clothed couples hugging in everyday winter settings, matched for composition, tone, expression, and quality, and could plausibly be created by either a human photographer or an AI application (see Fig. 1). No statistical differences were found in respondents' evaluations of the two couple images (see supplementary Table S1 at https://osf.io/mz4gp/).

In the second experiment, two similar sexuality-related couple counseling excerpts were used as stimuli and assigned to participants randomly with random assignment of the AI-generated or human-generated label in the second

Fig. 1 Stimulus material for Experiment 1: Romantic couple image. For the experimental variation, participants were randomly assigned to one of the two images with one of the two labels: "The following picture was created by a photographer [by an AI application]. Please look at the photo [AI-generated image] and then rate it." In this study, the couple images depicted mixed-gender-presenting couples. This pragmatic choice was made to maximize familiarity for a general population sample and is addressed in the Limitations section

Table 3 Stimulus material for Experiment 2: Couple counseling excerpts

Person seeking advice: I feel that my desire for intimacy has decreased lately, and I'm worried that this is putting a strain on our relationship.	Person seeking advice: Our sex life has felt very routine lately, and I don't know how to bring it up without it sounding hurtful.
Counselor: This is a very sensitive topic, and I understand that it's important to you. Do you feel that external factors, such as stress or routine, might be playing a role here?	Counselor: It's completely normal for a relationship to settle into a routine over time. Do you feel that both of you are experiencing it this way?
Person seeking advice: Yes, absolutely. I'm often tired or stressed, and sometimes it feels like intimacy is just another "obligation" in daily life.	Person seeking advice: I'm not sure. But for me, it's an issue, and I'd like to have more variety again.
Counselor: That's a common feeling when you're under pressure. How about trying to take the pressure off and create small, relaxed moments of closeness without it having to be about sexual intimacy right away?	Counselor: It might be helpful to frame this wish as an opportunity for new shared experiences. Maybe you could talk about what you both especially enjoyed in the past. Do you think that might be well received?
Person seeking advice: That sounds good. Sometimes I feel like it's an "all-or-nothing" topic, and that puts me under pressure.	Person seeking advice: Yes, I can imagine that. I think if I present it as a "we" topic, it would be easier.
Counselor: Exactly. If you focus on spending time together without a clear expectation, it can often bring back the space for natural desire. It might also help to find activities together that aren't sexual but still foster closeness—a walk together or a small surprise in everyday life can make a big difference.	Counselor: Exactly. Sometimes it helps to talk about what makes you both curious without focusing right away on "success." That way, playful new ideas can emerge to break the routine and bring fresh energy—without any pressure.

For the experimental variation, participants were randomly assigned to one of the two couple counseling excerpts with one of the two labels: "The following brief sexual counseling session was conducted by a trained sexual education professional [by an AI chatbot]. Please read the counseling provided by the professional [the AI counseling] and then evaluate it."

In this study, the couple counseling excerpts were phrased in gender-neutral language

experiment. The counseling excerpts depicted authentic dialogues in which a client discusses issues of reduced sexual desire in a long-term relationship with a counselor, using evidence-based techniques such as normalization and open-ended questioning. Both counseling excerpts were matched for length, emotional engagement, and complexity, and could plausibly involve a human counselor or an AI counselor (see Table 3). No statistical differences were found in respondents' evaluations of the two couple counseling excerpts (see supplementary Table S1 at https://osf.io/mz4gp/).

Dependent Variables

Evaluations of the romantic couple images were measured on four image-oriented items ("aesthetic," "erotic," "expressive," "vivid"), each answered on 5-point rating scales (1 = strongly disagree 5 = strongly disagree). The additive index of the four items for Image Evaluation (IE index) demonstrated good internal consistency (Cronbach's $\alpha = .80$; McDonald's $\omega = .82$) according to established internal consistency reliability norms (e.g., Kalkbrenner, 2023).

Evaluations of the sexuality-related couple counseling excerpts were measured on four counseling-oriented items ("helpful," "understanding," "empathetic," "competent") to

be answered on 5-point rating scales (1 = strongly disagree, 5 = strongly disagree). The additive index of the four items for Counseling Evaluation (CE index) showed good internal consistency (Cronbach's α = .93; McDonald's ω = .95).

Moderator Variables

AI attitudes were measured with the AI Attitude Scale (AIAS-4; Grassini, 2023) comprising of four items (e.g., "I believe that AI will improve my life," "I think AI technology is positive for humanity") to be answered on 10-point rating scales (1 = not at all; 10 = completely agree). The scale demonstrated good internal consistency (Cronbach's α = .95; McDonald's ω = .95).

AI literacy was measured with a short form of the Meta AI Literacy Scale (MAILS; Carolus et al., 2023) comprising of twelve items selected by the authors representing AI use (e.g., "I can operate AI applications"), AI knowledge (e.g., "I can assess what the limitations and opportunities of using an AI are"), ability to detect AI (e.g., "I can distinguish devices that use AI from devices that do not"), and AI ethics (e.g., "I can analyze AI-based applications for their ethical implications") to be answered on 11-point answer scales (0 = ability is not at all or hardly pronounced; 10 = ability is very well or (almost) perfectly pronounced).[2] The short scale showed good internal consistency (Cronbach's α = .97; McDonald's ω = .98).

AIAS-4 and our short version of MAILS were highly correlated (Pearson's $r = .68$). Therefore, we combined all items from both scales into a single AI Readiness scale, with higher scores indicating both more positive AI attitudes and greater AI literacy. The scale demonstrated good internal consistency (Cronbach's α = .98; McDonald's ω = .98).

Procedure

Participants first gave informed consent and then completed a questionnaire on their AI experiences including AI attitudes and AI literacy. After that, they were randomly assigned to the conditions of the two consecutive experiments. In the first experiment, participants were randomly assigned to one of the two romantic couple images and one of the two image-labeling conditions (image labeled as AI-generated versus human-generated). They were instructed to view the image and provide their image evaluation on rating scales. In the second experiment, they were randomly assigned to one of the two sexuality-related couple counseling excerpts and one of the two excerpt-labeling conditions (counseling excerpt labeled as involving an AI counselor versus a human counselor). They were instructed to read the excerpt and provide their counseling evaluation on rating scales. After the two experimental tasks, participants were debriefed, thanked, and dismissed.

Data Analysis

Given the large sample size and high statistical power, the significance level was set to α = .01, and the threshold for a meaningful effect was set at $\eta^2 \geq .01$, representing at least a small effect according to Cohen's benchmarks (Cohen, 1988). Data analysis was conducted with R 4.4.2 (packages afex 1.4–1, car 3.1–3, dplyr 1.1.4, expss 0.11.6, haven 2.5.5, labelled 2.14.1, lmtest 0.9–40, nortest 1.0–4, psych 2.5.6, sandwich 3.1–1, sjPlot 2.8.17, tidyverse 2.0.0). To test the four research hypotheses for the two experiments, *t*-tests and ANCOVAs were conducted, both of which assume normality of residuals and homogeneity of variances. Given the large sample size, the central limit theorem supports the assumption of normality for the sampling distributions. Homogeneity of variances was tested using Levene's test and showed only minimal deviations that did not reach the set level of significance (for couple images: $F(1, 2656) = 5.96$, $p = .015$; for couple counseling excerpts: $F(1, 2656) = 1.45$, $p = .119$). As a robustness check against potential heteroscedasticity, we repeated all mean comparisons using Welch's *t*-tests and estimated ANCOVAs with heteroskedasticity-consistent standard errors; the conclusions were unchanged.

Results

Results are presented separately for the two experiments.

Anti-AI Bias Toward Couple Images

Participants rated the romantic couple images labeled as AI-generated significantly more negatively than identical images labeled as human-generated, with a small effect size that explained about 1% of the variance ($d = .21$; see Table 4), supporting H_{1a}. This anti-AI bias toward AI couple images was moderated by AI readiness, as shown in the ANCOVA interaction effect between the label condition and AI readiness, with a small effect size that explained 1% of the variance ($\eta^2 = .01$; see Table 5), providing support for H_{1b}.

Anti-AI Bias Toward Couple Counseling

Participants rated the sexuality-related couple counseling excerpts labeled as AI-generated significantly more negatively than identical counseling excerpts labeled as human-generated, with a small effect size that explained about 1% of the variance ($d = .23$; see Table 4), supporting H_{2a}. This anti-AI bias toward AI couple counseling was not moderated by AI readiness, as shown in the ANCOVA interaction effect

[2] https://hci.uni-wuerzburg.de/research/MAILS/

Table 4 Evaluations of romantic couple images and couple counseling excerpts labeled as human-generated versus AI-generated

Content type	Human-generated			AI-generated							
	M	*SD*	*n*	*M*	*SD*	*n*	Δ*M*	*df*	*t*	*p*	Cohen 's *d*
Romantic Couple Images	3.62	0.79	1,356	3.45	0.86	1,302	0.17	2,656	5.39	< .001	.21
Couple Counseling Excerpts	3.64	0.95	1,302	3.42	0.99	1,356	0.22	2,656	5.83	< .001	.23

N = 2,658. Both evaluation indices range from 1 (low evaluation) to 5 (high evaluation). The delta value (Δ*M*) represents the anti-AI Bias with higher values representing a larger bias

Table 5 Analysis of covariance for evaluations of romantic couple images labeled as human-generated versus AI-generated, controlling for AI readiness

	df	*SS*	*MS*	*F*	*p*	η^2
AI Readiness (Covariate)	1	132.8	132.8	212.1	< .001	.074
Couple Image Label	1	21.4	21.4	34.2	< .001	.012
Couple Image Label x AI Readiness (Moderation)	1	8.9	8.9	14.2	< .001	.005
Error	2,654	1,661.4	0.6			

N = 2,658. The image evaluation index ranges from 1 (low evaluation) to 5 (high evaluation). Pearson's *r* (AI Readiness, Image Evaluation Index) = .27

Table 6 Analysis of covariance for evaluations of couple counseling excerpts labeled as human-generated versus AI-generated, controlling for AI readiness

	df	*SS*	*MS*	*F*	*p*	η^2
AI Readiness (Covariate)	1	308.4	308.4	374.7	< .001	.124
Couple Counseling Label	1	24.5	24.5	29.7	< .001	.011
Couple Counseling Label x AI Readiness (Moderation)	1	6.4	6.4	7.8	.005	.003
Error	2,654	2,184.4	0.8			

Note. *N* = 2,658. The couple counseling evaluation index ranges from 1 (low evaluation) to 5 (high evaluation). Pearson's *r* (AI Readiness, Counseling Evaluation Index) = .35

between the label condition and AI readiness that showed negligible explained variance ($\eta^2 = .003$; see Table 6), providing no support for H_{2b}. However, a descriptive analysis illustrates that participants with AI readiness below the median consistently exhibited a larger anti-AI bias than respondents with above median AI readiness, for both images and counseling excerpts (see Table 7).

Discussion

The discussion covers the interpretation of the findings, study limitations, and the conclusion.

Interpretation

Our study, based on two consecutive online experiments, confirmed a consistent but small anti-AI bias for romantic couple images and sexuality-related couple counseling dialogues among a national sample of respondents residing in Germany (see Table 2). Considering that (a) we chose intimacy-related stimuli, which might elicit a stronger negative machine heuristic than stimuli more closely related to rationality (e.g., news; see Table 1), and (b) we included participants from Germany, a country with a tradition of technology skepticism, one might have expected a stronger anti-AI bias. However, our effect sizes ($d = .21$ and $d = .23$) roughly mirror those reported in earlier studies in non-intimate content domains (e.g., Bellaiche et al., 2023 Horton et al., 2023; Reis et al., 2024).

AI readiness—the combination of positive AI attitudes and higher AI literacy—moderated the anti-AI bias: Participants with higher AI readiness exhibited a smaller anti-AI bias, with the moderation reaching a relevant size of 1% explained variance for images. One explanation for these findings could be that generative AI is increasingly normalized in everyday life, including in personal and relational domains, which may reduce resistance even in contexts traditionally associated with authenticity and emotional connection. The moderating effect of AI readiness further indicates

Table 7 Evaluation of romantic couple images and couple counseling excerpts labeled as human-generated versus AI-generated by participants with low AI readiness versus high AI readiness (median split)

	Low AI Readiness					High AI Readiness				
	Human-Generated		AI-generated			Human-Generated		AI-generated		
	M	*SD*	*M*	*SD*	Δ*M*	*M*	*SD*	*M*	*SD*	Δ*M*
Romantic Couple Images	3.50	0.82	3.23	0.89	0.27	3.74	0.72	3.65	0.72	0.09
Couple Counseling Excerpts	3.45	1.02	3.10	1.02	0.35	3.84	0.82	3.72	0.86	0.12

$N = 2{,}658$. Both evaluation indices range from 1 (low evaluation) to 5 (high evaluation). The delta value (Δ*M*) represents the anti-AI bias with higher values showing a larger bias. This descriptive median-split analysis is reported solely for illustrative purposes and should be interpreted with caution, as dichotomizing continuous variables can inflate group differences

that knowledge of, and positive attitudes toward AI can buffer against such negative bias.

Limitations

Our study benefits from its grounding in the experimental labeling tradition and its large national sample. Nonetheless, several limitations should be acknowledged.

1. Sampling: At the time of data collection, the panel provider Bilendi could only provide representative quotas for participants identifying as women or men, limiting the inclusion of gender-diverse perspectives. Panel providers are currently working on more gender-inclusive quota and recruitment plans to overcome this limitation in the future.

2. Independent Variables and Stimulus Material: We were able to work with gender-neutrally phrased couple counseling excerpts as stimulus material, but this was not possible for couple images. We made the pragmatic choice to work with images of mixed-gender-presenting couples to maximize familiarity for a general population sample. Future work should employ more diverse stimulus images that vary in gender composition and expression, and examine how these interact with evaluators' own gender and sexual identities. While our stimuli were intimacy-related, they were not sexually explicit—potentially yielding a smaller anti-AI bias than might occur with more erotic content (see Experiment 2 in Marini et al., 2024) or with pornographic imagery that will be investigated in the future.

3. Experimental Manipulation: We presented the labels identifying the stimuli as AI-generated or human-created directly together with the stimuli to ensure that participants noticed them. Given the clarity of the labels, we judged a separate manipulation check unnecessary. If some participants had overlooked, misunderstood, or doubted the label, this would have attenuated the experimental effect rather than inflated it. The fact that consistent label differences emerged across both experiments therefore suggests that the observed effects are robust, even under potentially attenuating conditions.

4. Dependent Variables: For image and counseling evaluations, we used two sets of four domain-specific items; future studies could expand the range of evaluation dimensions.

5. Moderator Variables: We used items from established psychometric scales to measure AI attitudes (AIAS-4, Grassini, 2023) and AI literacy (MAILS, Carolus et al., 2023). However, in line with most AI literacy measures developed so far, MAILS does not objectively test AI literacy, but relies on self-reported knowledge and skills with some overlap with attitudes. Future AI bias studies could use objective AI literacy tests.

6. Design: As both experiments were conducted consecutively and in a fixed order with the same sample, their results are not independent, and order effects are possible, although their magnitude and direction cannot be predicted. Future multiple-experiment studies can use multiple samples or randomize the order of experimental tasks within the same sample, an approach that was not implemented here due to logistical restrictions.

Conclusion

Our data show that anti-AI bias in intimacy-related contexts is prevalent in Germany, although its size is small. Our data further highlight the importance of moderating factors such as AI literacyand AI attitudes. Greater exposure and familiarity with AI systems can foster more positive attitudes and higher literacy, thereby reducing anti-AI bias in the future. From a theoretical perspective, greater personal experience with AI in intimacy- and sexuality-related domains may not only reduce the negative machine heuristic but also elicit a positive machineheuristic (Molina & Sundar, 2024; Sundar & Kim, 2019). Signs of such a shift are already visible in current debates, particularly when AI's non-human, fact-based, controllable, and rational qualities are framed as advantages.

These qualities may be beneficial in intimate contexts as well. For example, where human sexual counselors vary in knowledge, hold moralizing attitudes, or are less accessible, clients can turn to AI systems as low-threshold and potentially safer alternatives. In addition, human professionals and AI systems can collaborate in an efficient and meaningful way through hybrid counseling scenarios that combine constant 24/7, low-cost AI support with the option of transferring to human counselors whenever needed.

At the same time, however, increasing use of AI in intimate, romantic, and sexual contexts may trigger critical responses or even resistance and backlash. Illinois, for example, was the first U.S. state to ban AI from mental health care (Illinois General Assembly, 2025), potentially reflecting anti-AI bias. The German Ethics Council has emphasized that while AI can lower access barriers and reduce the stigma of psychotherapy, it cannot replace core elements such as trust, empathy, and authentic relational experiences with human professionals (Deutscher Ethikrat, 2023). Media reports of individuals forming strong attachments to AI chatbots that allegedly caused divorce, fatal accidents, or suicide (e.g., Brittain, 2025; Fike, 2025; Horwitz, 2025) have further fueled skepticism, even if causal explanations in these cases may be more nuanced. Together, these examples underline that openness and rejection, pro-AI and anti-AI biases, are likely to evolve in parallel rather than in a simple linear trajectory.

The way people apply negative or positive heuristics to AI is also shaped by their conceptualization of the technology itself. Research shows that technical explanations can reduce acceptance (Bara et al., 2025), whereas embodied demonstrations can enhance it (Chamberlain et al., 2018). To expand on our findings regarding the moderating role of AI literacy and AI attitudes, future research could examine in greater detail how the conceptualization of generative and conversational AI models (e.g., whether people believe they truly "think" and "understand", or view them as mere "stochastic parrots") influences AI evaluations and biases.

There remains substantial scope to further investigate the benefits and challenges of AI in the domain of sexuality, taking into account subjective experiences, diverse perspectives, and potential biases. While the present study focused on individual-level perceptions, future research should complement this with system-level analyses that address economic, legal, and societal dimensions (e.g., Anciaux & Gramaccia, 2025). Integrating these levels of analysis will be essential for a comprehensive understanding of how AI shapes intimate life and its broader implications.

Author contributions N.D. conceived and designed the research, coordinated the project, contributed to the development of the research instrument and the statistical analysis plan, and wrote the first draft of the manuscript. M.R.M. coordinated data collection and performed the statistical analyses. Both authors contributed to the interpretation of the findings, participated in reviewing and editing the manuscript, and approved the final version for submission.

Funding Open Access funding enabled and organized by Projekt DEAL. No funding was received for conducting this study.

Data availability The study instrument, data, and analysis script are publicly available at https://osf.io/mz4gp/.

Declarations

Conflict of interest The first author is one of the guest editors of the Archives of Sexual Behavior special section "Artificial Intelligence and Sexuality". The second author has no conflict of interest.

Ethical approval The study, which comprised the experimental component reported here and a survey component reported elsewhere (Döring et al., 2025b) was approved by the ethics committee of Technische Universität Ilmenau, and all participants provided informed consent.

Informed consent All participants gave informed consent.

Declaration of generative AI in scientific writing We included this AI declaration in response to the Guest Editors' request for the special issue on "Artificial Intelligence and Sexuality." The first author wrote the initial draft of the paper, with all ideas originating from her and being informed by prior research on AI-supported sexual activities, sexual content, and anti-AI bias. ChatGPT-4 was used for editing and translation purposes with prompts such as "translate to English," "improve flow," "proofread," or "shorten."

References

Alilunas, P. (2024). What we must be: AI and the future of porn studies. *Porn Studies, 11*(1), 99–112. https://doi.org/10.1080/23268743.2024.2312181

Anciaux, A., & Gramaccia, J. A. (2025). Imagining markets and crafting value: the emergence of an AI-generated pornographic content ecosystem. *Porn Studies*. https://doi.org/10.1080/23268743.2025.2492332

Ansani, A., Koehler, F., Giombini, L., Hämäläinen, M., Meng, C., Marini, M., & Saarikallio, S. (2025). AI performer bias: Listeners like music less when they think it was performed by an AI. *Empirical Studies of the Arts, 43*(2), 1137–1161. https://doi.org/10.1177/02762374241308807

Bandura, A. (2018). Toward a psychology of human agency: Pathways and reflections. *Perspectives on Psychological Science, 13*(2), 130–136. https://doi.org/10.1177/1745691617699280

Bara, I., Ramsey, R., & Cross, E. S. (2025). AI contextual information shapes moral and aesthetic judgments of AI-generated visual art. *Cognition, 257*, Article 106063. https://doi.org/10.1016/j.cognition.2025.106063

Bellaiche, L., Shahi, R., Turpin, M. H., Ragnhildstveit, A., Sprockett, S., Barr, N., Christensen, A., & Seli, P. (2023). Humans versus AI: Whether and why we prefer human-created compared to AI-created artwork. *Cognitive Research: Principles and Implications, 8*(1), Article 42. https://doi.org/10.1186/s41235-023-00499-6

Brauner, P., Glawe, F., Liehner, G. L., Vervier, L., & Ziefle, M. (2024, December 18). *AI perceptions across cultures: Similarities and differences in expectations, risks, benefits, tradeoffs, and value in Germany and China*. http://arxiv.org/pdf/2412.13841v1

Brittain, B. (2025, May 21). *Google, AI firm must face lawsuit filed by a mother over suicide of son, US court says.* https://www.reuters.com/sustainability/boards-policy-regulation/google-ai-firm-must-face-lawsuit-filed-by-mother-over-suicide-son-us-court-says-2025-05-21/

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS - Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans, 1*(2), Article 100014. https://doi.org/10.1016/j.chbah.2023.100014

Chamberlain, R., Mullin, C., Scheerlinck, B., & Wagemans, J. (2018). Putting the art in artificial: Aesthetic responses to computer-generated art. *Psychology of Aesthetics, Creativity, and the Arts, 12*(2), 177–192. https://doi.org/10.1037/aca0000136

Chatterjee, A. (2022). Art in an age of artificial intelligence. *Frontiers in Psychology, 13*, 1024449. https://doi.org/10.3389/fpsyg.2022.1024449

Cloudy, J., Banks, J., & Bowman, N. D. (2022). AI journalists and reduction of perceived hostile media bias: Replication and extension considering news organization cues. *Technology, Mind, and Behavior*, *3*. https://doi.org/10.1037/tmb0000083

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Deutscher Ethikrat [German Ethics Council]. (2023). *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz:* Stellungnahme [Human and Machine–Challenges of Artificial Intelligence: Position Paper]. Berlin: Deutscher Ethikrat. https://www.ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

Döring, N., Le, T. D., Vowels, L. M., Vowels, M. J., & Marcantonio, T. L. (2025a). The impact of artificial intelligence on human sexuality: A five-year literature review 2020–2024. *Current Sexual Health Reports*, *17*. https://doi.org/10.1007/s11930-024-00397-y

Döring, N., Mikhailova, V., & Mohseni, M. R. (2025b). *Prevalence of AI-supported sexual activities among adults in Germany: Results from a national online survey*. Manuscript submitted for publication.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., & Wright, R. (2023). Opinion paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Easterbrook-Smith, G. (2025). Pornographic aura, AI and the value of authenticity. *Porn Studies*. https://doi.org/10.1080/23268743.2025.2511311

Fike, A. (2025). *75-year-old man asks wife for divorce after falling in love with AI chatbot.* https://www.vice.com/en/article/75-year-old-man-asks-wife-for-divorce-after-falling-in-love-with-ai-chatbot/

Gondlach, K. A., & Regneri, M. (2023). The ghost of German angst: Are we too skeptical for AI development? In I. Knappertsbusch & K. Gondlach (Eds.), *Work and AI 2030* (pp. 3–10). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-40232-7_1

Göring, S., Ramachandra Rao, R. R., Merten, R., & Raake, A. (2023). Analysis of appeal for realistic AI-generated photos. *IEEE Access, 11*, 38999–39012. https://doi.org/10.1109/ACCESS.2023.3267968

Grassini, S. (2023). Development and validation of the AI Attitude Scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology, 14*, 1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Grassini, S., & Koivisto, M. (2024). Understanding how personality traits, experiences, and attitudes shape negative bias toward AI-generated artworks. *Scientific Reports, 14*(1), 4113. https://doi.org/10.1038/s41598-024-54294-4

Guingrich, R. E., & Graziano, M. S. A. (2025). P(doom) versus AI optimism: Attitudes toward artificial intelligence and the factors that shape them. *Journal of Technology in Behavioral Science*. https://doi.org/10.1007/s41347-025-00512-3

Hatch, S. G., Goodman, Z. T., Vowels, L., Hatch, H. D., Brown, A. L., Guttman, S., Le, Y., Bailey, B., Bailey, R. J., Esplin, C. R., Harris, S. M., Holt, D. P., McLaughlin, M., O'Connell, P., Rothman, K., Ritchie, L., Top, D. N., & Braithwaite, S. R. (2025). When ELIZA meets therapists: A Turing test for the heart and mind. *PLoS Mental Health, 2*(2), Article e0000145. https://doi.org/10.1371/journal.pmen.0000145

He, R., Cao, J., & Tan, T. (2025). Generative artificial intelligence: A historical perspective. *National Science Review, 12*(5), Article nwaf050. https://doi.org/10.1093/nsr/nwaf050

Horton, C. B., White, M. W., & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports, 13*(1), 19001. https://doi.org/10.1038/s41598-023-45202-3

Horwitz, J. (2025). *Meta's "flirty" AI chatbot invited a retiree to New York. He never made it home*. https://www.reuters.com/investigates/special-report/meta-ai-chatbot-death/

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. Yale University Press.

Illinois General Assembly. (2025). *Wellness and Oversight for Psychological Resources Act* (Public Act No. 104–0054) [Online Document]. Retrieved from https://www.ilga.gov/Documents/Legislation/PublicActs/104/PDF/104-0054.pdf

Jia, H., Appelman, A., Wu, M., & Bien-Aimé, S. (2024). News bylines and perceived AI authorship: Effects on source and message credibility. *Computers in Human Behavior: Artificial Humans, 2*(2), Article 100093. https://doi.org/10.1016/j.chbah.2024.100093

Kalkbrenner, M. T. (2023). Alpha, omega, and h internal consistency reliability estimates: Reviewing these options and when to use them. *Counseling Outcome Research and Evaluation, 14*(1), 77–88. https://doi.org/10.1080/21501378.2021.1940118

Karinshak, E., Liu, S. X., Park, J. S., & Hancock, J. T. (2023). Working with AI to persuade: Examining a large language model's ability to generate pro-vaccination messages. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW1), 1–29. https://doi.org/10.1145/3579592

Kernis, M. H., & Goldman, B. M. (2006). A multicomponent conceptualization of authenticity: Theory and research. *Advances in Experimental Social Psychology*, *38*, 283–357.

Lapointe, V. A., Dubé, S., Rukhlyadyev, S., Kessai, T., & Lafortune, D. (2025). The present and future of adult entertainment: A content analysis of AI-generated pornography websites. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-025-03099-1

Lim, S., & Schmälzle, R. (2024). The effect of source disclosure on evaluation of AI-generated messages. *Computers in Human Behavior: Artificial Humans, 2*(1), Article 100058. https://doi.org/10.1016/j.chbah.2024.100058

Liu, T., Giorgi, S., Aich, A., Lahnala, A., Curtis, B., Ungar, L., & Sedoc, J. (2025). The illusion of empathy: How AI chatbots shape conversation perception. *Proceedings of the AAAI Conference on Artificial Intelligence, 39*(13), 14327–14335. https://doi.org/10.1609/aaai.v39i13.33569

Marini, M., Ansani, A., Demichelis, A., Mancini, G., Paglieri, F., & Viola, M. (2024). Real is the new sexy: The influence of perceived realness on self-reported arousal to sexual visual stimuli. *Cognition and Emotion, 38*(3), 348–360. https://doi.org/10.1080/02699931.2023.2296581

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). A survey on bias and fairness in machine learning. *Acm Computing Surveys, 54*(6). https://doi.org/10.1145/3457607

Molina, M. D., & Sundar, S. S. (2024). Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation. *New Media & Society, 26*(6), 3638–3656. https://doi.org/10.1177/14614448221103534

Monahan, L., Martin, M., Zaitsev, A., Bartelt, V., & Rahman Noordeen, A. (2025). *Do I spy AI?* First Monday. Advance online publication. https://doi.org/10.5210/fm.v30i4.13799

Montag, C., Schulz, P. J., Zhang, H., & Li, B. J. (2025). On pessimism aversion in the context of artificial intelligence and locus of control: Insights from an international sample. *AI and Society, 40*(5), 3349–3356. https://doi.org/10.1007/s00146-025-02186-0

Nadarzynski, T., Bayley, J., Llewellyn, C., Kidsley, S., & Graham, C. A. (2020). Acceptability of artificial intelligence (AI)-enabled chatbots, video consultations and live webchats as online platforms for sexual health advice. *BMJ Sexual & Reproductive Health, 46*(3), 210–217. https://doi.org/10.1136/bmjsrh-2018-200271

Nadarzynski, T., Lunt, A., Knights, N., Bayley, J., & Llewellyn, C. (2023). "But can chatbots understand sex?" Attitudes towards artificial intelligence chatbots amongst sexual and reproductive health professionals: An exploratory mixed-methods study. *International Journal of STD and AIDS, 34*(11), 809–816. https://doi.org/10.1177/09564624231180777

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Ovsyannikova, D., Mello, VOde, & Inzlicht, M. (2025). Third-party evaluators perceive AI as more compassionate than expert humans. *Communications Psychology, 3*(1), Article 4. https://doi.org/10.1038/s44271-024-00182-6

Ragot, M., Martin, N., & Cojean, S. (2020). AI-generated vs. human artworks. A perception bias towards artificial intelligence? In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–10). ACM. https://doi.org/10.1145/3334480.3382892

Reis, M., Reis, F., & Kunde, W. (2024). Influence of believed AI involvement on the perception of digital medical advice. *Nature Medicine, 30*(11), 3098–3100. https://doi.org/10.1038/s41591-024-03180-7

Richter, V., Katzenbach, C., & Zeng, J. (2025). Negotiating AI(s) futures: competing imaginaries of AI by stakeholders in the US, China, and Germany. *Journal of Science Communication*, *24*. https://doi.org/10.22323/2.24020208

Rubin, M., Li, J. Z., Zimmerman, F., Ong, D. C., Goldenberg, A., & Perry, A. (2025). Comparing the value of perceived human versus AI-generated empathy. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-025-02247-w

Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2025). Generative artificial intelligence: A systematic review and applications. *Multimedia Tools and Applications, 84*(21), 23661–23700. https://doi.org/10.1007/s11042-024-20016-1

Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K., & Belfi, A. M. (2023). AI composer bias: Listeners like music less when they think it was composed by an AI. *Journal of Experimental Psychology. Applied, 29*(3), 676–692. https://doi.org/10.1037/xap0000447

Smith, N., & Southerton, C. (2025). AI and aesthetic alienation: The image and creativity in contemporary culture. *Social Science Computer Review*. https://doi.org/10.1177/08944393251361449

Storey, V. C., Yue, W. T., Zhao, J. L., & Lukyanenko, R. (2025). Generative artificial intelligence: Evolving technology, growing societal impact, and opportunities for information systems research. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-025-10581-7

Sundar, S. S., & Kim, J. (2019). Machine heuristic: When we trust computers more than humans with our personal information. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–9). ACM. https://doi.org/10.1145/3290605.3300768

Vowels, L. M., Francois-Walcott, R. R., & Darwiche, J. (2024). AI in relationship counselling: Evaluating ChatGPT's therapeutic capabilities in providing relationship advice. *Computers in Human Behavior: Artificial Humans, 2*(2), Article 100078. https://doi.org/10.1016/j.chbah.2024.100078

Waddell, T. F. (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. *Journalism & Mass Communication Quarterly, 96*(1), 82–100. https://doi.org/10.1177/1077699018815891

Wischnewski, M., & Krämer, N. (2024). Does polarizing news become less polarizing when written by an AI? *Journal of Media Psychology*. https://doi.org/10.1027/1864-1105/a000441

Zagalo, N., & Keller, D. (2026). *Artificial media: Emerging trends in narratives, education and creative practice*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-89037-6