
AI and Agency: Benefiting Users by Understanding and Serving Explicit Aspirations

Emily Xing, Ravi Iyer, Rachel Xu, Yuning Liu, Beth Goldberg,
and Nathanael J. Fast

This paper introduces a framework for understanding how generative AI affects user agency, drawing on psychological research distinguishing between expressed preferences and immediate, often-regretted desires. We conducted a study examining how AI usage relates to users' self-reported capacity for accomplishing goals along with the potential benefits of providing users with a sense of control over AI tools. This study with a nationally representative sample of U.S. generative AI users examined users' aspirational preferences for AI usage; where they perceive capability (outcome control) gains from AI; users' perceived control over AI tool experiences (process control); the relationship between perceived control over AI experiences and reported capability gains; where users locate their concerns about AI; and whether reported gains differ across demographic groups. Results indicate that users primarily both value and report benefit from AI in instrumental domains (e.g., task completion, information gathering, and learning). Aspirations for, and experiences of, benefits related to social and affective domains were less common. Users most commonly reported concerns about AI eroding independent thinking and agency. Importantly, users who felt more empowered to shape AI outputs (process control) also reported greater capability gains (outcome control), although exploratory analyses suggest these gains are unevenly distributed across demographic groups. We argue that centering user agency, encompassing control over both product experience and product outcomes, should be a primary design principle.

1 Introduction

User concerns about AI development are substantial, widespread, and often correlated with feelings of powerlessness over AI's future (Gerlich 2024). A Pew Research Center

survey found that 50% of Americans are more concerned than excited about the increased use of AI in daily life, up from 37% in 2021 (Kennedy et al. 2025). Approximately 56% of the public is extremely or very concerned about AI-related job loss (McClain et al. 2025), and 71% of U.S. adults report concern about AI fueling job displacement (Lange and Alper 2025). Neither AI usage nor concern about AI is evenly distributed, with both factors differing by gender, education, age, and race (Angrisani et al. 2026; Cranney, Delecourt, and Koning 2026; Gottfried et al. 2026).

Many have noted a tension between the short-term productivity gains that AI tools offer and their long-term implications for job security and autonomy, which can exacerbate feelings of powerlessness. Surveys of academics (Muscio and Shibayama 2025) and medical professionals (Banerjee et al. 2021) have found that professionals are excited about productivity gains but concerned about those gains coming at the cost of reduced personal capabilities due to reliance on AI. Accounting for perceived sense of control over how AI systems are used in their work may help resolve this tension. In one longitudinal study of Korean workers, perceptions of “human-centered AI,” which included perceived user control over AI systems, were found to buffer the negative effects of AI-related insecurity on subsequent workplace innovation (Kim and Lee 2026). Similarly, employees’ perceived control over how they use AI in their work has been found to buffer the negative effects of ethical concerns about generative AI on employee innovation (Bai et al. 2025). Together, these findings suggest that preserving user control over AI systems may be key to ensuring these tools support users, rather than undermine them.

To design AI systems that benefit users, developers may wish to reflect on the lessons of previous technological innovations, where a perceived lack of agency similarly contributed to widespread pessimism. In particular, AI developers may want to avoid repeating the patterns observed with social media, where engagement-driven design, optimized for attention capture rather than user control, (Milli et al. 2025; Zac et al. 2025) has contributed to not only more emotionally charged (Milli et al. 2025), extreme, and polarized content (Zimmerman et al. 2024; Brady et al. 2025), but also prolonged usage (Guess et al. 2023), widespread user regret (Terzimehić et al. 2023), feelings of manipulation (Sahebi and Formosa 2022), and diminished sense of control (Zahrai et al. 2022). To highlight the importance of agency even within social media use, research has demonstrated that, compared to more passive use, users report less regret for intentional social media usage such as direct messaging (Guo et al. 2025). Similarly, those who hold an agentic mindset toward their social media use report better psychological well-being than those who do not (Lee and Hancock 2024). Drawing on this parallel, preserving user agency in AI system design appears critical to ensuring positive outcomes for users.

AI companies have articulated commitments to being more deliberate about design that serves users’ stated intentions. OpenAI has invested in systems that remember long-term user preferences over time and has stated its aim of optimizing products for user satisfaction (Altman 2025). DeepMind CEO Demis Hassabis has cautioned that AI

risks replicating social media’s harms if developers prioritize speed over responsibility, criticizing the prevailing “move fast and break things” mentality (Spirlet 2025). Meta has conducted community forums to understand how representative samples of the public in five countries believe AI agents should behave and interact with users (Chang et al. 2025). These statements suggest growing industry attention to users’ sense of agency, though they are company self-descriptions rather than independent evidence of practice.

1.1 User Agency over Both Outcomes and Process

Agency has been studied across multiple disciplines, and conceptions of agency cluster along two primary dimensions: outcome control and process control. The first concerns the capacity for action that produces desired outcomes, sometimes highlighted in summaries of Bandura’s (2001) social-cognitive account of agency as acting to produce intended effects, which self-efficacy research operationalizes (Prunkl 2024). The second dimension concerns control over processes concerning what a person intends and what follows moment to moment. This sense of agency refers to whether a person experiences themselves as directing what is happening while it happens (Synofzik, Vosgerau, and Voss 2013; Tapal et al. 2017). Translated to AI development, these two dimensions roughly correspond to ensuring that (a) AI products improve users’ desired capabilities (outcome control), and (b) AI products give users meaningful control over their technology experiences (process control).

Technology developers have often refrained from being explicit about the outcome metrics that their products are designed to influence—in deference to the diversity of use cases that users may discover for general-use products. Companies have sought user input on what they prefer, find valuable, or find to be “worth their time” (Clegg 2021)—but those user inputs are often provided in relation to individual experiences with a specific feature or stimulus, rather than holistically, reflecting upon users’ broader goals for the product (Schulz van Endert and Mohr 2020; Siebold 2021). Generative AI developers have continued this approach, regularly soliciting user input on whether a given AI chatbot response is preferable or mood-enhancing (Johnson 2025).

However, aligning a product to immediate user inputs primarily refines user experience at the surface level, potentially at the expense of long-term empowerment. For example, in one study of 1.5 million Claude conversations, interactions with greater disempowerment potential received higher user approval ratings, such as emphatic validation of users’ grandiose or persecutory self-narratives (Sharma et al. 2026). Anchoring on user approval ratings of granular AI responses may fail to adequately capture users’ broader outcome goals, and thus may not shape the products to reflect true user agency, as defined by the longer-term outcomes users wish to bring about.

1.2 Designing for Agency

The current research aims to advance the practice of measuring and advancing agency in the context of technology development. In doing so, we hope to support centering user agency as a practical product choice for companies. Rather than adding controls in later stages in response to user complaints, designing for agency may help improve perceptions of AI, which is especially important at a time when many users worry about AI's societal effects. Technology companies often state this general goal. For example, Y Combinator, an accelerator that has launched many leading technology companies, encapsulates its philosophy as "Make something people want."¹ Yet, companies often do not consider the difference between long-term, aspirational desires, and short-term, potentially transient wants. This may lead companies to optimize for engagement, rather than aspiration.

Serving broader aspirations is technically more challenging than optimizing for short-term engagement, as feedback cycles are longer and harder to measure, so many companies have not publicly invested in such initiatives. However, some companies have indeed moved toward measuring user-centered preferences: Pinterest, for example, has developed multi-embedding retrieval frameworks that model both implicit behavioral signals and explicit user interests, such as followed topics, encoding users' interests across thousands of interactions (Fan et al. 2025).

1.3 The Current Research

The study of the human impact of AI is moving quickly, in an attempt to keep pace with the development of AI systems. Even as new research is generated, the observation that users value agency is likely to remain true, given that agency and autonomy are fundamental human needs (Vansteenkiste, Ryan, and Soenens 2020). AI's reported benefits continue to concentrate in instrumental domains, with more developed evidence on how AI provides users with outcome, rather than process, control. The present research examines user agency in generative AI use, conceptualized as both outcome and process control, as explicitly reported by users. The study draws on a probability-based national panel of U.S. adults who use generative AI ($n = 2,126$) to examine U.S. users' explicit aspirations and concerns for AI usage and perceived capabilities gained through AI. Within this panel, we examine the relationship between these two agency dimensions, or whether users who report greater AI control also report larger gains from AI usage. We therefore, in this study, examine five questions:

RQ1: In which life domains do users report that AI has changed their agency, compared to how able they report being in those same domains generally?

RQ2: What do users say they want from AI, and does it align with the domains in which they report benefit?

1. <https://www.ycombinator.com/>

RQ3: How much control do users report over their AI tool outputs, does it differ across products, and is it associated with larger agency gains?

RQ4: Where do users report concern, and to what extent are those concerns about agency?

RQ5: Do reported agency gains from AI differ across demographic groups?

To address these questions we fielded a survey combining closed-ended and open-ended items. The closed-ended measures of perceived agency, change attributable to AI, and control address RQ1 and RQ3. To capture the full range of users' intentions and concerns about AI, RQ2 and RQ4 adopt an inductive approach that allows these categories to emerge from users' own open-ended responses via thematic analysis. Our design suits RQ1, RQ2, and RQ4, which concern population distributions and are answerable with quantified uncertainty from a weighted probability sample. RQ3's comparative and relational components require two qualifications: Because users choose their own tool, product differences partly reflect who chooses what. Generally efficacious users may both gain more from AI and feel more in control of it, so we report the association between control and agency. RQ5 is exploratory and reported as hypothesis-generating. All measures are self-reported perceptions, whose limitations we address in the Discussion (Section 5).

2 Prior Work on AI and User Agency

Early research on AI systems suggests two emerging patterns. First, AI usage generally has been shown to produce benefits in outcome control for instrumental domains while producing mixed results for outcomes in social, emotional, and relational domains. Second, greater perceived control over AI products has been shown to predict user reports of greater overall outcome control.

2.1 Mixed Results in Instrumental and Non-Instrumental Domains

Early research demonstrates that using AI can enhance outcomes in instrumental tasks, defined by Korsgaard (1983) as tasks that are a means toward some larger goal. For example, in a work environment, using AI predicted subsequent self-reported work satisfaction, flourishing, and reduced frustration (Ngwenyama, Kanita, and Rowe 2025). Other research has demonstrated that AI chatbots can increase creativity as measured by idea generation (Lee and Chung 2024). Chatterji et al. (2025) found that the most common uses of ChatGPT at the time of their study included work, practical guidance, information, or writing, with only 4.3% of queries falling into self-expression or social categories. Thus, there is emerging early evidence indicating that people are likely to both use and report value related to the use of AI for help with instrumental outcomes.

In contrast to the above outcome benefits in instrumental domains, early research on AI's impact in social, emotional, and moral domains yields mixed results. In one study, AI built with the purpose of companionship has been shown to reduce loneliness as much as interacting with another person, and more than passive alternatives such as watching videos, with users underestimating the benefit (De Freitas et al. 2026). Others have found that, while AI can increase social communication by automating responses, people may judge non-authentic replies negatively (Hohenstein et al. 2023). Moreover, general chatbots vary in their ability to help users with mental health issues, as measured by correcting cognitive biases (Rządeczka et al. 2025). In other studies, large language models (LLMs) were noticeably worse at understanding questions with a moral dimension (Hilliard et al. 2025) and AI-driven virtual companionship reduced social anxiety online while increasing it offline (Xie and Wang 2024). And, in contrast to the idea that LLMs reduce loneliness, Fang et al. (2025) and Phang et al. (2025) reported that higher-intensity LLM usage is associated with more loneliness, higher emotional dependence on chatbots, more problematic usage, and reduced socialization, with some individuals/heavy users being more vulnerable to these effects. In sum, unlike task completion use cases, LLM benefits for social and emotional capabilities are mixed, and appear to depend on deliberate product design, making it important to establish where users themselves report agency gains and where gaps lie.

2.2 The Importance of User Control

Despite the mixed findings, prior work has shown that benefits from AI usage often depend on deliberate tool design. Careful design can improve results even in affective domains: conversational agents can produce measurable clinical gains (H. Li et al. 2023; Heinz et al. 2025), and structured, purpose-built applications show promise in early mental-health detection (Dehbozorgi et al. 2025), mindfulness training (Mitsea, Drigas, and Skianis 2023), and social-skills instruction (Adako, Adeusi, and Alaba 2025). However, these same capabilities can work against users when AI tools are built without intentional safeguards against over-reliance. In education, Bastani et al. (2025) found that AI assistance improved short-term performance while impairing longer-term learning, and Darvishi et al. (2024) showed that students become reliant on that assistance in ways self-regulation strategies only partly offset. Frequent use predicts over-reliance more broadly, particularly among younger users (Gerlich 2025; Xu et al. 2025).

Some research suggests that increasing perceived control over AI tools could be a way to improve users' experience of benefit. For example, a study of Taiwanese students found that a perceived sense of self-determination and control over use of ChatGPT was related to both usage frequency and satisfaction with the product experience (Tsai et al. 2026). Y. Li et al. (2025) showed how chatbot customization in a mental health study led to reports of increased emotional connection, trust, engagement, and satisfaction with a mental health chatbot. Others have found that enhancing perceived control through choice visibility for personalized recommendations led to increased recommendation

satisfaction and platform trust (Abinesh and Dulloo 2024; Xu and Chen 2025).

However, limited studies have investigated how perceived control leads to more tangible user benefits (Oz et al. 2024; Yang et al. 2024). Based on these findings, we suggest that active, managed AI usage will be experienced as more beneficial than passive over-reliance on AI. This is consistent with student and academic sentiment that AI can enhance learning but also that its benefits depend on ethical, active use and avoiding over-reliance (Qian 2025; Vieriu and Petrea 2025). Relatedly, users can feel threatened when information systems act more autonomously, further underscoring the need for user control in AI product design (Adam et al. 2024). Across these early studies, it appears that users understand AI usage introduces both benefit and risk, and desire to actively manage their usage to mitigate the risk.

2.3 Measuring AI Experience Variation across Demographics

Historically, adoption and skilled use of new technologies have varied across demographic groups differentiated by factors such as gender, age, and income. In the case of AI, Cranney, Delecourt, and Koning (2026) estimated that, across more than 100 countries and 318,924 survey respondents, adoption sat at 47.8% among men and 39.3% among women, a relative gap of 22% that has narrowed but not closed. This gender disparity persists in a national U.S. sample, where 31% of men but 18% of women were expected to personally benefit from AI, with men also being twice as likely as women to report being more excited than concerned regarding AI (McClain et al. 2025). Similarly, varying racial groups experience usage differences, with 70% of Asian adults reporting chatbot use against 46–49% of White, Black, and Hispanic adults (Gottfried et al. 2026); a nationally representative longitudinal panel similarly found Asian respondents were 11.1 percentage points more likely than White respondents to have used LLMs, while these gaps by race, gender, age, income, and education persisted or widened, rather than narrowed, across survey waves (Angrisani et al. 2026).

Surprisingly, younger adults under 30 are considerably more likely than adults 65 and older to expect AI to worsen human abilities, including thinking creatively and forming meaningful relationships (Kennedy et al. 2025). AI user base compositions are constantly shifting, with the early gender gap in ChatGPT use closing over roughly two years, and the balance of work, practical-guidance, and personal use varying substantially across users (Chatterji et al. 2025); however, this is not true of all demographic gaps, as one longitudinal U.S. panel found none of the usage gaps it tracked narrowed between 2023 and 2024, with age- and education-based gaps in particular widening over time (Angrisani et al. 2026). Together these findings indicate that research on AI and AI agency requires demographic stratification.

3 Methods

3.1 Study Sample

We designed and fielded a survey as a part of the Understanding America Study (UAS), a probability-based internet panel administered by the Center for Economic and Social Research (CESR) at the University of Southern California. The UAS recruits participants via Address-Based Sampling (ABS), in which household addresses are drawn from postal records, providing each U.S. household a non-zero probability of selection. To minimize coverage bias, the UAS provides internet-connected tablets to households lacking existing internet access, ensuring representation of offline populations. At the time of analysis, the panel comprised approximately 15,000 U.S. residents. We did not control panel recruitment, retention, or weighting; those are panel properties and documented by administrators. Panel members are compensated at a standard rate of \$20 per 30 minutes of survey time, prorated for shorter surveys. The UAS and its surveys operate under University of Southern California IRB approval #UP-14-00148.

Our study (UAS 738) was in the field from August 4 to September 4, 2025, and was administered in English only. Invitations were restricted to active panel members who had indicated on previous UAS surveys in this series (UAS 607 and UAS 689, which span April 2024–May 2025) that they had used AI applications. Of roughly 14,000 active panel members at the time, 2,706 were eligible and invited. The released analytic file contains 2,126 respondents. Survey weights were constructed in two stages: first, base weights correcting for unequal probabilities of household selection; and second, post-stratification weights calibrating the sample to U.S. Census benchmarks on gender, race/ethnicity, age, education, and geographic region to correct for unequal probabilities of household response to any given survey.

These post-stratification weights differ amongst each UAS study to capture the reference population (Understanding America Study, [n.d.](#); see also Alattar, Messel, and Rogofsky 2018). Therefore, Census calibration depends on recruitment and response rates in the panel. We also correct for systematic nonresponse within our eligible AI-using sample, but do not adjust for AI-use itself in screening for eligibility. Our weighted estimates should thus be interpreted as describing U.S. adults who use generative AI and not all U.S. adults.

When comparing the weighted study sample with the weighted eligible AI-using sample, we find they closely resemble each other (for instance 36.2% versus 35.7% of respondents are aged 18 to 34, and 45.4% versus 47.1% are women; Table A6). Both, however, differ from the full weighted panel, which matches marginals from the Census. Within our panel, AI users are generally younger (36.2% in AI users versus 23.2% in the full panel aged 18 to 34), more educated (62.2% versus 36.5% holding a bachelor's degree or higher), and higher-income (51.4% versus 30.0% at or above \$100,000). These differences reflect the eligibility criterion rather than panel composition.

3.2 Survey Measures

In previous qualitative interviews conducted across multiple countries to understand what capabilities users desired from AI (Jigsaw 2025), we identified five dimensions of agency that users associated with AI use: cognitive (making decisions confidently), affective (managing feelings), structural (influencing one's community), relational (improving relationships), and instrumental (learning new things, completing tasks). These dimensions informed development of survey items assessing both current perceived capabilities of users and current perceived effects of AI usage on outcomes.

Baseline perceived outcome control was measured with 13 items asking how able respondents currently felt to bring about desired outcomes in each of the five agency domains, rated from 1 ("not at all able") to 5 ("very able"). Changes in perceived outcome control attributable to AI use were measured with six parallel items asking whether, as a result of using AI, respondents felt more or less able in each domain (e.g., "As a result of using AI in your life, do you feel more or less able to reach your goals?"), rated from 1 ("much less capable") to 5 ("much more capable") with a midpoint of 3 ("neither more nor less capable"). Process control, the second dimension, was measured separately at the product level, as described below.

Items were assigned to the five agency dimensions a priori, with the cognitive domain represented by five baseline items and the instrumental, affective, social, and structural domains by two each. The six AI-attribution items comprise one per domain plus one general item on reaching goals.

We asked users "Which of the following genAI tools have you used?" within a four-week period, presenting 26 named tools identified by previous surveys as most common: Gemini, ChatGPT, Claude, Microsoft Copilot, Perplexity, Jasper, DeepSeek, Meta AI/Llama, Character.AI, Replika, DALL-E, Grammarly, ElevenLabs, Poe, DeepAI, Grok, Prezi, Image Creator from Microsoft Designer, Gretel.ai, Midjourney, Murf.AI, QuillBot, Remini Photo Enhancer, SnapAI, DreamStudio, and Stable Diffusion, along with "other" and "none of these" options. In our sample, ChatGPT was the most widely used (75.4%), followed by Gemini (28.5%) and Microsoft Copilot (26.1%); 9.5% weighted reported using none of the listed tools in that window. For each tool they indicated using, we asked users how often they used them (daily, several times per week, a few times or less, or not at all). We then asked users questions about their experiences using their most-used AI product, including perceived process control and perceived outcome control. Respondents who reported using more than one tool were asked which they used most often, and that response was piped into the subsequent block, so all product-level estimates refer to each respondent's self-identified primary tool.

Respondents for whom a most-used AI tool could be identified (1,903 of 2,126, 89.5%) were asked five questions about that specific product: whether they felt in control over the results it created ("Do you feel like you are in control over the results that are created?";

1=“not at all in control” to 4=“very in control”), whether it made them more or less able to reach their goals (1=“much less” to 5=“much more able”), whether they felt they were building a personal bond with it (1=“not at all” to 4=“very much”), whether they felt more or less lonely than in everyday life (1=“much less alone” to 5=“much more alone”), and whether they ever felt manipulated by it (1=“not at all” to 4=“very much”). The first two items constitute our measures of perceived process control and perceived outcome control. A full list of question text is in Appendix C.

3.3 Analytic Strategy

All estimates are weighted with Taylor-series linearized design-based standard errors, with the design treated as weighted simple random sampling. Thirty-six respondents held a weight of zero, so they are excluded. The design effect is 1.85, yielding an effective sample size of approximately 1,153 (Henry and Valliant 2015). We report 95% confidence intervals for all point estimates and per-item n throughout. Distributions are summarized with weighted proportions and means. The AI-attribution items are each tested against the scale midpoint of 3.

To relate perceived process control to perceived outcome control, we regressed each capability item on process control entered as a linear 1–4 term, using weighted least squares with HC1 robust standard errors, first unadjusted and then adjusting for co-variables age, gender, education, and household income. As a robustness check, each model was refitted as a weighted ordinal logistic regression, and conclusions did not differ. Associations among agency items are weighted Spearman correlations.

We employed a Kruskal–Wallis omnibus test to analyze product comparisons across the three most-used tools, followed by weighted ordinal logistic contrasts against the modal product (ChatGPT). We also tested each capability item against gender, age, and education as weighted Pearson correlations (point-biserial for the binary gender indicator) with design-based standard errors, and applied Holm–Bonferroni correction across the resulting family as exploratory demographic analysis. We coded age continuously in years, education continuously on the UAS 16-point scale, gender using binary indicators (1 = male, 0 = female). Because the data file for this wave provided only this binary measure, we did not analyze responses beyond men and women separately. We tested race as a categorical variable. For each capability item, we ran a Kruskal–Wallis omnibus test across the four groups large enough to estimate (White $n = 1,414$; Asian $n = 297$; Black $n = 177$; Mixed $n = 131$) with contrasts against White respondents as the reference group. Other groups such as American Indian or Alaska Native ($n = 23$) and Hawaiian or Pacific Islander ($n = 10$) had too few respondents to analyze separately. Holm–Bonferroni correction was applied within each of the three families: the 18 correlations, the six omnibus tests, and the 18 contrasts. Nine respondents with missing race data were excluded from the race analyses. Subgroup differences in theme prevalence were tested using two-proportion z -tests by comparing each demographic group against all other respondents, with Holm–Bonferroni correction within each open-ended question.

Uncorrected results are reported in Appendix E. Analyses were conducted in Python using statsmodels and scipy.

3.4 LLM-Assisted Thematic Analysis with Human-in-the-Loop Validation

We asked four open-ended questions. The excitement and concern questions were asked of all respondents (1,919 and 1,924 responses analyzed, respectively, after excluding responses under 10 characters); the goal-achievement question yielded 387 analyzable responses, and the control question, asked about the respondent's most-used tool, yielded 502:

- Excitement about AI: In a sentence or two, please tell us what you are most excited about regarding AI, for yourself or society at large (Results presented in Figure 3; Table A5 and Figure A1).
- When AI supports users in goal achievement: In a sentence or two, please describe a time when AI made you feel more able to reach your goals (Results presented in Figure 4; Table A5).
- Concerns about AI: In a sentence or two, please tell us your biggest concern regarding AI, for yourself or society at large (Results presented in Figure 8; Table A5 and Figure A2).
- When users feel in control of AI use: In a sentence or two, please describe a time when you used [most used AI], where you felt very much in control (Results presented in Figure 7; Table A5).

To analyze these open-ended responses, we developed a multi-stage computational pipeline using an LLM as a coding assistant under iterative human input. The approach integrates two core steps: inductive thematic identification and deductive text classification. The pipeline combines repeated independent induction, consensus merging, and structured human review to produce a stable and interpretable codebook. This design follows emerging practice for LLM-assisted annotation and content analysis, which emphasizes human oversight of codebook development, repeated classification runs to assess stability, and validation against human coders (Chew et al. 2023; Gilardi, Alizadeh, and Kubli 2023; Ziems et al. 2024). The procedure was applied separately to open-ended questions, generating both a shared codebook across all respondents and demographic-specific codebooks capturing subgroup-specific themes (details in Appendix A). We applied gpt-5.4-mini (temperature=0) as the LLM assistant in the study.

Codebook construction, classification, and validation proceeded in six phases, specified fully in Appendix A. In brief, the LLM generated candidate themes in three independent runs that were merged into a consensus codebook, reviewed and finalized by researchers (Checkpoint 1), and pruned of labels assigned to fewer than 3% of responses (Checkpoint 2). The finalized codebook was then applied to every response five times, with a label retained only if assigned in at least three of five runs; per-response agreement

and label entropy were computed as stability metrics. Validity was assessed by comparing LLM labels against those of a researcher who blind-coded a stratified random sample of responses (precision, recall, and F1). Finally, a gap-screening procedure identified demographic-specific themes not captured by the shared codebook, which were pruned, researcher-reviewed (Checkpoint 3), and classified with the same majority-vote procedure.

The theme labels and their prevalences are reported in Appendix B; the full codebooks, including theme definitions and illustrative example responses, are available from the authors upon request. The performance of the LLM-assisted thematic analysis pipeline is reported in Appendix B.

Across all four survey questions, the LLM-based thematic classification demonstrated satisfactory stability, with per-theme Mean Agreement ranging from 0.79 to 0.96, indicating that the majority-vote label assignment was highly consistent across repeated classification runs. Inter-rater validation, in which an independent human researcher blind-coded a stratified sample of responses, yielded macro F1 scores between 0.74 and 0.84 across the four questions, reflecting acceptable to good agreement between the two classification systems (see details in Appendix B).

4 Results

4.1 Baseline Agency and Reported Effects of AI on Outcome Control

AI users generally reported moderate to high baseline agency across dimensions, although with substantial heterogeneity. Respondents expressed high levels of self-reported capability in the cognitive and instrumental domains (Figure 1): 84% responded either very able or often able to “make up my own mind,” 83% for “trust my own judgement,” 83% for “learn anything that is useful or important,” and 82% for “complete the tasks I set out to achieve.” Respondents reported baseline agency as lower in structural, social, and affective domains. For instance, only 43% responded very able or often able to “get people in power to listen,” 54% to “make new friends,” 54% to “get help from formal support systems...when I have a problem,” and 69% to “get over negative emotions.” The pattern overall indicates that U.S. AI users feel more agentic in controlling personal cognitive, informational, and task completion outcomes compared to those contingent on social or structural settings.

Users’ reported effects of AI on their perceived outcome control differed substantially depending on the domain (Figure 2). It was rare for users to report that AI decreased their capability in any domain. A majority of users reported the highest positive value of AI in the instrumental domain, with 66% responding that AI renders them a little more/much more capable of learning anything useful or important. A majority (53%) also responded that AI makes them more capable of reaching their goals. In contrast, users

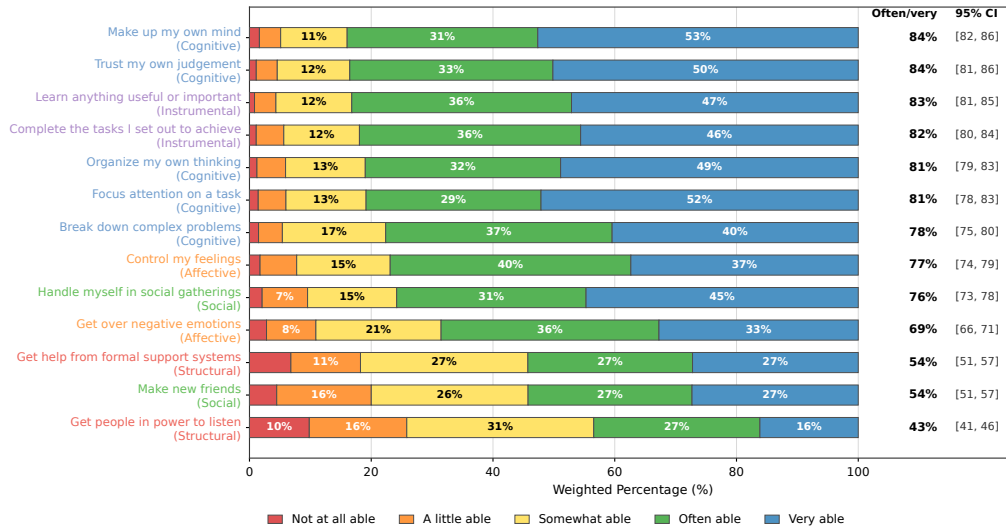


Figure 1: Baseline perceived agency across domains among U.S. AI users. *Note.* Data are from UAS 738, weighted to be nationally representative of English-speaking U.S. adults who use generative AI ($n = 2,062$ – $2,068$ per item). Respondents rated how able they are to perform each capability from 1 (not at all able) to 5 (very able). Bars show the weighted response distribution; the right column gives the percentage responding often able or very able with its 95% confidence interval. Items are colored by agency domain and ordered by positive response rate.

reported relatively low current benefit from AI for capabilities in the social and affective domains, with only 7% responding AI increases capabilities to make new friends, 14% for controlling feelings, and 21% for influencing powerful people.

We tested each item against the scale midpoint. Five of six differ from no net change at $p < 0.001$. The exception is AI impact on making new friends ($M = 3.02$, $SE = 0.015$, $p = 0.30$), meaning there was no significant effect in either direction.

4.2 AI-Related Aspirations and Enabled Goal Achievements

Users' open-ended descriptions of what excites them about AI were related to task-oriented and information-seeking goals (Figure 3). LLM-assisted thematic content analysis identified 16 shared themes across respondents. The most frequently endorsed theme was information access (20.7%, $n = 398$), followed by saving time (17.4%, $n = 333$), task automation (12.5%, $n = 239$), and reducing cognitive load (11.5%, $n = 220$). Other commonly identified themes included uncertainty and concerns (10.5%), learning support (10.4%), reading and writing support (8.4%), and future transformation (8.1%). Notably, themes about social companionship or emotional connection with AI did not appear among the common themes across the full sample. Excitement varied across demographic groups (see Table A5). Older adults (65+) show a higher proportion of excitement related to information access (29.9% vs. 19.3% among all other respondents, $p_{\text{Holm}} = 0.0003$). In contrast, young men (aged 18–35) exhibit elevated rates of excitement for both task automation (17.6% vs. 11.9%, $p_{\text{Holm}} = 0.024$) and future

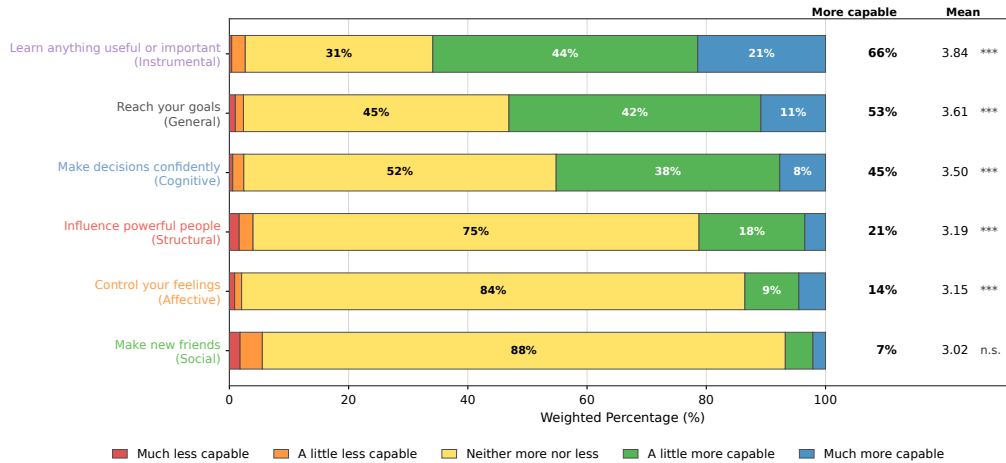


Figure 2: Reported effects of AI on perceived capability across domains. *Note.* Data are from UAS 738, weighted ($n = 2,059$ – $2,061$ per item). Respondents rated whether AI has made them more or less able in each domain from 1 (much less capable) to 5 (much more capable). Bars show the weighted response distribution; the right columns give the percentage responding a little or much more capable and the weighted mean. Each item's weighted mean is tested against the scale midpoint of 3 with Holm–Bonferroni correction; reported p -values are corrected. *** $p < 0.001$, n.s.=not significant.

transformation with AI's potential to drive large-scale societal and industrial change (15.0% vs. 7.3%, $p_{\text{Holm}} < 0.001$).

Thematic analysis also identified demographic-specific themes that were less prevalent in the full sample but salient within particular groups (Figure A1). Among older adults, three additional themes emerged: cross-domain pattern finding ($n = 38$), reflecting AI's ability to connect ideas across domains; conversational understanding ($n = 18$), referring to natural, dialogue-like interaction that users experienced; and deepening ideas ($n = 14$), capturing AI's support for more thorough thinking. Young women uniquely emphasized human-AI collaboration ($n = 95$), indicating interest in working with AI as a partner, alongside reduced social friction ($n = 31$), and perspective-taking ($n = 8$), reflecting interest in exposure to multiple viewpoints with AI. Young men were characterized by excitement around AI's impact on human role reallocation ($n = 70$), reflecting on AI restructuring human labor, and technical problem solving ($n = 70$).

Users' open-ended descriptions of occasions when AI enhanced their ability to achieve goals were concentrated on language and information-related tasks (Figure 4). The most frequently reported use was writing, editing, and polishing text (e.g., emails, reports, summaries; 26.6%, $n = 103$), followed by finding and summarizing information (20.9%, $n = 81$), task and workflow management (15.0%, $n = 58$), and organizing ideas and planning (14.2%, $n = 55$). Other themes, including technical and coding assistance, academic support, health and wellness guidance, brainstorming, and creative production, each accounted for approximately 5–6% of responses.

The contexts in which AI enhanced goal attainment varied across demographic groups

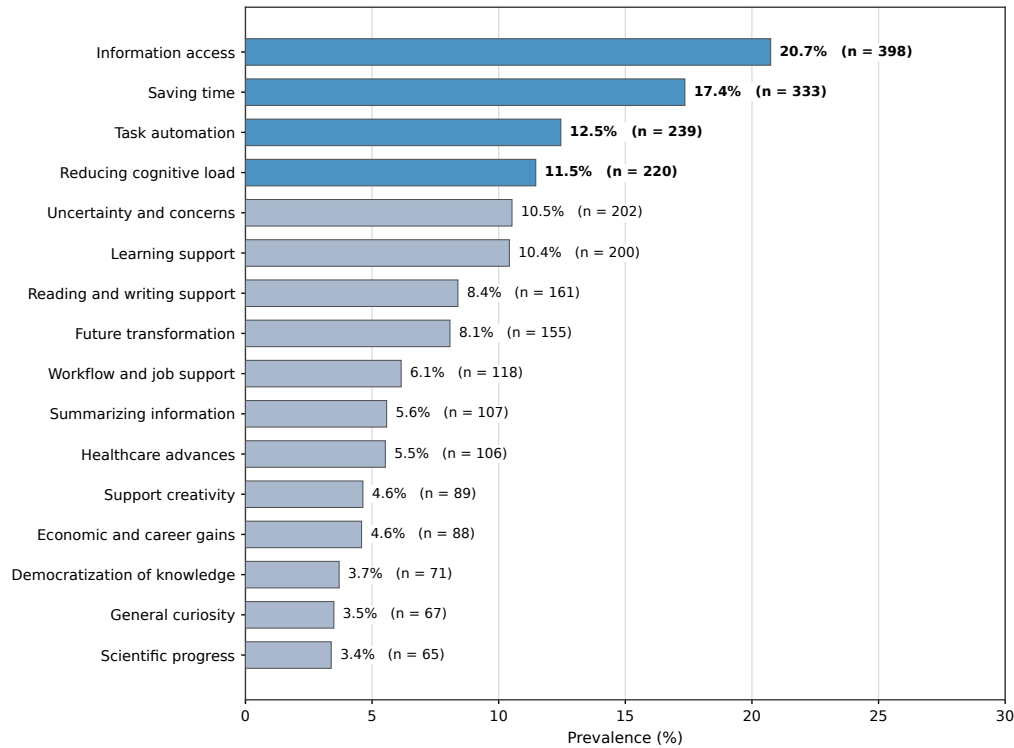


Figure 3: Prevalence of shared themes in responses about AI excitement. *Note.* Open-ended responses to: “In a sentence or two, please tell us what you are most excited about regarding AI, for yourself or society at large.” Themes were derived by the LLM-assisted pipeline described in Methods (Section 3), with human review at each checkpoint; they are not mutually exclusive and thus do not sum to 100. Bars show the percentage of analyzed responses assigned each theme, with counts printed on the bars; highlighted bars mark themes discussed in the text. Demographic breakdowns appear in Table A5 and Figure A1.

(Table A5). Two of eight subgroup contrasts survive correction, both among young women: adjusting communication tone (15.0% vs. 3.1% among all other respondents, $p_{\text{Holm}} < 0.001$) and task and workflow management (26.7% vs. 12.8%, $p_{\text{Holm}} = 0.041$). The remaining differences do not survive correction and are reported descriptively. Older adults (65+) were more likely to report finding and summarizing information as their primary use (27.3% vs. 19.9%), while task and workflow management was less common (9.1% vs. 16.0%). Young women also showed higher rates of organizing ideas and planning (23.3% vs. 12.5%). In contrast, young men reported higher engagement in finding and summarizing information (28.3% vs. 19.9%), solving practical problems (17.4% vs. 8.5%), and learning and skill development (13.0% vs. 4.4%). Fewer demographic-specific themes emerged for goal achievement, partly reflecting the lower response rate to this question: persuasive writing support ($n = 7$) and option comparison support ($n = 6$) among older adults, financial management ($n = 7$) among young men, and none unique to young women.

Across quantitative and qualitative results, aspiration and experience aligned closely. The domains users reported being most excited about—information access, time saving,

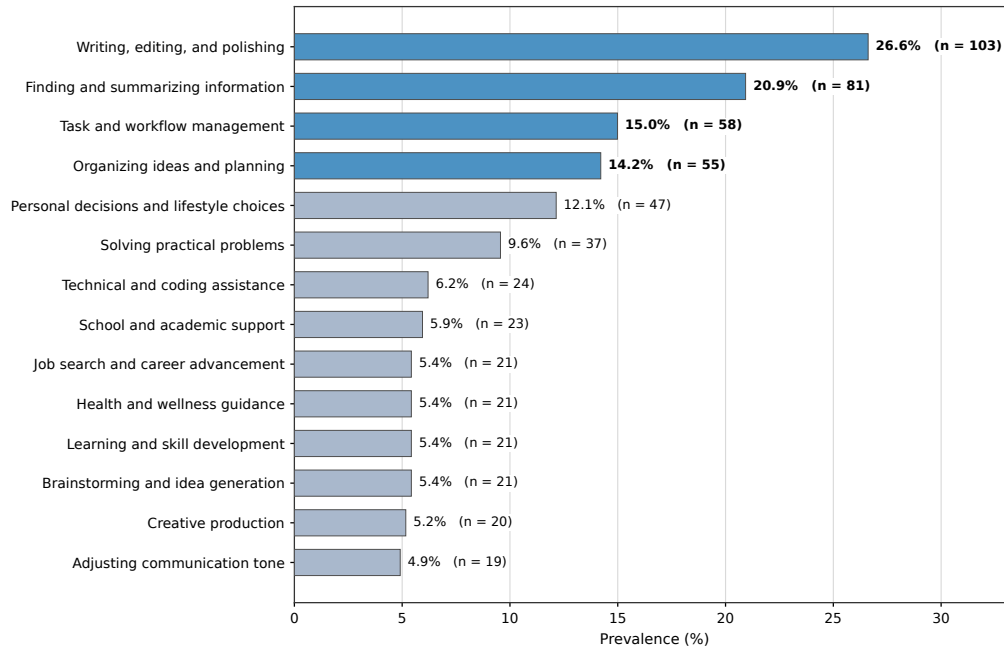


Figure 4: Prevalence of shared themes in descriptions of AI-supported goal achievement. *Note.* Open-ended responses to: “In a sentence or two, please describe a time when AI made you feel more able to reach your goals.” Coding as in Figure 3. Demographic breakdowns appear in Table A5.

task automation, cognitive offloading—were the domains in which they report the largest capability gains, and the near-absence of companionship themes parallels the null result for making new friends.

4.3 Perceived User Control (Outcome and Process) and Association with Reported Gains

Three products were used broadly enough to be analyzed independently: ChatGPT ($n = 1,089$), Gemini ($n = 223$), and Copilot ($n = 212$), each having more than 200 users who indicated it was their most-used AI tool. Perceived control was moderate and similar across products (Figure 5): on the 1–4 scale, ChatGPT $M = 2.73$ ($SD = 0.90$), Gemini $M = 2.64$ ($SD = 1.00$), and Copilot $M = 2.74$ ($SD = 0.97$), with “somewhat in control” the modal response for each.

Product differences in perceived control did not reach statistical significance, $H(2) = 4.44$, $p = 0.109$, nor did differences in perceived outcome control, $H(2) = 4.38$, $p = 0.112$, or in feeling manipulated, $H(2) = 2.25$, $p = 0.325$. Only reported personal bond differed significantly by product, $H(2) = 21.92$, $p < 0.001$, with ChatGPT users reporting more bonding than Gemini or Copilot users. Relative to ChatGPT users, both Gemini users ($OR = 0.47$, $p < 0.001$, $p_{\text{Holm}} = 0.001$) and Copilot users ($OR = 0.57$, $p = 0.003$, $p_{\text{Holm}} = 0.023$) had lower odds of reporting a bond, and both contrasts survive Holm correction ($n = 1,524$). Loneliness effects were marginally significant,

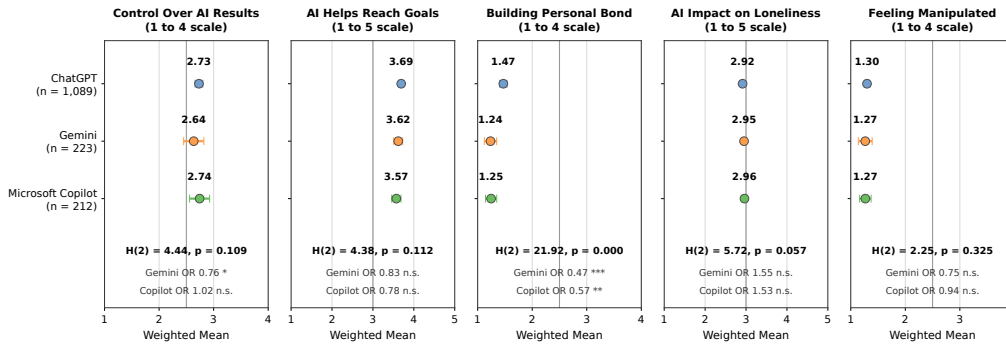


Figure 5: Product experience among users of the three most-used AI tools. *Note.* Data are from UAS 738, weighted to be nationally representative of U.S. adults who use generative AI. Product groups are defined by the product each respondent reported using most, which was piped into these items at fielding. Respondents rated their most-used product (ChatGPT $n = 1,089$; Gemini $n = 223$; Copilot $n = 212$) on perceived control over results, effect on reaching goals, personal bond, loneliness, and feeling manipulated. Error bars are linearized design-based 95% confidence intervals; the n shown is for the control item. Kruskal–Wallis omnibus tests across products are reported in the text; only personal bonds differ reliably. Because respondents select their own primary tool, contrasts are descriptive.

$H(2) = 5.72, p = 0.057$.

Perceived process control was positively associated with perceived gains in outcome control across all six domains (Figure 6). Regressing each capability item on process control (entered 1–4), all six slopes are positive and survive Holm correction: influencing powerful people ($b = 0.102$, 95% CI [0.062, 0.142]), reaching your goals ($b = 0.102$ [0.053, 0.151]), controlling your feelings ($b = 0.083$ [0.043, 0.122]), making decisions confidently ($b = 0.072$ [0.028, 0.116]), learning anything useful or important ($b = 0.069$ [0.017, 0.120]), and making new friends ($b = 0.060$ [0.028, 0.092]).

All six associations survive adjustment for age, gender, education, and household income, with coefficients changing by no more than 0.006. The associations are small with a three-point move across the control scale corresponding to roughly 0.2 scale points of perceived capability. The strongest associations are with reaching goals and influencing powerful people, and the weakest is with making new friends.

Given the importance of reported process control in both our study and in previous research, we examined our open-ended responses to see when users reported feeling the most in control. As shown in Figure 7, open-ended responses indicate that users feel most in control when they can iteratively guide AI through prompting (34.1%, $n = 171$) and evaluate or edit its outputs (31.5%, $n = 158$). Other cases in which users perceived control of AI include AI supporting writing or polishing text (19.1%, $n = 96$) and finding information (14.5%, $n = 73$). Of four subgroup contrasts, the only relationship that survives correction is young men reporting greater control when using AI for specific, task-oriented purposes (25.0% vs. 12.2% among all other respondents, $p_{\text{Holm}} = 0.043$). Descriptively, older adults reported greater feelings of control when using

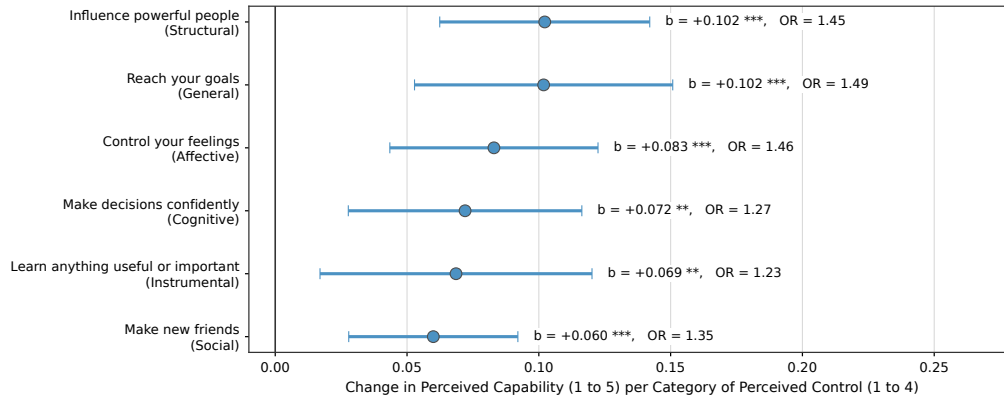


Figure 6: Perceived process control is associated with perceived capability gains in all six domains. *Note.* Weighted least-squares slopes (with HC1 robust 95% confidence intervals) from regressing each capability item on perceived control over AI results, entered 1 (not at all in control) to 4 (very much in control). All six slopes are positive and survive Holm–Bonferroni correction, and are unchanged by adjustment for age, gender, education, and household income (coefficients shift ≤ 0.006).

AI for information retrieval (20.0% vs. 13.8%) and when relying on their own domain knowledge to evaluate AI outputs (13.3% vs. 7.7%), enabling them to judge accuracy; young women more frequently reported feeling in control when effectively guiding AI through prompting (43.9% vs. 32.6%).

When examining users’ experiences based on the specific AI tool they most used, findings were consistent with the general pattern: users reported that AI products had very little effect on the social and affective domains and were most useful for helping them achieve instrumental goals. Specifically for ChatGPT users, while 83.0% reported no effect on loneliness and 70.6% reported no “personal bond” with the product, the majority (58.2%) said ChatGPT made them a little/much more capable in helping “reach your goals.” As detailed later in our analysis of descriptions of these experiences, described goals were generally in the instrumental domain. Results were similar for Gemini and Copilot users.

4.4 Examining Users’ Reported Concerns about AI

Users’ open-ended descriptions of their primary concerns about AI centered on over-reliance, malicious use, and job loss due to automation (Figure 8). Descriptively, the most frequently reported concern was reduced thinking and agency (24.3%, $n = 467$), reflecting fears that AI may undermine independent reasoning and decision-making. This was followed by manipulation and malicious use (19.6%, $n = 378$), job loss and automation (16.2%, $n = 311$), inaccurate or fabricated outputs (15.5%, $n = 298$), and loss of human control (13.1%, $n = 252$). Young men reported markedly higher concern about job loss due to AI (28.7% vs. 14.8% among all other respondents, $p_{\text{Holm}} < 0.001$), and young women expressed elevated concern about the environmental costs of AI (13.7% vs. 4.0%, $p_{\text{Holm}} < 0.001$). The following results did not survive correction and are

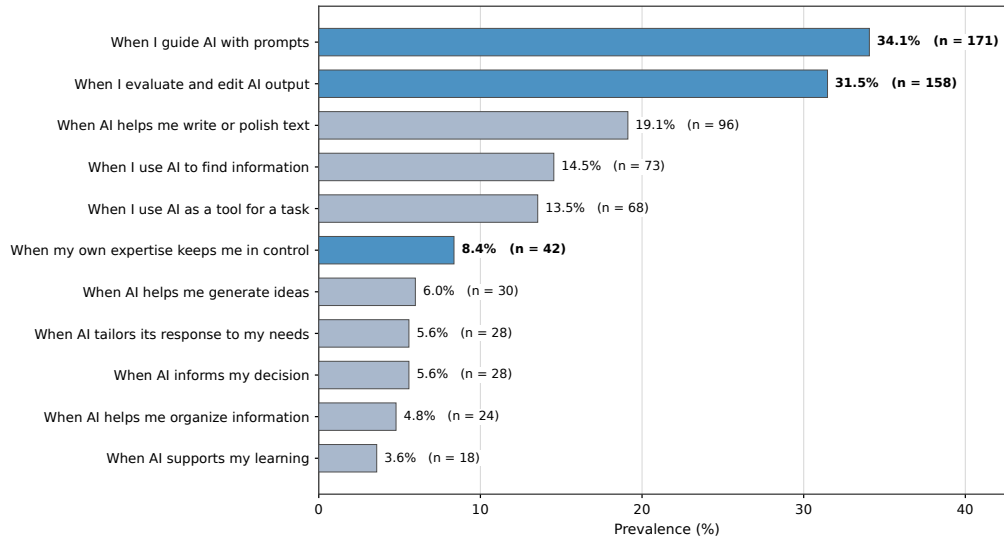


Figure 7: Prevalence of shared themes in descriptions of feeling in control of AI. *Note.* Open-ended responses to: “In a sentence or two, please describe a time when you used [most-used AI], where you felt very much in control.” Coding as in Figure 3. Demographic breakdowns appear in Table A5.

reported descriptively. Concerns about reduced thinking and agency were nominally more prevalent among young men (30.3%) and young women (27.8%) than among older adults (19.5%), and older adults expressed nominally greater concern about manipulation and malicious use (24.3% vs. 18.9%) and inaccurate outputs (19.1% vs. 14.9%).

Beyond the shared concerns, we identified several demographic-specific concerns for AI (Figure A2). Older adults expressed concerns about belief reinforcement ($n = 42$), concerns about treating AI as a tool rather than a person ($n = 9$), and lack of AI content disclosure ($n = 7$). Young women exhibited the most diverse set of unique concerns, including decline in work ethic ($n = 86$), inadequate social protection for affected individuals ($n = 70$), premature deployment ($n = 47$), loss of humanity ($n = 47$), and emotional dependence and mental health harm ($n = 36$), highlighting concerns about psychological and societal impacts. Young men also reported concerns about erosion of human creativity ($n = 53$) and polarization and echo chambers ($n = 12$).

4.5 Who Reports Capability Gains from AI

Three associations between the six capability items and gender, age, and education survive Holm–Bonferroni correction (Figure 9). Men reported greater benefit than women for learning anything useful or important from AI use ($r = 0.125$, $p_{\text{Holm}} < 0.001$) and for making decisions confidently ($r = 0.097$, $p_{\text{Holm}} = 0.020$). Education was also positively associated with reported gains in confident decision-making ($r = 0.098$, $p_{\text{Holm}} = 0.018$). Reported capability gains from AI also differed by race on four of the six items (all $p_{\text{Holm}} < 0.001$): making decisions confidently, reaching goals, controlling feelings, and learning anything useful or important. Race was also contrasted against

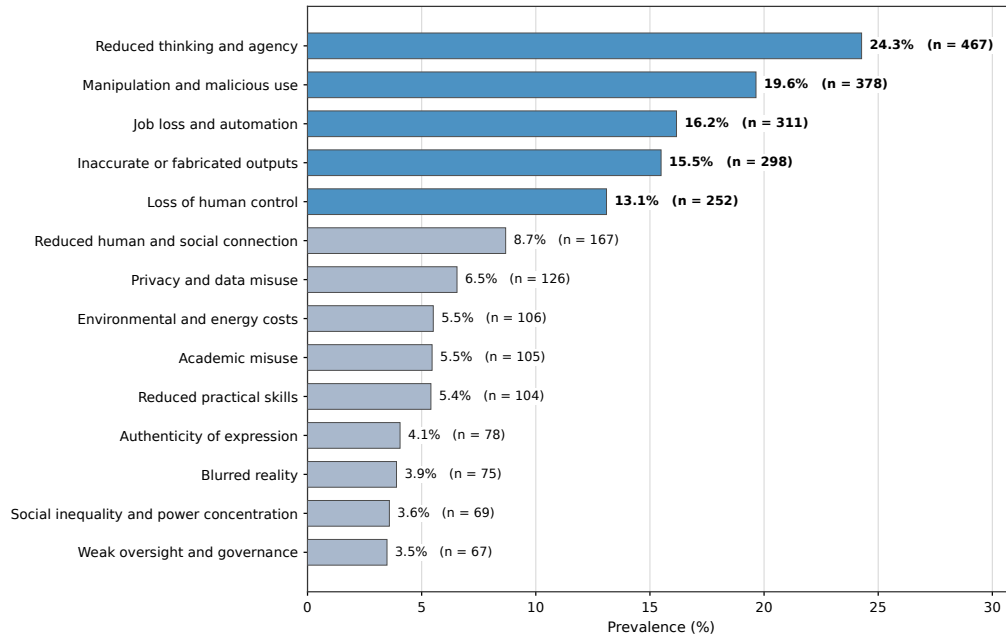


Figure 8: Prevalence of shared themes in responses about AI concerns. *Note.* Open-ended responses to: “In a sentence or two, please tell us your biggest concern regarding AI, for yourself or society at large.” Coding as in Figure 3. Demographic breakdowns appear in Table A5 and Figure A2.

White respondents, with two surviving correction, both for Asian respondents: making decisions confidently (+0.20 scale points, $p_{\text{Holm}} = 0.019$) and learning anything useful or important (+0.22, $p_{\text{Holm}} = 0.019$). Asian respondents reported the highest learning gains of any group even after adjusting for other covariates ($b = 0.20$, $p = 0.002$).

5 Discussion

Across a nationally representative sample of U.S. adults who use generative AI, we examined what users want from AI products and what makes their interactions with AI feel empowering. The findings are consistent with themes that are beginning to emerge in early generative AI research: (a) that users’ aspirational goals for AI center on instrumental outcomes, (b) that process control matters to users, and (c) that increased control over the user experience (process control) is linked to tangible perceived AI-driven gains in perceived outcome control. Results vary by demographic population.

5.1 Users’ Outcome Goals and Reported Benefits Are Predominantly Instrumental

The domains in which users most want AI to expand their capabilities to produce outcomes are generally consistent: information access, learning, task completion, and goal achievement. Likewise, users’ open-ended descriptions of what excited them

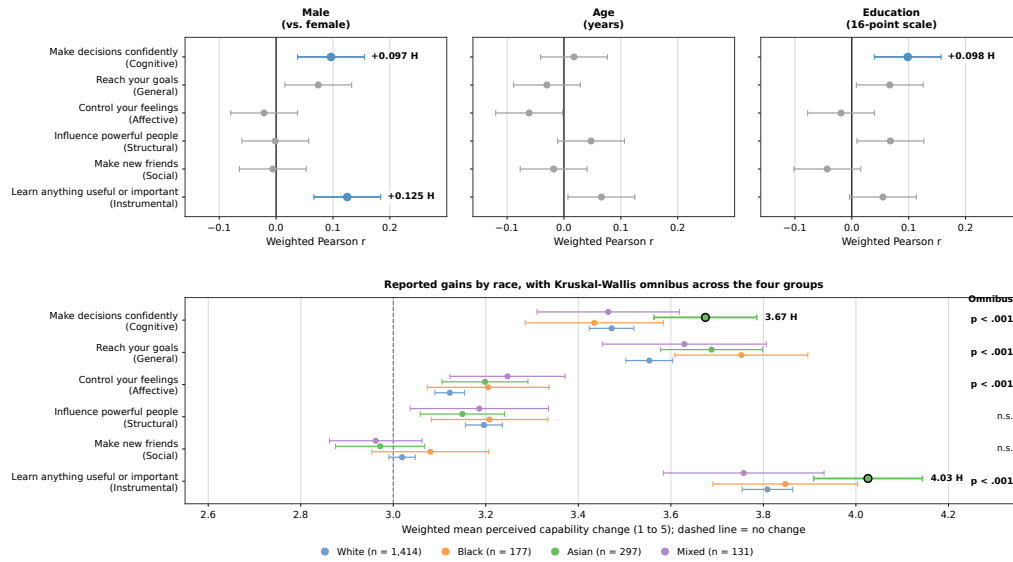


Figure 9: Demographic differences in reported capability gains from AI. *Note.* Top row shows weighted Pearson correlations between each capability item and gender, age, and education, with design-based 95% confidence intervals; Holm–Bonferroni correction was applied across the 18 correlations. Bottom row shows weighted mean perceived capability change by race with 95% confidence intervals, and the Kruskal–Wallis omnibus p (Holm-corrected across the six items) in the right-hand column. Contrasts are against White respondents, Holm-corrected. These analyses are exploratory and hypothesis-generating. Uncorrected results appear in Appendix E.

most about AI were dominated by themes of information access, saving time, and task automation; subsequently, they reported AI rendered them more capable of learning anything or reaching goals that were largely instrumental (see Figures 2–4). This focus on instrumental agency, defined in past work as tasks that serve as a means toward some larger end (Korsgaard 1983), was consistent across demographic subgroups and aligned with research showing that AI yields its clearest and most replicable benefits in work productivity, information gathering, creativity, and problem-solving contexts (Ngwenyama, Kanita, and Rowe 2025; Lee and Chung 2024; Chatterji et al. 2025).

This does not mean, however, that AI has no role to play in cognitive, social, affective, or structural domains. Some users did express excitement about AI’s potential for emotional support. The literature, too, shows that AI can deliver meaningful benefits in non-instrumental areas, but generally only when the application is purposefully designed. Intentionally built AI interventions have improved clinical outcomes (Heinz et al. 2025), supported mindfulness practice (Mitsea, Drigas, and Skianis 2023), developed social skills in individuals with autism (Adako, Adeusi, and Alaba 2025), and facilitated learning at individualized paces (Xu et al. 2025). In contrast, when general-purpose AI products are used in relational and affective domains without intentional design scaffolding, harms can emerge: higher-intensity LLM usage has been associated with loneliness, emotional dependence, and reduced socialization (Fang et al. 2025; Phang et al. 2025). Our social dimension results agree that usage of a general-purpose AI yields no statistically

significant loss or gain on friend-making for the typical user. The contrasting findings suggest that opportunities outside the instrumental domain are real, but they require more conscious and deliberate design investment than simply building capable models and deploying them broadly. Most users do not report benefits in these domains and several high-profile cases have suggested the potential for harm when general-purpose models are used in these ways.

Whatever benefits AI confers both in our sample and beyond, they do not appear evenly distributed. Existing literature documents that adoption is stratified by gender (Cranney, Delecourt, and Koning 2026), by race and ethnicity (Gottfried et al. 2026), and by expectation of benefit (McClain et al. 2025). Our demographic analyses were exploratory and hypothesis-generating rather than confirmatory, and most did not survive correction for multiple comparisons: of the 18 correlations with gender, age, and education, three survived, and of the 18 race contrasts, two survived. With that caveat, the surviving associations point in a consistent direction: the groups for whom adoption has been greater, including men, those with higher education, and Asian respondents, report the greatest perceived benefits. Future work should explore these associations through confirmatory designs and mechanism studies, to understand how to mitigate such disparities.

5.2 Process Control Is Both Desired and Consequential

A second finding is that users want to feel in control of their AI experiences, and that it is associated with greater reported capability gains. Concerns about over-reliance and loss of agency were among the top fears users expressed, cited by nearly a quarter of all respondents as their primary concern about AI (Figure 8). Users who felt more empowered to shape AI outputs reported significantly greater capability gains across all six domains. This extends the perceived-control literature reviewed above, which has linked control chiefly to satisfaction-related outcomes, rather than reported gains in capability. The uniformity of positive associations suggests that control functions as a general condition of perceived benefit.

Users' own descriptions of when they felt most in control offer inspiration for product designs that may increase perceived process control. The two most prevalent themes from open-ended responses were when they guided AI with prompts and when they evaluated and edited AI outputs before accepting them (see Figure 7). These are active, iterative behaviors that reflect users functioning as informed collaborators rather than passive recipients of information. That users' sense of control tracks with these behaviors, and that the presence of these behaviors runs counter to concerns about over-reliance (see Figure 8), indicates that investment in process control could yield benefits for product developers, beyond improved satisfaction. Students who regard their own agency as important have likewise been found to engage with AI more resourcefully and reflectively (Yang et al. 2024), aligning with our respondents. Such investment could address prominent concerns about AI over-reliance.

The convergence across our survey results, the broader psychological literature on agency (Bandura 2001; Vansteenkiste, Ryan, and Soenens 2020), and early experimental work on perceived control (Oz et al. 2024) lends confidence to the conclusion that process control is a design feature worth investing in substantively. However, our study design cannot determine whether users who reported increased perceived control were actually able to control their usage better. Previous work finding that perceived and functional control can dissociate (Oz et al. 2024) also prompts further need for experiments that manipulate actual control affordances and measure performance. Future work should explore whether process control enables outcome control, outcome gains foster a sense of process control, or some other more complex relationship exists.

Our survey results on experiences with individual products indicate that there is room for improvement. Current levels of perceived process control across products, while associated with greater benefit, are not uniformly high (see Figure 5). Product designers can take the descriptions users offered of experiences where they felt most in control (Figure 7) as early signals about the interaction patterns that feel most empowering. Building these patterns into core product architecture from the outset, rather than retrofitting them after complaints regarding over-reliance, may be one of the more impactful investments AI companies can make. Our study suggests that these investments may shift users from being concerned about AI toward being more excited.

5.3 Agency-Centered Design May Help Bridge the Trust Gap

A broader implication of these findings concerns the divide between technologists' optimism about AI and the general public's wariness (Rainie and Anderson 2025). Surveys consistently show that while AI experts and developers anticipate transformative benefits, the general public remains substantially more skeptical and concerned: 50% of U.S. adults are more concerned than excited about AI (Kennedy et al. 2025), and fears of job loss, manipulation, and reduced human agency are widespread (McClain et al. 2025; Lange and Alper 2025). This gap is not simply a matter of information deficit. Many concerns reflect reasonable observations, held by everyday users and not only by skeptics, about how technology has let people down before.

Insofar as this gap is linked to a lack of agency, the findings here suggest a potential pathway toward closing that gap. First, investing in process control, giving users accessible pathways to influence how AI responds to them, can help build the kind of trust that previous technologies failed to earn. Users reported that they felt more in control when they felt empowered to reject results and iterate on them, and therefore AI systems that explicitly invite such process control may engender more agency. Social media platforms that deployed engagement-driven algorithms without meaningful user controls generated regret, feelings of manipulation, and diminished autonomy (Milli et al. 2025; Zahrai et al. 2022); AI products with features that center users' ability to shape their own experience could offer a different model. AI companies could design products primarily for user control, rather than for increased usage.

Second, designing for the outcomes users actually express wanting, primarily in the instrumental domain, can align products with users' broader aspirations rather than more transient desires, such as eliciting praise and validation from AI products. Explicit disclosure of optimization targets, such as digital transparency requirements by the EU (2024), would allow researchers and regulators to identify products designed toward empowerment. And third, where companies choose to extend AI into social, emotional, or relational domains, intentional and explicitly purposeful design is essential, rather than encouraging the use of general-purpose models for those purposes. In practice, this implies an evidence standard where products with a social or emotional component should articulate and show the intentional benefit they hope to create.

5.4 Limitations and Future Directions

Several limitations should be noted. Our data are cross-sectional, so while we can establish that perceived control, reported capability, and demographics have correlations, we cannot determine the direction of that relationship. Furthermore, all measures are self-reported perceptions, so reported capability gains may also reflect personal satisfaction or expectancy rather than actual performance change. Because predictors and outcomes were collected in a single instrument, common-method variance may inflate our associations. Our results are therefore best interpreted as where users perceive benefit and control but not as evidence of performance gains.

Our measure of process control asked whether respondents felt "in control over the results that are created," which could be interpreted as control over outcomes rather than control over the interaction itself. Users' open-ended descriptions suggest the process reading predominates, with the two most common themes being iteratively guiding AI through prompts (34.1%) and evaluating/editing AI outputs before accepting them (31.5%). Future work should consider a purpose-built measure distinguishing the two constructs.

The open-ended questions also rely on a smaller sample than the closed-ended items; for instance, the excitement and concern questions were answered by roughly 90% of the sample (1,919 and 1,924 analyzable responses) but the goal-achievement question yielded 387 responses (18% of the sample) and the control question 502 (26% of those with an identified most-used tool). Theme prevalences are correspondingly less precise and should be read as indicative, not population estimates.

Product comparisons are confounded by self-selection. As noted in our methodology, respondents choose their own primary tool, so differences between product user bases may also reflect differences between the people who choose each product.

Our sample is restricted to U.S. adults who can speak English and already use generative AI. However, preferences and concerns may differ among other language communities, in other cultural contexts, and among non-users. Post-stratification weighting yields

an effective sample size of approximately 1,153 rather than the nominal 2,126, so our estimates are less precise than the raw sample size suggests. Eligibility was based on having reported AI use on an earlier survey rather than on current use; up to sixteen months elapsed between those responses and our fielding window. Thus, our respondents should be viewed as individuals with AI experience rather than as active users (9.5% reported using none of the listed tools in the past four weeks). Our sample nonetheless underrepresents people who tried AI and abandoned it before the previous surveys. Our estimates should be interpreted as upper bounds on what a broader population of ever-users would report.

Finally, AI products are changing quickly, and perceptions measured at one point may not persist. Future research should employ longitudinal and experimental designs to test causal pathways, examine cross-cultural variation, and develop validated measures of AI-related agency that go beyond self-report.

6 Conclusion

We encourage the continued study of users' explicit aspirations regarding AI, as well as the development and evaluation of design features that provide users with greater control over their experiences. By centering user agency, serving expressed goals and providing meaningful control, AI developers have the opportunity to build technologies that enhance human capabilities without the regret, manipulation, and loss of autonomy that have characterized previous technological revolutions (Doctorow 2025). Doing so will require understanding users' long-term aspirational goals, not just what increases usage. We are hopeful that such efforts can help bridge the divide between the general public's current wariness of AI and many technologists' optimistic view of what AI can one day enable.

References

- Abinesh, R. C., and Rhytheema Dulloo. 2024. "The Impact of AI-Driven Personalization on Customer Satisfaction in E-Commerce: Balancing Technology, Transparency, and Control." *Journal of Computational Analysis and Applications* 33 (2): 649–55. <https://www.eudoxuspress.com/index.php/pub/article/view/382>.
- Adako, Oyeyemi Patricia, Oluwafemi Clement Adeusi, and Peter Adeniyi Alaba. 2025. "Enhancing Education for Children with ASD: A Review of Evaluation and Measurement in AI Tool Implementation." *Disability and Rehabilitation: Assistive Technology* 20 (6): 1578–95. <https://doi.org/10.1080/17483107.2025.2477678>.
- Adam, Martin, Christopher Diebel, Marc Goutier, and Alexander Benlian. 2024. "Navigating Autonomy and Control in Human-AI Delegation: User Responses to Technology-versus User-Invoked Task Allocation." *Decision Support Systems* 180. <https://doi.org/10.1016/j.dss.2024.114193>.
- Alattar, Laith, Matt Messel, and David Rogofsky. 2018. "An Introduction to the Understanding America Study Internet Panel." *Social Security Bulletin* 78 (2). <https://www.ssa.gov/policy/docs/ssb/v78n2/v78n2p13.html>.
- Altman, Sam. 2025. "Sora 2." Sam Altman (blog), September 30, 2025. Accessed August 31, 2026. <https://blog.samaltman.com/sora-2>.
- Angrisani, Marco, Maria Casanova, Nathanael J. Fast, Jimmy Narang, and Juliana Schroeder. 2026. "Gaps in Large Language Model Awareness, Usage, and Perceptions in the United States: Evidence from a Nationally Representative Longitudinal Survey." *PNAS Nexus* 5 (2). <https://doi.org/10.1093/pnasnexus/pgag007>.
- Bai, Jing Yi, IpKin Anthony Wong, Tzung-Cheng T. C. Huan, Fevzi Okumus, and Aliana Man Wai Leong. 2025. "Ethical Perceptions of Generative AI Use and Employee Work Outcomes: Role of Moral Rumination and AI-Supported Autonomy." *Tourism Management* 111. <https://doi.org/10.1016/j.tourman.2025.105242>.
- Bandura, Albert. 2001. "Social Cognitive Theory: An Agentic Perspective." *Annual Review of Psychology* 52 (1): 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>.
- Banerjee, Maya, Daphne Chiew, Keval T. Patel, Ieuan Johns, Digby Chappell, Nick Linton, Graham D. Cole, et al. 2021. "The Impact of Artificial Intelligence on Clinical Education: Perceptions of Postgraduate Trainee Doctors in London (UK) and Recommendations for Trainers." *BMC Medical Education* 21 (1). <https://doi.org/10.1186/s12909-021-02870-x>.
- Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakçı, and Rei Mariman. 2025. "Generative AI without Guardrails Can Harm Learning: Evidence from High School Mathematics." *Proceedings of the National Academy of Sciences* 122 (26). <https://doi.org/10.1073/pnas.2422633122>.

- Brady, William J., Joshua Conrad Jackson, Meriel Doyle, and Silvan Baier. 2025. "Engagement-Based Algorithms Disrupt Human Social Norm Learning." *OSF Preprints*, https://doi.org/10.31219/osf.io/mgdwq_v1.
- Chang, Samuel, James S. Fishkin, Ricky Hernandez Marquez, Ayushi Kadakia, Alice Siu, and Robert Taylor. 2025. *Meta Community Forum: Results Analysis*. Stanford Deliberative Democracy Lab, April. <https://drive.google.com/file/d/1vlw22n2E8ESFbuDjfr6tvSDUmpH90PL1/view>.
- Chatterji, Aaron, Thomas Cunningham, David Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman. 2025. *How People Use ChatGPT*. w34255. National Bureau of Economic Research. <https://doi.org/10.3386/w34255>.
- Chew, Robert, John Bollenbacher, Michael Wenger, Jessica Speer, and Annice Kim. 2023. "LLM-Assisted Content Analysis: Using Large Language Models to Support Deductive Coding." *arXiv*, <https://doi.org/10.48550/arXiv.2306.14924>.
- Clegg, Nick. 2021. "You and the Algorithm: It Takes Two to Tango." Medium, March 31, 2021. Accessed August 31, 2026. <https://nickclegg.medium.com/you-and-the-algorithm-it-takes-two-to-tango-7722b19aa1c2>.
- Cranney, Katelyn, Solène Delecourt, and Rembrand Koning. 2026. *Global Evidence on Gender Gaps and Generative AI over Time*. Working Paper 25-023. Harvard Business School. https://www.hbs.edu/ris/Publication%20Files/25-023_be8fb517-3dd5-40aa-97f9-4e42e1c8e6ff.pdf.
- Darvishi, Ali, Hassan Khosravi, Shazia Sadiq, Dragan Gašević, and George Siemens. 2024. "Impact of AI Assistance on Student Agency." *Computers & Education* 210. <https://doi.org/10.1016/j.compedu.2023.104967>.
- De Freitas, Julian, Zeliha Oğuz-Uğuralp, Ahmet Kaan Uğuralp, and Stefano Puntoni. 2026. "AI Companions Reduce Loneliness." *Journal of Consumer Research* 52 (6): 1126–48. <https://doi.org/10.1093/jcr/ucaf040>.
- Dehbozorgi, Raziye, Sanaz Zangeneh, Elham Khooshab, Donya Hafezi Nia, Hamid Reza Hanif, Pooya Samian, Mahmoud Yousefi, et al. 2025. "The Application of Artificial Intelligence in the Field of Mental Health: A Systematic Review." *BMC Psychiatry* 25 (1). <https://doi.org/10.1186/s12888-025-06483-2>.
- Doctorow, Cory. 2025. *Enshittification: Why Everything Suddenly Got Worse and What to Do about It*. New York: MCD.
- European Parliament and Council of the European Union, Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence ... (Artificial Intelligence Act), Official Journal of the European Union (2024), <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

- Fan, Zhibo, Hongtao Lin, Haoyu Chen, Bowen Deng, Hedi Xia, Yuke Yan, and James Li. 2025. "Synergizing Implicit and Explicit User Interests: A Multi-Embedding Retrieval Framework at Pinterest." In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 4396–405. <https://doi.org/10.1145/3711896.3737265>.
- Fang, Cathy Mengying, Auren R. Liu, Valdemar Danry, Eunhae Lee, Samantha W. T. Chan, Pat Pataranutaporn, Pattie Maes, et al. 2025. "How AI and Human Behaviors Shape Psychosocial Effects of Extended Chatbot Use: A Longitudinal Randomized Controlled Study." *arXiv*, <https://doi.org/10.48550/arXiv.2503.17473>.
- Gerlich, Michael. 2024. "Public Anxieties about AI: Implications for Corporate Strategy and Societal Impact." *Administrative Sciences* 14 (11). <https://doi.org/10.3390/admsci14110288>.
- . 2025. "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking." *Societies* 15 (1). <https://doi.org/10.3390/soc15010006>.
- Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli. 2023. "ChatGPT Outperforms Crowd Workers for Text-Annotation Tasks." *Proceedings of the National Academy of Sciences* 120 (30). <https://doi.org/10.1073/pnas.2305016120>.
- Gottfried, Jeffrey, William Bishop, Monica Anderson, Michelle Faverio, Eugenie Park, and Colleen McClain. 2026. *Americans and AI 2026: Chatbots, Smart Devices and Views on Impact*. Pew Research Center, June 17, 2026. https://www.pewresearch.org/wp-content/uploads/sites/20/2026/06/PI_2026.06.17_Americans-and-AI_REPORT.pdf.
- Guess, Andrew M., Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio, et al. 2023. "How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?" *Science* 381 (6656): 398–404. <https://doi.org/10.1126/science.abp9364>.
- Guo, Longjie, Yue Fu, Xiran Lin, Xuhai Xu, Yung-Ju Chang, and Alexis Hiniker. 2025. "What Social Media Use Do People Regret? An Analysis of 34K Smartphone Screenshots with Multimodal LLM." In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–23. <https://doi.org/10.1145/3706598.3713724>.
- Heinz, Michael V., Daniel M. Mackin, Brianna M. Trudeau, Sukanya Bhattacharya, Yinzhou Wang, Haley A. Banta, Abi D. Jewett, Abigail J. Salzhauer, Tess Z. Griffin, and Nicholas C. Jacobson. 2025. "Randomized Trial of a Generative AI Chatbot for Mental Health Treatment." *NEJM AI* 2 (4). <https://doi.org/10.1056/AIoA2400802>.
- Henry, Kimberly A., and Richard Valliant. 2015. "A Design Effect Measure for Calibration Weighting in Single-Stage Samples." *Survey Methodology* 41 (2): 315–31. <https://www150.statcan.gc.ca/n1/pub/12-001-x/2015002/article/14236-eng.pdf>.

- Hilliard, Elizabeth, Akshaya Jagadeesh, Alex Cook, Steele Billings, Nicholas Skytland, Alicia Llewellyn, Jackson Paull, et al. 2025. "Measuring AI Alignment with Human Flourishing." *arXiv*, <https://doi.org/10.48550/arXiv.2507.07787>.
- Hohenstein, Jess, Rene F. Kizilcec, Dominic DiFranzo, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, Jeffrey Hancock, and Malte F. Jung. 2023. "Artificial Intelligence in Communication Impacts Language and Social Relationships." *Scientific Reports* 13 (1). <https://doi.org/10.1038/s41598-023-30938-9>.
- Jigsaw. 2025. "A New Perspective on Human Agency for the AI Era." Medium, September 10, 2025. Accessed August 31, 2026. <https://medium.com/jigsaw/a-new-perspective-on-human-agency-for-the-ai-era-cd785faab026>.
- Johnson, Bobbie. 2025. "What Is Sora Slop For, Exactly?" *New York Times*, October 19, 2025. <https://www.nytimes.com/2025/10/19/opinion/ai-sora-slop.html>.
- Kennedy, Brian, Eileen Yam, Emma Kikuchi, Isabelle Pula, and Javier Fuentes. 2025. *How Americans View AI and Its Impact on People and Society*. Pew Research Center, September 17, 2025. https://www.pewresearch.org/wp-content/uploads/sites/20/2025/09/PS_2025.9.15_AI-and-its-impact_report.pdf.
- Kim, Byung-Jik, and Julak Lee. 2026. "How Human-Centered AI Buffers the Negative Effects of AI-Induced Job Insecurity on Sustainable Innovation." *Sustainable Development* 34 (5): 6987–7006. <https://doi.org/10.1002/sd.70705>.
- Korsgaard, Christine M. 1983. "Two Distinctions in Goodness." *Philosophical Review* 92 (2): 169–95. <https://doi.org/10.2307/2184924>.
- Lange, Jason, and Alexandra Alper. 2025. "Americans Fear AI Permanently Displacing Workers, Reuters/Ipsos Poll Finds." Reuters, August 19, 2025. Accessed September 24, 2026. <https://www.reuters.com/world/us/americans-fear-ai-permanently-displacing-workers-reutersipsos-poll-finds-2025-08-19/>.
- Lee, Angela Y., and Jeffrey T. Hancock. 2024. "Social Media Mindsets: A New Approach to Understanding Social Media Use and Psychological Well-Being." *Journal of Computer-Mediated Communication* 29 (1). <https://doi.org/10.1093/jcmc/zmad048>.
- Lee, Byung Cheol, and Jaeyeon Chung. 2024. "An Empirical Investigation of the Impact of ChatGPT on Creativity." *Nature Human Behaviour* 8 (10): 1906–14. <https://doi.org/10.1038/s41562-024-01953-1>.
- Li, Han, Renwen Zhang, Yi-Chieh Lee, Robert E. Kraut, and David C. Mohr. 2023. "Systematic Review and Meta-Analysis of AI-Based Conversational Agents for Promoting Mental Health and Well-Being." *npj Digital Medicine* 6 (1). <https://doi.org/10.1038/s41746-023-00979-5>.

- Li, Yi, Xuanxuan Ding, Yifan Chen, Yeye Li, and Nan Ma. 2025. "Customizable AI for Depression Care: Improving the User Experience of Large Language Model-Driven Chatbots." In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 1844–66. <https://doi.org/10.1145/3715336.3735795>.
- McClain, Colleen, Brian Kennedy, Jeffrey Gottfried, Monica Anderson, and Giancarlo Pasquini. 2025. *How the U.S. Public and AI Experts View Artificial Intelligence*. Pew Research Center, April 3, 2025. https://www.pewresearch.org/wp-content/uploads/sites/20/2025/04/pi_2025.04.03_us-public-and-ai-experts_report.pdf.
- Milli, Smitha, Micah Carroll, Yike Wang, Sashrika Pandey, Sebastian Zhao, and Anca D. Dragan. 2025. "Engagement, User Satisfaction, and the Amplification of Divisive Content on Social Media." *PNAS Nexus* 4 (3). <https://doi.org/10.1093/pnasnexus/pgaf062>.
- Mitsea, Eleni, Athanasios Drigas, and Charalabos Skianis. 2023. "Digitally Assisted Mindfulness in Training Self-Regulation Skills for Sustainable Mental Health: A Systematic Review." *Behavioral Sciences* 13 (12). <https://doi.org/10.3390/bs13121008>.
- Muscio, Alessandro, and Sotaro Shibayama. 2025. "Does AI Lead to Significant Skilling or Deskilling of Academics? Results from a Large-Scale Survey." *SSRN*, <https://doi.org/10.2139/ssrn.5572865>.
- Ngwenyama, Ojelanki, Nada Kanita, and Frantz Rowe. 2025. "Can Generative AI Contribute to Both Productivity Gains and Human Flourishing, and in Fine Satisfaction at Work? Research on GitHub Copilot Use in Software Development." In *Proceedings of the 58th Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2025.719>.
- Oz, Burak, Alexander Karran, Jared Boasen, Constantinos Coursaris, and Pierre-Majorique Léger. 2024. "The Role of User Control in Enhancing Human-AI Collaboration Effectiveness: Insights from a Pilot Study." In *Information Systems and Neuroscience*, 185–93. Cham: Springer. https://doi.org/10.1007/978-3-031-58396-4_15.
- Phang, Jason, Michael Lampe, Lama Ahmad, Sandhini Agarwal, Cathy Mengying Fang, Auren R. Liu, Valdemar Danry, et al. 2025. "Investigating Affective Use and Emotional Well-Being on ChatGPT." *arXiv*, <https://doi.org/10.48550/arXiv.2504.03888>.
- Prunkl, Carina. 2024. "Human Autonomy at Risk? An Analysis of the Challenges from AI." *Minds and Machines* 34 (3). <https://doi.org/10.1007/s11023-024-09665-1>.
- Qian, Yufeng. 2025. "Pedagogical Applications of Generative AI in Higher Education: A Systematic Review of the Field." *TechTrends* 69 (5): 1105–20. <https://doi.org/10.1007/s11528-025-01100-1>.

- Rainie, Lee, and Janna Anderson. 2025. *Many Americans Expect AI to Have Significant Negative Impact on Human Capacities and Behaviors Such as Social and Emotional Intelligence, Analytical Thinking and Agency by 2035*. Imagining the Digital Future Center, Elon University, September 17, 2025. Accessed September 24, 2026. <https://imaginingthedigitalfuture.org/wp-content/uploads/2025/08/LLM-survey-public-views-on-being-human-in-2035-9-17-25.pdf>.
- Rządeczka, Marcin, Anna Sterna, Julia Stolińska, Paulina Kaczyńska, and Marcin Moskalewicz. 2025. "The Efficacy of Conversational AI in Rectifying the Theory-of-Mind and Autonomy Biases: Comparative Analysis." *JMIR Mental Health* 12. <https://doi.org/10.2196/64396>.
- Sahebi, Siavosh, and Paul Formosa. 2022. "Social Media and Its Negative Impacts on Autonomy." *Philosophy & Technology* 35 (3). <https://doi.org/10.1007/s13347-022-00567-7>.
- Schulz van Endert, Tim, and Peter N. C. Mohr. 2020. "Likes and Impulsivity: Investigating the Relationship between Actual Smartphone Use and Delay Discounting." *PLOS ONE* 15 (11). <https://doi.org/10.1371/journal.pone.0241383>.
- Sharma, Mrinank, Miles McCain, Raymond Douglas, and David Duvenaud. 2026. "Who's in Charge? Disempowerment Patterns in Real-World LLM Usage." *arXiv*, <https://doi.org/10.48550/arXiv.2601.19062>.
- Siebold, Nicole. 2021. "Reference Points for Business Model Innovation in Social Purpose Organizations: A Stakeholder Perspective." *Journal of Business Research* 125:710–19. <https://doi.org/10.1016/j.jbusres.2020.01.032>.
- Spirlet, Thibault. 2025. "The CEO of Google DeepMind Warns AI Companies Not to Fall into the Same Trap as Early Social Media Firms." *Business Insider*, September 15, 2025. Accessed August 31, 2026. <https://www.businessinsider.com/google-deep-mind-ceo-warns-ai-could-repeat-social-medias-mistakes-2025-9>.
- Synofzik, Matthis, Gottfried Vosgerau, and Martin Voss. 2013. "The Experience of Agency: An Interplay between Prediction and Postdiction." *Frontiers in Psychology* 4. <https://doi.org/10.3389/fpsyg.2013.00127>.
- Tapal, Adam, Ela Oren, Reuven Dar, and Baruch Eitam. 2017. "The Sense of Agency Scale: A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment." *Frontiers in Psychology* 8. <https://doi.org/10.3389/fpsyg.2017.01552>.
- Terzimehić, Nada, Florian Bemmman, Miriam Halsner, and Sven Mayer. 2023. "A Mixed-Method Exploration into the Mobile Phone Rabbit Hole." *Proceedings of the ACM on Human-Computer Interaction* 7 (MHCI): 1–29. <https://doi.org/10.1145/3604241>.

- Tsai, Chiung-Ying, Tien-Chi Huang, Yu Shu, and Yu-Hung Chiang. 2026. "Investigating the Influence of Autonomy, Competence, Relatedness, and Excitement on User Satisfaction with ChatGPT: A Self-Determination Theory Perspective." In *Innovative Technologies and Learning*, 383–92. Cham: Springer. https://doi.org/10.1007/978-3-031-98197-5_41.
- Understanding America Study. n.d. "Weights." University of Southern California, Center for Economic and Social Research. Accessed September 25, 2026. <https://uasdata.usc.edu/page/weights>.
- Vansteenkiste, Maarten, Richard M. Ryan, and Bart Soenens. 2020. "Basic Psychological Need Theory: Advancements, Critical Themes, and Future Directions." *Motivation and Emotion* 44 (1): 1–31. <https://doi.org/10.1007/s11031-019-09818-1>.
- Vieriu, Aniella Mihaela, and Gabriel Petrea. 2025. "The Impact of Artificial Intelligence (AI) on Students' Academic Development." *Education Sciences* 15 (3). <https://doi.org/10.3390/educsci15030343>.
- Xie, Zehang, and Zeyu Wang. 2024. "Longitudinal Examination of the Relationship between Virtual Companionship and Social Anxiety: Emotional Expression as a Mediator and Mindfulness as a Moderator." *Psychology Research and Behavior Management* 17:765–82. <https://doi.org/10.2147/PRBM.S447487>.
- Xu, Yan, and Linmin Chen. 2025. "Personalized Recommendations and Consumer Trust: The Role of Perceived Control and Locus of Control." *Acta Psychologica* 261. <https://doi.org/10.1016/j.actpsy.2025.105936>.
- Xu, Ying, Yenda Prado, Rachel L. Severson, Silvia Lovato, and Justine Cassell. 2025. "Growing Up with Artificial Intelligence: Implications for Child Development." In *Handbook of Children and Screens*, edited by Dimitri A. Christakis and Lauren Hale, 611–17. Cham: Springer. https://doi.org/10.1007/978-3-031-69362-5_83.
- Yang, Yunying, Jinwen Luo, Miaoyan Yang, Runde Yang, and Jiayin Chen. 2024. "From Surface to Deep Learning Approaches with Generative AI in Higher Education: An Analytical Framework of Student Agency." *Studies in Higher Education* 49 (5): 817–30. <https://doi.org/10.1080/03075079.2024.2327003>.
- Zac, Amit, Yu-Chun Huang, Amédée von Moltke, Christopher Decker, and Ariel Ezrachi. 2025. "Dark Patterns and Consumer Vulnerability." *Behavioural Public Policy*, 1–50. <https://doi.org/10.1017/bpp.2024.49>.
- Zahrai, Kseniia, Ekant Veer, Paul William Ballantine, Huibert Peter de Vries, and Girish Prayag. 2022. "Either You Control Social Media or Social Media Controls You: Understanding the Impact of Self-Control on Excessive Social Media Use from the Dual-System Perspective." *Journal of Consumer Affairs* 56 (2): 806–48. <https://doi.org/10.1111/joca.12449>.

- Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. 2024. "Can Large Language Models Transform Computational Social Science?" *Computational Linguistics* 50 (1): 237–91. https://doi.org/10.1162/coli_a_00502.
- Zimmerman, Federico, David D. Bailey, Goran Muric, Emilio Ferrara, Jonas Schöne, Robb Willer, Eran Halperin, Joaquín Navajas, James J. Gross, and Amit Goldenberg. 2024. "Attraction to Politically Extreme Users on Social Media." *PNAS Nexus* 3 (10). <https://doi.org/10.1093/pnasnexus/pgae395>.

Authors

Emily Xing is an undergraduate at Harvard University, where she studies intersections of media, technology, and human health.

Ravi Iyer is Managing Director of the Neely Center for Ethical Leadership and Decision Making at the University of Southern California's Marshall School of Business, where he studies technology design and its societal impact. Email: raviyer@marshall.usc.edu.

Rachel Xu is a Research Manager at Google Jigsaw, where she studies misinformation resilience and interventions that support user agency online.

Yuning Liu is a Postdoctoral Researcher at Harvard Business School, where she studies digital well-being and user agency.

Beth Goldberg is Head of Research and Development at Google Jigsaw, where she studies online harms and strategies for mitigating them.

Nathanael J. Fast is Professor of Management and Organization at the University of Southern California's Marshall School of Business and Director of the Neely Center for Ethical Leadership and Decision Making, where he studies power, hierarchy, and the psychology of technology.

Acknowledgments

We thank the staff of the Center for Economic and Social Research at the University of Southern California for programming and fielding UAS 738.

Data and Code Availability

The UAS 738 survey data are available to researchers through the Understanding America Study at <https://uasdata.usc.edu> under the panel's data access terms. Code for the quantitative analyses and figures is available at <https://github.com/em-xing/xing2026ai>. The LLM-assisted thematic analysis pipeline is available at https://github.com/Yuning-Liu-712/thematic_analysis_AI_agency. No preanalysis plan was registered.

Ethics Statement

The Understanding America Study and its surveys operate under University of Southern California IRB approval #UP-14-00148. Panel members provide informed consent on joining the panel and may decline any survey or any individual item. The released analytic file contains no direct identifiers; respondents are identified only by a panel ID.

Funding Statement

This work was funded by Google Jigsaw. The authors received additional support through the Neely Center at the University of Southern California, which receives funding from

the Knight Foundation and the Omidyar Foundation.

Disclosure Statement

Rachel Xu and Beth Goldberg were employed by Google Jigsaw at the time of writing. The study reports comparative results for several generative AI products, including Google's Gemini.

AI/LLMs were used to support the creation of this manuscript, including in the coding of qualitative data, management of references, suggestions for sentence restructuring, and proofreading.

Keywords

Agency; human-AI interaction; user preferences; user control; technology design

Appendix

A. Method: LLM-Assisted Thematic Analysis and Deductive Classification with Human-in-the-Loop Validation

We developed a human-in-the-loop LLM-assisted pipeline combining inductive theme generation with deductive text classification. The pipeline produced two types of codebooks: a shared codebook capturing themes that emerged across the full sample, and three subgroup-specific codebooks capturing themes unique to Older Adults (ages 65+), Young Women (ages 18–35), and Young Men (ages 18–35). Each codebook entry includes a theme label, a structured definition, and three representative example responses. Responses were not mutually exclusive. Resulting theme distributions are visualized and reported in the main manuscript and the appendix.

A.1. Phase 0: Data Loading and Preprocessing

Raw open-text survey responses are loaded from the UAS panel dataset and filtered to the target question. Responses are split into four subsets: the full sample (all respondents) and the three demographic subgroups. Responses below a minimum length threshold (10 characters) are discarded.

A.2. Phase 1: Construct Shared Codebook

A shared thematic codebook is built from the full sample of responses. The LLM (gpt-5.4-mini, temperature=0) is called three times in parallel, each time with a different random seed, processing responses in chunks of 100. Each run independently generates a set of thematic labels with names and definitions, reducing sensitivity to any single prompt execution.

The three run-level codebooks are merged into a consensus codebook via a dedicated LLM call that identifies overlapping and complementary themes across runs. A further simplification call removes redundancy and standardizes label wording in the consensus codebook.

Human Checkpoint 1 (CP1): At this stage, the codebook is prepared for human review. Prior to this step, an LLM-based Merge Advisor (gpt-5.4-mini, temperature=0) is invoked to identify semantically redundant labels and propose potential consolidations. These merging suggestions are stored in a JSON file and displayed in the console for researcher evaluation. Researchers then review, revise, and finalize the codebook.

A.3. Phase 2: Prevalence Pruning

To remove rare themes that carry little analytical weight, a single-run classification pass applies the Phase 1 codebook to all responses. Any label assigned to fewer than 3% of coded responses is dropped. Surviving labels are renumbered sequentially (D1, D2, ...). This produces the pruned shared codebook.

Human Checkpoint 2 (CP2): The pruned codebook, along with up to five example responses per label and the total response count (N), is saved to both a JSON and a CSV file, which are automatically opened for the researcher. The pipeline pauses until the researcher edits the JSON and confirms the codebook. The confirmed codebook (shared_codebook) is saved as a REVIEWED file. On any subsequent run, if the REVIEWED file exists, this checkpoint is skipped automatically.

A.4. Phase 3: Majority-Vote Classification of Shared Themes

The reviewed codebook (shared_codebook) is applied to each demographic subgroup (all, old, young_female, young_male) independently. For each subgroup, the LLM (gpt-5.4-mini, temperature=0) classifies every response five times using different random seeds. Each classification run assigns one or more label codes to each response, or assigns "None" if no label applies. For each response, we record the following information across the five runs:

- Majority vote: a label code is included in the final assignment for a response only if it appears in at least 3 of the 5 runs. This reduces the influence of idiosyncratic single-run decisions and increases classification stability.
- Agreement score: for each response, the proportion of runs that agree exactly on the full label set is recorded as an agreement score.
- Stability flag: responses with agreement below 0.6 are flagged as "unstable"; those between 0.6 and 0.8 as "uncertain"; those at or above 0.8 as "stable".
- Label entropy: per-response Shannon entropy over the distribution of distinct label-set outcomes across the five runs is computed as a supplementary reliability metric.

Based on the results, a stability report summarizing the mean agreement, percentage of unstable responses, and mean label entropy for each label is saved for each demographic subgroup.

A.5. Phase 4: Human Validation

When enabled, a stratified random sample of 100 responses is drawn from the full-sample labeled output, with proportional representation of each label. The LLM-assigned labels are withheld from the human-facing CSV to prevent anchoring bias. The researcher labels each response independently; then the LLM and human labels are merged automatically.

Inter-rater agreement metrics (precision, recall, and F1) are computed using scikit-learn and saved as a metrics CSV.

A.6. Phase 5: Unique Theme Detection for Each Demographic

For each of the three demographic subgroups (people with age 65+, female with age 18–35, male with age 18–35), the pipeline identifies themes present in that group’s responses that are not already captured by the shared codebook. This process runs independently for each group, producing fully separate unique codebooks with their own U-coded labels (U1, U2, ...). The steps are as follows:

1. Step 1—Gap screening: Every response in the demographic subgroup is evaluated by the LLM in batches of 10. For each response, the LLM (gpt-5.4-mini, temperature=0) is asked whether it contains substantive content not captured by the shared codebook. Responses are flagged liberally and a brief gap description is recorded for each flagged response.
2. Step 2—Candidate theme clustering: Gap descriptions collected from all flagged responses are treated as a new text corpus and passed through the same three-run induction, consensus-merge, and simplification pipeline used in Phase 1. This produces a set of candidate unique themes.
3. Step 3—Validation against shared codebook: The candidate unique themes are reviewed by the LLM (gpt-5.4-mini, temperature=0) against the shared codebook. Any candidate that substantially overlaps with or is already covered by a shared label is discarded. Only genuinely novel themes are retained.
4. Step 4—Frequency pruning: A single-run classification (gpt-5.4-mini, temperature=0) applies the validated unique themes to all responses in the subgroup. Any theme assigned to fewer than 5 responses is dropped. Surviving themes are renumbered U1, U2, ... This step ensures that only themes with sufficient empirical support are presented for human review.

Human Checkpoint 3 (CP3): The pruned unique-theme codebook, along with example gap descriptions that contributed to each theme, is presented to the researcher. The researcher can delete false positives, rename labels, and rewrite definitions. The confirmed codebook is saved as a REVIEWED file. In total, three unique codebooks were saved for each of the demographics.

Harmonization Checkpoint (CP4): After all three demographic unique codebooks have been reviewed independently at CP3, a cross-demographic harmonization step opens all three REVIEWED files simultaneously alongside a combined view. The researcher checks that themes appearing across multiple groups use consistent label names and definitions, and edits the individual files directly. Once confirmed, the harmonized codebooks are saved. All subsequent steps use the harmonized version (unique_codebook).

A.7. Phase 6: Majority-Vote Classification of Unique Themes

The harmonized unique-theme codebook (unique_codebook) for each demographic subgroup is applied to that subgroup's responses using the same five-run majority-vote classifier described in Phase 3. Agreement scores, stability flags, and label entropy are computed and saved in a stability report. The resulting labeled CSV contains the unique-theme assignments for every response in the subgroup.

A.8. Phase 7: Figures, Tables, and Summary Outputs

After all classification steps are complete, the following outputs are generated:

- Figure 1: A four-panel horizontal bar chart comparing shared theme prevalence across all four groups (All, Older Adults, Young Female, Young Male), with themes ordered by full-sample prevalence. These per-group prevalence charts were generated by the pipeline but are not reproduced here; the corresponding subgroup comparisons are reported as significance tests in Table A5.
- Figure 2: A three-panel horizontal bar chart displaying unique themes for each demographic group, with theme name, prevalence within the subgroup, and response count (Figure A1 and Figure A2).
- Excel codebook table: A four-sheet spreadsheet with the shared-theme codebook and one sheet per demographic unique codebook, including label codes, names, definitions, N counts, and up to three example responses per label.

Summary performance of the pipeline: a plain-text results summary, generated for each question, documenting the ranked prevalence of all shared themes in the full sample, per-label reliability statistics derived from the five-run majority-vote classification (mean agreement score, percentage of unstable assignments, and label entropy), and the demographic-specific unique themes with their assigned response counts and mean classification agreement.

B. Performance of LLM-assisted thematic analysis pipeline

B.1. Question sc003

Question wording: In a sentence or two, please tell us what you are most excited about regarding AI (for yourself or society at large).

B.1.1. Shared Themes (full sample)

We identified 16 shared theme(s) across respondents.

D1—Information access: 20.7% (n=398)

D3—Saving time: 17.4% (n=333)

D5—Task automation: 12.5% (n=239)

D4—Reducing cognitive load: 11.5% (n=220)

D11—Uncertainty and concerns: 10.5% (n=202)

D7—Learning support: 10.4% (n=200)

D6—Reading and writing support: 8.4% (n=161)

D10—Future transformation: 8.1% (n=155)

D16—Workflow and job support: 6.1% (n=118)

D2—Summarizing information: 5.6% (n=107)

D8—Healthcare advances: 5.5% (n=106)

D13—Support creativity: 4.6% (n=89)

D14—Economic and career gains: 4.6% (n=88)

D12—Democratization of knowledge: 3.7% (n=71)

D15—General curiosity: 3.5% (n=67)

D9—Scientific progress: 3.4% (n=65)

B.1.2. Classification Stability (shared codebook, full sample)

Mean Agreement (the average proportion of the 5 classification runs that fully agree with the majority-vote label assignment for each response) ranged from 0.79 (D16—Workflow and job support) to 0.95 (D15—General curiosity), with 94% of themes achieving Mean Agreement $\geq 80\%$. [] indicates the theme met the stability threshold (Mean Agreement $\geq 80\%$). [UNCERTAIN] indicates Mean Agreement between 60%-80%; [UNSTABLE] indicates Mean Agreement below 60%.

Per-label stability:

D1—Information access: Mean Agreement=0.91, % Unstable=2.8%, Label Entropy=0.26
[]

D2—Summarizing information: Mean Agreement=0.87, % Unstable=5.6%, Label Entropy=0.38 []

D3—Saving time: Mean Agreement=0.88, % Unstable=3.9%, Label Entropy=0.33
[]

D4—Reducing cognitive load: Mean Agreement=0.89, % Unstable=3.2%, Label Entropy=0.31 []

D5—Task automation: Mean Agreement=0.89, % Unstable=3.8%, Label Entropy=0.33 []

D6—Reading and writing support: Mean Agreement=0.89, % Unstable=3.1%, Label Entropy=0.31 []

D7—Learning support: Mean Agreement=0.88, % Unstable=6.0%, Label Entropy=0.33 []

D8—Healthcare advances: Mean Agreement=0.91, % Unstable=3.8%, Label Entropy=0.27 []

D9—Scientific progress: Mean Agreement=0.92, % Unstable=0.0%, Label Entropy=0.26 []

D10—Future transformation: Mean Agreement=0.89, % Unstable=4.5%, Label Entropy=0.33 []

D11—Uncertainty and concerns: Mean Agreement=0.94, % Unstable=1.0%, Label Entropy=0.18 []

D12—Democratization of knowledge: Mean Agreement=0.86, %Unstable=4.2%, Label Entropy=0.39 []

D13—Support creativity: Mean Agreement=0.91, % Unstable=4.5%, Label Entropy=0.27 []

D14—Economic and career gains: Mean Agreement=0.81, % Unstable=5.7%, Label Entropy=0.54 []

D15—General curiosity: Mean Agreement=0.95, % Unstable=0.0%, Label Entropy=0.17 []

D16—Workflow and job support: Mean Agreement=0.79, % Unstable=5.9%, Label Entropy=0.61 []

B.1.3. Human-AI Validation

Comparing human and LLM-judge labels on a stratified sample of 100 responses: F1=0.740.

Per-label metrics:

D1—Information access: Precision=0.950, Recall=0.864, F1=0.905, Support=22

D2—Summarizing information: Precision=0.778, Recall=1.000, F1=0.875, Support=7

D3—Saving time: Precision=0.583, Recall=0.778, F1=0.667, Support=18

D4—Reducing cognitive load: Precision=0.833, Recall=0.714, F1=0.769, Support=14

D5—Task automation: Precision=0.875, Recall=0.636, F1=0.737, Support=11

D6—Reading and writing support: Precision=0.917, Recall=0.846, F1=0.880, Support=13

D7—Learning support: Precision=0.769, Recall=0.909, F1=0.833, Support=11

D8—Healthcare advances: Precision=1.000, Recall=0.500, F1=0.667, Support=2

D9—Scientific progress: Precision=1.000, Recall=0.667, F1=0.800, Support=3

D10—Future transformation: Precision=0.800, Recall=0.889, F1=0.842, Support=9

D11—Uncertainty and concerns: Precision=1.000, Recall=1.000, F1=1.000, Support=10

D12—Democratization of knowledge: Precision=1.000, Recall=0.600, F1=0.750, Support=5

D13—Support creativity: Precision=1.000, Recall=0.750, F1=0.857, Support=8

D14—Economic and career gains: Precision=0.000, Recall=0.000, F1=0.000, Support=3

D15—General curiosity: Precision=0.333, Recall=0.333, F1=0.333, Support=3

D16—Workflow and job support: Precision=1.000, Recall=0.857, F1=0.923, Support=7

B.1.4. Unique Themes By Demographic

Old (65+) 3 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 90.4%.

Themes:

1. U12—Deepening ideas (n assigned=14)
2. U15—Conversational understanding (n assigned=18)
3. U18—Cross-domain pattern finding (n assigned=38)

Young Female (18-35) 3 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 79.6%.

Themes:

1. U6—Perspective-taking (n assigned=8)
2. U13—Reduced social friction (n assigned=31)
3. U23—Human-AI collaboration (n assigned=95)

Young Male (18-35) 2 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 94.0%.

Themes:

1. U5—Human role reallocation (n assigned=70)
2. U7—Technical problem solving (n assigned=70)

B.2. Question sc004

Question wording: In a sentence or two, please tell us your biggest concern regarding AI (for yourself or society at large).

B.2.1. Shared Themes (full sample)

We identified 14 shared theme(s) across respondents.

Themes ranked from most to least prevalent:

- D1—Reduced thinking and agency: 24.3% (n=467)
- D8—Manipulation and malicious use: 19.6% (n=378)
- D3—Job loss and automation: 16.2% (n=311)
- D5—Inaccurate or fabricated outputs: 15.5% (n=298)
- D10—Loss of human control: 13.1% (n=252)
- D11—Reduced human and social connection: 8.7% (n=167)
- D4—Privacy and data misuse: 6.5% (n=126)
- D12—Environmental and energy costs: 5.5% (n=106)
- D13—Academic misuse: 5.5% (n=105)
- D2—Reduced practical skills: 5.4% (n=104)

D14—Authenticity of expression: 4.1% (n=78)

D6—Blurred reality: 3.9% (n=75)

D7—Social inequality and power concentration: 3.6% (n=69)

D9—Weak oversight and governance: 3.5% (n=67)

B.2.2. Classification Stability (shared codebook, full sample)

Mean Agreement (the average proportion of the 5 classification runs that fully agree with the majority-vote label assignment for each response) ranged from 0.87 (D2—Reduced practical skills) to 0.96 (D3—Job loss and automation), with 100% of themes achieving Mean Agreement $\geq 80\%$. [] indicates the theme met the stability threshold (Mean Agreement $\geq 80\%$). [UNCERTAIN] indicates Mean Agreement between 60%-80%; [UNSTABLE] indicates Mean Agreement below 60%.

Per-label stability:

D1—Reduced thinking and agency: Mean Agreement=0.94, % Unstable=2.1%, Label Entropy=0.19 []

D2—Reduced practical skills: Mean Agreement=0.87, % Unstable=1.9%, Label Entropy=0.36 []

D3—Job loss and automation: Mean Agreement=0.96, % Unstable=0.6%, Label Entropy=0.11 []

D4—Privacy and data misuse: Mean Agreement=0.95, % Unstable=1.6%, Label Entropy=0.15 []

D5—Inaccurate or fabricated outputs: Mean Agreement=0.93, %Unstable=2.0%, Label Entropy=0.19 []

D6—Blurred reality: Mean Agreement=0.89, % Unstable=1.3%, Label Entropy=0.33 []

D7—Social inequality and power concentration: Mean Agreement=0.90, % Unstable=2.9%, Label Entropy=0.29 []

D8—Manipulation and malicious use: Mean Agreement=0.92, % Unstable=1.6%, Label Entropy=0.22 []

D9—Weak oversight and governance: Mean Agreement=0.91, % Unstable=0.0%, Label Entropy=0.25 []

D10—Loss of human control: Mean Agreement=0.94, % Unstable=2.0%, Label Entropy=0.17 []

D11—Reduced human and social connection: Mean Agreement=0.93, % Unstable=4.2%, Label Entropy=0.20 []

D12—Environmental and energy costs: Mean Agreement=0.94, %Unstable=1.9%, Label Entropy=0.15 []

D13—Academic misuse: Mean Agreement=0.94, % Unstable=1.9%, Label Entropy=0.16 []

D14—Authenticity of expression: Mean Agreement=0.89, % Unstable=2.6%, Label Entropy=0.32 []

B.2.3. Human-AI Validation

Comparing human and LLM-judge labels on a stratified sample of 100 responses: F1=0.841.

Per-label metrics:

D1—Reduced thinking and agency: Precision=0.857, Recall=0.882, F1=0.870, Support=34

D2—Reduced practical skills: Precision=0.545, Recall=0.750, F1=0.632, Support=8

D3—Job loss and automation: Precision=0.962, Recall=0.926, F1=0.943, Support=27

D4—Privacy and data misuse: Precision=0.917, Recall=1.000, F1=0.957, Support=11

D5—Inaccurate or fabricated outputs: Precision=0.966, Recall=0.903, F1=0.933, Support=31

D6—Blurred reality: Precision=0.750, Recall=0.750, F1=0.750, Support=8

D7—Social inequality and power concentration: Precision=0.857, Recall=0.667, F1=0.750, Support=9

D8—Manipulation and malicious use: Precision=0.941, Recall=0.889, F1=0.914, Support=36

D9—Weak oversight and governance: Precision=1.000, Recall=0.667, F1=0.800, Support=6

D10—Loss of human control: Precision=0.933, Recall=1.000, F1=0.966, Support=14

D11—Reduced human and social connection: Precision=1.000, Recall=0.917, F1=0.957, Support=12

D12—Environmental and energy costs: Precision=1.000, Recall=1.000, F1=1.000, Support=10

D13—Academic misuse: Precision=1.000, Recall=0.818, F1=0.900, Support=11

D14—Authenticity of expression: Precision=1.000, Recall=0.250, F1=0.400, Support=4

B.2.4. Unique Themes By Demographic

Old (65+) 5 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 88.5%.

Themes:

1. U1—Belief reinforcement (n assigned=42)
2. U2—Deceptive self-presentation (n assigned=6)
3. U3—Lack of AI content disclosure (n assigned=7)
4. U4—AI as tool, not person (n assigned=9)
5. U5—Pseudo-relationships and emotional dependency (n assigned=14)

Young Female (18-35) 8 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 91.0%.

Themes:

1. U1—Erosion of human creativity (n assigned=16)
2. U2—Emotional dependence and mental health harm (n assigned=36)
3. U3—Intellectual property and creator rights (n assigned=6)
4. U4—Identity theft and impersonation (n assigned=19)
5. U5—Inadequate social protection (n assigned=70)
6. U6—Premature deployment (n assigned=47)
7. U7—Loss of humanity (n assigned=47)
8. U8—Decline in work ethic (n assigned=86)

Young Male 4 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 91.8%.

Themes:

1. U1—Polarization and echo chambers (n assigned=12)
2. U2—Low-quality content flood (n assigned=4)
3. U3—Erosion of human creativity (n assigned=53)
4. U4—Covert surveillance (n assigned=9)

B.3. Question ai002b

Question wording: In a sentence or two, please describe a time when AI made you feel more able to reach your goals.

B.3.1. Shared Themes (full sample)

We identified 14 shared theme(s) across respondents.

Themes ranked from most to least prevalent:

- D1—Writing, editing, and polishing: 26.6% (n=103)
- D3—Finding and summarizing information: 20.9% (n=81)
- D5—Task and workflow management: 15.0% (n=58)
- D4—Organizing ideas and planning: 14.2% (n=55)
- D14—Personal decisions and lifestyle choices: 12.1% (n=47)
- D11—Solving practical problems: 9.6% (n=37)
- D12—Technical and coding assistance: 6.2% (n=24)
- D6—School and academic support: 5.9% (n=23)
- D13—Job search and career advancement: 5.4% (n=21)
- D8—Health and wellness guidance: 5.4% (n=21)
- D7—Learning and skill development: 5.4% (n=21)
- D10—Brainstorming and idea generation: 5.4% (n=21)
- D9—Creative production: 5.2% (n=20)
- D2—Adjusting communication tone: 4.9% (n=19)

B.3.2. Classification Stability (shared codebook, full sample)

Mean Agreement (the average proportion of the 5 classification runs that fully agree with the majority-vote label assignment for each response) ranged from 0.84 (D6—School and academic support) to 0.94 (D7—Learning and skill development), with 100% of themes achieving Mean Agreement $\geq 80\%$. [] indicates the theme met the stability threshold (Mean Agreement $\geq 80\%$). [UNCERTAIN] indicates Mean Agreement between 60%-80%; [UNSTABLE] indicates Mean Agreement below 60%.

Per-label stability:

D1—Writing, editing, and polishing: Mean Agreement=0.93, % Unstable=1.9%, Label Entropy=0.21 []

D2—Adjusting communication tone: Mean Agreement=0.94, % Unstable=0.0%, Label Entropy=0.20 []

D3—Finding and summarizing information: Mean Agreement=0.88, % Unstable=4.9%, Label Entropy=0.34 []

D4—Organizing ideas and planning: Mean Agreement=0.87, % Unstable=3.6%, Label Entropy=0.36 []

D5—Task and workflow management: Mean Agreement=0.88, % Unstable=6.9%, Label Entropy=0.35 []

D6—School and academic support: Mean Agreement=0.84, % Unstable=8.7%, Label Entropy=0.43 []

D7—Learning and skill development: Mean Agreement=0.94, % Unstable=0.0%, Label Entropy=0.20 []

D8—Health and wellness guidance: Mean Agreement=0.89, % Unstable=0.0%, Label Entropy=0.36 []

D9—Creative production: Mean Agreement=0.89, % Unstable=0.0%, Label Entropy=0.32 []

D10—Brainstorming and idea generation: Mean Agreement=0.92, % Unstable=4.8%, Label Entropy=0.22 []

D11—Solving practical problems: Mean Agreement=0.89, % Unstable=2.7%, Label Entropy=0.35 []

D12—Technical and coding assistance: Mean Agreement=0.90, % Unstable=0.0%, Label Entropy=0.28 []

D13—Job search and career advancement: Mean Agreement=0.93, % Unstable=4.8%,

Label Entropy=0.19 []

D14—Personal decisions and lifestyle choices: Mean Agreement=0.89, % Unstable=2.1%, Label Entropy=0.31 []

B.3.3. Human-AI Validation

Comparing human and LLM-judge labels on a stratified sample of 100 responses: F1=0.776.

Per-label metrics:

D1—Writing, editing, and polishing: Precision=0.882, Recall=0.811, F1=0.845, Support=37

D2—Adjusting communication tone: Precision=0.727, Recall=0.800, F1=0.762, Support=10

D3—Finding and summarizing information: Precision=0.778, Recall=0.677, F1=0.724, Support=31

D4—Organizing ideas and planning: Precision=0.958, Recall=0.920, F1=0.939, Support=25

D5—Task and workflow management: Precision=0.806, Recall=0.806, F1=0.806, Support=31

D6—School and academic support: Precision=0.818, Recall=0.750, F1=0.783, Support=12

D7—Learning and skill development: Precision=0.600, Recall=0.600, F1=0.600, Support=5

D8—Health and wellness guidance: Precision=0.900, Recall=0.900, F1=0.900, Support=10

D9—Creative production: Precision=0.250, Recall=0.500, F1=0.333, Support=2

D10—Brainstorming and idea generation: Precision=0.857, Recall=0.857, F1=0.857, Support=7

D11—Solving practical problems: Precision=0.737, Recall=0.700, F1=0.718, Support=20

D12—Technical and coding assistance: Precision=1.000, Recall=0.909, F1=0.952, Support=11

D13—Job search and career advancement: Precision=0.909, Recall=0.909, F1=0.909,

Support=11

D14—Personal decisions and lifestyle choices: Precision=1.000, Recall=0.579, F1=0.733, Support=19

B.3.4. Unique Themes By Demographic

Old (65+) 2 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 97.1%.

Themes:

1. U1—Persuasive writing support (n assigned=7)
2. U2—Option comparison support (n assigned=6)

Young Male (18-35) 1 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 97.1%.

Themes:

1. U1—Financial management (n assigned=7)

B.4. Question as001b

Question wording: In a sentence or two, please describe a time when you used [most used AI], where you felt very much in control.

B.4.1. Shared Themes (full sample)

We identified 11 shared theme(s) across respondents.

Themes ranked from most to least prevalent:

- D1—When I guide AI with prompts: 34.1% (n=171)
- D2—When I evaluate and edit AI output: 31.5% (n=158)
- D4—When AI help me write or polish text: 19.1% (n=96)
- D6—When I use AI to find information: 14.5% (n=73)
- D3—When I used AI as a tool for a task: 13.5% (n=68)
- D10—When my own expertise kept me in control: 8.4% (n=42)
- D7—When AI help me generate ideas: 6.0% (n=30)

D5—When AI tailor its response to my needs: 5.6% (n=28)

D8—When AI inform my decision: 5.6% (n=28)

D11—When AI helped me organize information: 4.8% (n=24)

D9—When AI support my learning: 3.6% (n=18)

B.4.2. Classification Stability (shared codebook, full sample)

Mean Agreement (the average proportion of the 5 classification runs that fully agree with the majority-vote label assignment for each response) ranged from 0.83 (D3—When i used ai as a tool for a task) to 0.92 (D6—When i use ai to find information), with 100% of themes achieving Mean Agreement $\geq 80\%$. [] indicates the theme met the stability threshold (Mean Agreement $\geq 80\%$). [UNCERTAIN] indicates Mean Agreement between 60%-80%; [UNSTABLE] indicates Mean Agreement below 60%.

Per-label stability:

D1—When i guide ai with prompts: Mean Agreement=0.91, % Unstable=4.1%, Label Entropy=0.24 []

D2—When i evaluate and edit ai output: Mean Agreement=0.88, %Unstable=3.2%, Label Entropy=0.32 []

D3—When i used ai as a tool for a task: Mean Agreement=0.83, % Unstable=8.8%, Label Entropy=0.45 []

D4—When ai help me write or polish text: Mean Agreement=0.90, % Unstable=1.0%, Label Entropy=0.28 []

D5—When ai tailor its response to my needs: Mean Agreement=0.86, % Unstable=3.6%, Label Entropy=0.39 []

D6—When i use ai to find information: Mean Agreement=0.92, % Unstable=1.4%, Label Entropy=0.26 []

D7—When ai help me generate ideas: Mean Agreement=0.89, % Unstable=6.7%, Label Entropy=0.25 []

D8—When ai inform my decision: Mean Agreement=0.89, % Unstable=3.6%, Label Entropy=0.32 []

D9—When ai support my learning: Mean Agreement=0.89, % Unstable=5.6%, Label Entropy=0.23 []

D10—When my own expertise kept me in control: Mean Agreement=0.83, % Unstable=2.4%, Label Entropy=0.47 []

D11—When ai helped me organize information: Mean Agreement=0.83, % Unstable=4.2%, Label Entropy=0.49 []

B.4.3. Human-AI Validation

Comparing human and LLM-judge labels on a stratified sample of 100 responses: F1=0.745.

Per-label metrics:

D1—When I guide AI with prompts: Precision=0.815, Recall=0.863, F1=0.838, Support=51

D2—When I evaluate and edit AI output: Precision=0.889, Recall=0.696, F1=0.780, Support=46

D3—When I used AI as a tool for a task: Precision=0.809, Recall=0.567, F1=0.667, Support=30

D4—When AI help me write or polish text: Precision=0.867, Recall=0.703, F1=0.776, Support=37

D5—When AI tailor its response to my needs: Precision=0.300, Recall=0.375, F1=0.333, Support=8

D6—When I use AI to find information: Precision=0.750, Recall=0.522, F1=0.615, Support=23

D7—When AI help me generate ideas: Precision=0.923, Recall=1.000, F1=0.960, Support=12

D8—When AI inform my decision: Precision=1.000, Recall=1.000, F1=1.000, Support=6

D9—When AI support my learning: Precision=1.000, Recall=0.857, F1=0.923, Support=7

D10—When my own expertise kept me in control: Precision=0.800, Recall=0.667, F1=0.727, Support=12

D11—When AI helped me organize information: Precision=0.800, Recall=0.444, F1=0.571, Support=9

B.4.4. Unique Themes By Demographic

Young Female 1 unique theme(s) identified.

Mean majority-vote agreement across unique themes: 92.6%.

Themes:

1. U1—When AI provides clear answers (n assigned=43)

C. AI and Agency Survey Question Text

C.1. Baseline agency

This section measures participants' baseline level of perceived agency in their lives, and how that has changed in recent years.

#	Question text	Response options
	[Introduction text] Read each of the following items carefully, thinking about how it relates to your life, and then indicate how capable/able you are and whether you feel that over the past year, this has changed. How able are you to do this?	Over the past year, have you become more or less able?
1	I am able to trust my own judgement. [Cognitive - decision making]	Not at all able A little able Somewhat able Often able Very able Much less able A little less able Neither more nor less able A little more able Much more able
2	I am able to control my feelings. [Affective]	
3	I am able to learn anything that is useful or important [Instrumental / Informational]	
4	I am able to make new friends [Social]	
5	I am able to get people in power to listen to what I say. [Structural]	
6	I am able to handle myself well in social gatherings. [Social]	
7	I am able to get help from formal support systems (organizations, companies, government, etc.) when I have a problem. [Structural]	
8	I am able to get over negative emotions when I lose out on something. [Affective]	
9	I am able to complete the tasks I set out to achieve. [Instrumental / Informational]	
10	I am able to break down complex problems into simpler ones. [Cognitive - clarity and organization of thought]	

#	Question text	Response options
11	I am able to focus my attention on a task for more than a few minutes at a time. [Cognitive - ability to focus]	
12	I am able to make up my own mind on things that are important. [Cognitive - general]	
13	I am able to organize my own thinking. [Cognitive - general]	

C.2. AI Usage

This section measures participants' current use of AI in their day to day lives.

#	Question text	Response options
	[Introduction text] Our next questions ask you about your opinions and experiences with AI, which are computer programs that are designed to learn tasks that humans typically do. As with previous questions, your answers will be kept anonymous by the researchers of this study and will not be shared with anyone you know.	
1	In the past 4 weeks, which of the following genAI tools have you used? Select all that apply.	Gemini ChatGPT Claude Microsoft Copilot Perplexity Jasper Deepseek Meta AI / Llama Character.AI Replika DALL-E Grammarly ElevenLabs Poe DeepAI Grok Prezi Image Creator from Microsoft Designer Gretel.ai Midjourney Murf.AI QuillBot Remini Photo Enhancer SnapAI DreamStudio Stable Diffusion Other, please specify [Open text]

#	Question text	Response options
2	[Grid] For each of the tools you indicated using, how often did you use them? [Ask for each tool selected in item 1]	Daily Several times per week A few times or less Not at all
3	[Grid] For each of the tools you indicated using, what did you use them for? Select all that apply. [Present in a grid with tool across the top and use case down the LHS]	Out of curiosity For entertainment As a companion or friend To learn something new about the world To help with tasks at a job To complete a homework assignment for school To better understand a new word or concept To generate income To learn about a specific health condition or treatment To create content for social media To assist with personal tasks, such as planning activities or getting gift ideas, etc. To improve communications (for instance, help in writing emails, letters, etc.) To help work through thoughts or emotions Other, please specify [Open text]

C.3. Agency with AI [Life]

This section measures participants' level of perceived agency in their usage of AI. These questions are adapted from the Basic Psychological Needs Scale for Technology Use.

#	Question text	Response options
1	As a result of using AI in your life, do you feel more or less able to make decisions confidently? [cognitive] [If selected "Much less / a little less"] In a sentence or two, please describe a time when AI made you feel less able to make decisions confidently. [If selected "Much more / a little more / neither more nor less"] In a sentence or two, please describe a time when AI made you feel more able to make decisions confidently.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]

#	Question text	Response options
2	As a result of using AI in your life, do you feel more or less able to reach your goals? [general] [If selected "Much less / a little less"] In a sentence or two, please describe a time when AI made you feel less able to reach your goals. [If selected "Much more / a little more / neither more nor less"] In a sentence or two, please describe a time when AI made you feel more able to reach your goals.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]
3	As a result of using AI in your life, do you feel more or less able to control your feelings? [affective] [If selected "Much less / a little less"] In a sentence or two, please describe a time when AI made you feel less able to control your feelings. [If selected "Much more / a little more / neither more nor less"] In a sentence or two, please describe a time when AI made you feel more able to control your feelings.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]
4	As a result of using AI in your life, do you feel more or less able to influence powerful people? [structural] [If selected "Much less / a little less"] In a sentence or two, please describe a time when AI made you feel less able to influence powerful people. [If selected "Much more / a little more" / neither more nor less] In a sentence or two, please describe a time when AI made you feel more able to influence powerful people.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]
5	As a result of using AI in your life, do you feel more or less able to make new friends? [social] [If selected "Much less / a little less"] In a sentence or two, please describe a time when AI made you feel less able to make new friends. [If selected "Much more / a little more / neither more nor less"] In a sentence or two, please describe a time when AI made you feel more able to make new friends.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]

#	Question text	Response options
6	As a result of using AI in your life, do you feel more or less able to learn anything useful or important? [informational/instrumental] [If selected “Much less / a little less”] In a sentence or two, please describe a time when AI made you feel less able to learn anything useful or important. [If selected “Much more / a little more / neither more nor less”] In a sentence or two, please describe a time when AI made you feel more able to learn anything useful or important.	Much less capable A little less capable Neither more nor less capable A little more capable Much more capable [Open text for follow ups]

C.4. Agency with Specific AI [Life]

This section measures participants' level of perceived agency in their usage of AI. These questions are adapted from the Basic Psychological Needs Scale for Technology Use.

#	Question text	Response options
1	When you use [AI product most used], do you feel like you are in control over the results that are created? [If selected “Not at all in control / A little in control”] In a sentence or two, please describe a time when you used [AI product most used], where you felt a lack of control. [If selected “Somewhat in control / Very in control”] In a sentence or two, please describe a time when you used [AI product most used], where you felt very much in control.	Not at all in control A little in control Somewhat in control Very in control [Open text for follow ups]
2	When you use [AI product most used], do you feel more or less able to reach your goals? [If selected “Much less / a little less capable”] In a sentence or two, please describe a time when you used [AI product most used], where you felt less able to reach your goals. [If selected “Much more / a little more able / Neither more nor less able”] In a sentence or two, please describe a time when you used [AI product most used], where you felt more able to reach your goals.	Much less able A little less able Neither more nor less able A little more able Much more able [Open text for follow ups]

#	Question text	Response options
3	When you use [AI product most used], do you feel like you are building a personal bond with the [AI product most used]? [If selected “A little”, “Somewhat”, or “Very much”] In a sentence or two, describe a time when you felt like you were building a personal bond with AI.	Not at all A little Somewhat Very much [Open text for follow up]
4	When you use [AI product most used], do you feel more or less lonely compared to your experience in your everyday life? [If selected “Much more alone / a little more alone”] In a sentence or two, please describe a time when you felt more lonely when using [AI product most used]. [If selected “Much less / a little less alone”] In a sentence or two, please describe a time when you felt less lonely when using [AI product most used].	Much less alone A little less alone Neither more nor less alone A little more alone Much more alone [Open text for follow up]
5	When you use [AI product most used], do you ever feel manipulated? [If selected “A little”, “Somewhat”, or “Very much”] In a sentence or two, describe a time when you felt like you were manipulated by [AI product most used].	Not at all A little Somewhat Very much

C.5. Societal implications

This section measures participants' perception of the impact of AI on society and other people.

#	Question text	Response options
	[Introduction text] When answering the following questions, consider your feelings towards AI applications such as ChatGPT, Meta AI, Gemini, or CoPilot.	
1	Do you believe that AI applications are beneficial?	Not at all useful Not very useful Somewhat useful Very useful Extremely useful
2	Do you believe that AI applications are harmful?	Not at all harmful Not very harmful Somewhat harmful Very harmful Extremely harmful

#	Question text	Response options
3	In a sentence or two, please tell us what you are most excited about regarding AI (for yourself or society at large).	[Open text]
4	In a sentence or two, please tell us your biggest concern regarding AI (for yourself or society at large).	[Open text]

D. Additional Figures

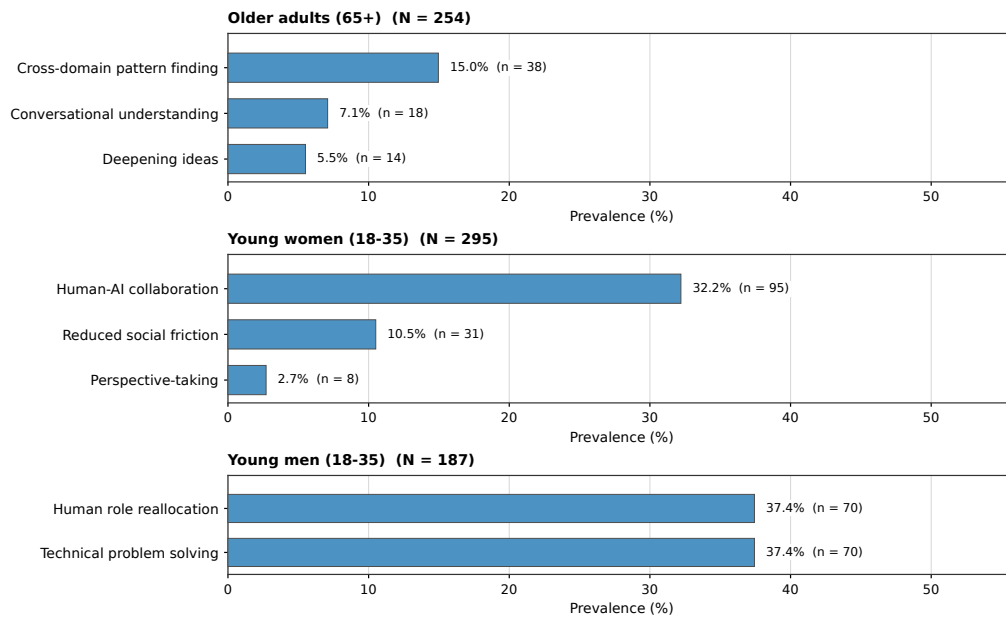


Figure A1: Demographic-specific excitement themes identified by the gap-screening procedure (Appendix A). Bars show the percentage of each subgroup's responses assigned the theme, with counts in parentheses (older adults 65+ N=254; young women 18–35 N=295; young men 18–35 N=187). Themes were retained after researcher review if assigned to at least five responses. “Demographic-specific” means the theme was not captured by the shared codebook, not that it is exclusive to one subgroup.

E. Additional Tables

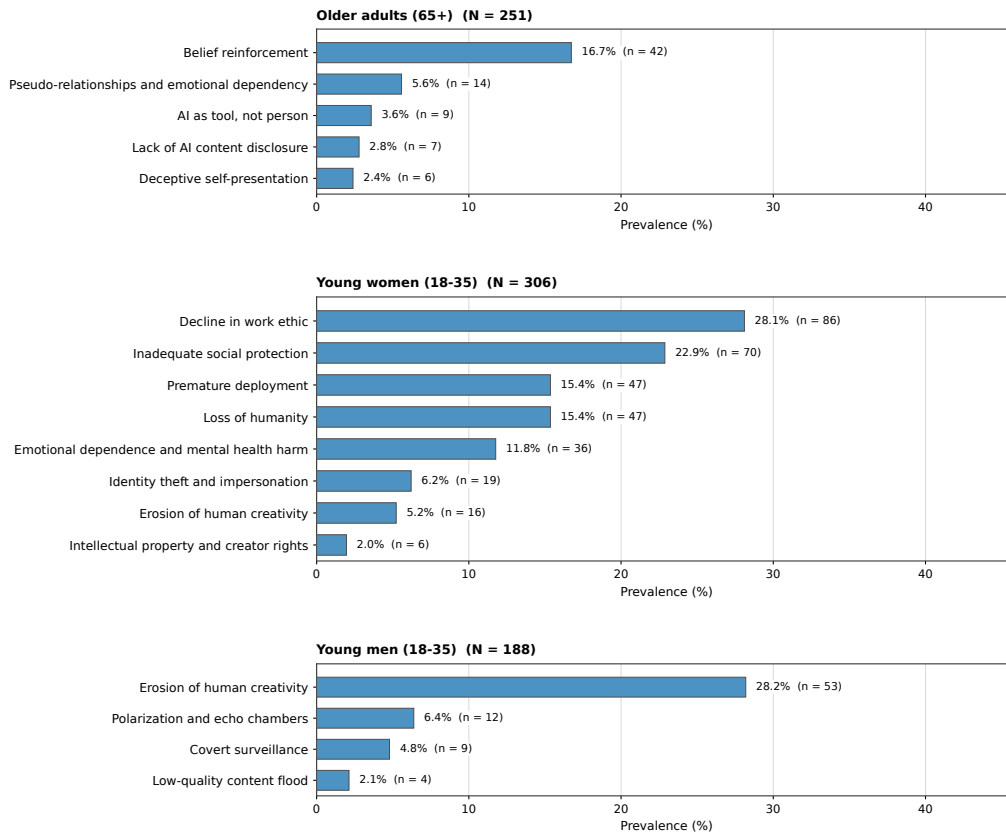


Figure A2: Demographic-specific concern themes identified by the gap-screening procedure (older adults 65+ N=251; young women 18–35 N=306; young men 18–35 N=188). Bars show the percentage of each subgroup's responses assigned the theme, with counts in parentheses. Themes were retained after researcher review if assigned to at least five responses. "Demographic-specific" means the theme was not captured by the shared codebook, not that it is exclusive to one subgroup—erosion of human creativity appears for both young women (5.2%) and young men (28.2%).

Table A1: Weighted Pearson correlations (point-biserial for the binary gender indicator) between each of the six AI-attribution items and gender, age, and education, with design-based standard errors. Holm–Bonferroni correction applied across this family of 18 tests. Three associations survive correction; all others do not and are not interpreted in the main text.

Capability item	Demographic	Weighted Pearson r	p (uncorrected)	p (Holm, 18 tests)	n
Make decisions confidently	Gender (male = 1)	+0.097	1.27e-03	0.0203	2061
Make decisions confidently	Age (years)	+0.018	5.55e-01	1.0000	2058
Make decisions confidently	Education (16-point scale)	+0.098	1.03e-03	0.0175	2061
Reach your goals	Gender (male = 1)	+0.074	1.34e-02	0.2011	2060
Reach your goals	Age (years)	−0.030	3.18e-01	1.0000	2057
Reach your goals	Education (16-point scale)	+0.067	2.60e-02	0.3382	2060
Control your feelings	Gender (male = 1)	−0.021	4.82e-01	1.0000	2060
Control your feelings	Age (years)	−0.061	4.14e-02	0.4552	2057
Control your feelings	Education (16-point scale)	−0.019	5.27e-01	1.0000	2060
Influence powerful people	Gender (male = 1)	−0.001	9.62e-01	1.0000	2059
Influence powerful people	Age (years)	+0.048	1.12e-01	1.0000	2056
Influence powerful people	Education (16-point scale)	+0.068	2.35e-02	0.3294	2059
Make new friends	Gender (male = 1)	−0.006	8.48e-01	1.0000	2059
Make new friends	Age (years)	−0.018	5.46e-01	1.0000	2056
Make new friends	Education (16-point scale)	−0.043	1.50e-01	1.0000	2059
Learn anything useful or important	Gender (male = 1)	+0.125	3.00e-05	0.0005	2059
Learn anything useful or important	Age (years)	+0.066	2.82e-02	0.3390	2056
Learn anything useful or important	Education (16-point scale)	+0.055	6.81e-02	0.6807	2059

Table A2: Kruskal–Wallis omnibus tests of each AI-attribution item across the four race groups large enough to estimate (White $n = 1,414$; Asian $n = 297$; Black $n = 177$; Mixed $n = 131$), with Holm–Bonferroni correction across the six items. Four of the six items differ by race.

Capability item	Kruskal–Wallis H	df	p (uncorrected)	p (Holm, 6 tests)
Make decisions confidently	39.97	3	1.08e-08	< 0.001
Reach your goals	31.32	3	7.27e-07	< 0.001
Control your feelings	20.96	3	1.07e-04	0.0003
Influence powerful people	2.32	3	5.08e-01	0.9666
Make new friends	2.46	3	4.83e-01	0.9666
Learn anything useful or important	26.95	3	6.02e-06	< 0.001

Table A3: Weighted mean differences in reported capability change against White respondents as the reference group, with design-based standard errors and Holm–Bonferroni correction across the 18 contrasts. Two contrasts survive correction, both for Asian respondents. Nine respondents with missing race data were excluded.

Capability item	Group (vs. White)	Weighted mean diff.	SE	p (uncorrected)	p (Holm, 18 tests)	n
Make decisions confidently	Black	−0.038	0.080	6.39e-01	1.0000	177
Make decisions confidently	Asian	+0.203	0.062	1.05e-03	0.0185	297
Make decisions confidently	Mixed	−0.007	0.082	9.29e-01	1.0000	131
Reach your goals	Black	+0.199	0.078	1.05e-02	0.1680	177
Reach your goals	Asian	+0.135	0.062	2.95e-02	0.4424	297
Reach your goals	Mixed	+0.076	0.094	4.18e-01	1.0000	131
Control your feelings	Black	+0.083	0.069	2.31e-01	1.0000	177
Control your feelings	Asian	+0.076	0.050	1.29e-01	1.0000	297
Control your feelings	Mixed	+0.125	0.066	5.71e-02	0.7995	131
Influence powerful people	Black	+0.012	0.067	8.56e-01	1.0000	177
Influence powerful people	Asian	−0.046	0.051	3.61e-01	1.0000	297
Influence powerful people	Mixed	−0.010	0.079	9.01e-01	1.0000	131
Make new friends	Black	+0.061	0.066	3.56e-01	1.0000	177
Make new friends	Asian	−0.047	0.051	3.55e-01	1.0000	297
Make new friends	Mixed	−0.057	0.053	2.85e-01	1.0000	131
Learn anything useful or important	Black	+0.038	0.085	6.49e-01	1.0000	176
Learn anything useful or important	Asian	+0.217	0.066	1.03e-03	0.0185	297
Learn anything useful or important	Mixed	−0.051	0.093	5.81e-01	1.0000	131

Table A4: Valid response counts for every survey item used in the analysis. The five product-specific items were asked only of the 1,903 respondents with an identified most-used tool. Weighted estimates additionally exclude 36 respondents whose released weight is zero, so analytic n is up to 36 lower than the valid n shown.

var	item	n_in_file	n_asked	n_valid	pct_valid_of_file
ag001a	baseline agency: trust my own judgement	2126	2126	2103	98.9
ag001b	baseline agency: control my feelings	2126	2126	2101	98.8
ag001c	baseline agency: learn anything useful or important	2126	2126	2102	98.9
ag001d	baseline agency: make new friends	2126	2126	2101	98.8
ag001e	baseline agency: get people in power to listen	2126	2126	2099	98.7
ag001f	baseline agency: handle myself in social gatherings	2126	2126	2103	98.9
ag001g	baseline agency: get help from formal support systems	2126	2126	2101	98.8
ag001h	baseline agency: get over negative emotions	2126	2126	2099	98.7
ag001i	baseline agency: complete the tasks I set out to achieve	2126	2126	2097	98.6
ag001j	baseline agency: break down complex problems	2126	2126	2101	98.8
ag001k	baseline agency: focus attention on a task	2126	2126	2101	98.8
ag001l	baseline agency: make up my own mind	2126	2126	2101	98.8
ag001m	baseline agency: organize my own thinking	2126	2126	2101	98.8
ai001	AI effect: make decisions confidently	2126	2126	2096	98.6
ai002	AI effect: reach your goals	2126	2126	2095	98.5
ai003	AI effect: control your feelings	2126	2126	2095	98.5
ai004	AI effect: influence powerful people	2126	2126	2094	98.5
ai005	AI effect: make new friends	2126	2126	2094	98.5
ai006	AI effect: learn anything useful or important	2126	2126	2094	98.5
as001	most-used product: perceived process control over results	2126	2126	1898	89.3
as002	most-used product: more able to reach goals with this product	2126	2126	1899	89.3
as003	most-used product: building a personal bond	2126	2126	1898	89.3
as004	most-used product: more/less lonely (hi=more alone)	2126	2126	1900	89.4
as005	most-used product: feel manipulated	2126	2126	1898	89.3
sc001	AI beneficial	2126	2126	2090	98.3
sc002	AI harmful	2126	2126	2091	98.4

Table A5: Two-proportion z-tests comparing theme prevalence in each demographic subgroup against all other respondents, with Holm–Bonferroni correction applied within each open-ended question. Uncorrected p -values are shown alongside corrected values; the final column indicates survival of Holm correction; contrasts that do not survive are reported descriptively in the main text and are not interpreted as reliable subgroup differences.

Question	Theme	Demo-graphic group	n in group	% in group	% among all others	p (unc.)	p (Holm)	Survives
Excite-ment	Information access	Older adults 65+	254	29.9	19.3	0.0001	0.0003	Yes
Excite-ment	Task automation	Young men 18–35	187	17.6	11.9	0.0236	0.0236	Yes
Excite-ment	Future transformation	Young men 18–35	187	15.0	7.3	0.0003	0.0005	Yes
Goal achieve-ment	Finding and summarizing information	Older adults 65+	55	27.3	19.9	0.2119	0.5579	No
Goal achieve-ment	Task and workflow management	Older adults 65+	55	9.1	16.0	0.1860	0.5579	No
Goal achieve-ment	Task and workflow management	Young women 18–35	60	26.7	12.8	0.0058	0.0408	Yes
Goal achieve-ment	Organizing ideas and planning	Young women 18–35	60	23.3	12.5	0.0277	0.1386	No
Goal achieve-ment	Adjusting communication tone	Young women 18–35	60	15.0	3.1	0.0001	0.0007	Yes
Goal achieve-ment	Finding and summarizing information	Young men 18–35	46	28.3	19.9	0.1929	0.5579	No
Goal achieve-ment	Solving practical problems	Young men 18–35	46	17.4	8.5	0.0543	0.2174	No
Goal achieve-ment	Learning and skill development	Young men 18–35	46	13.0	4.4	0.0151	0.0907	No
Feeling in control	Use AI to find information	Older adults 65+	60	20.0	13.8	0.2012	0.2773	No
Feeling in control	Own expertise kept me in control	Older adults 65+	60	13.3	7.7	0.1387	0.2773	No
Feeling in control	Guide AI with prompts	Young women 18–35	66	43.9	32.6	0.0693	0.2079	No
Feeling in control	Used AI as a tool for a task	Young men 18–35	52	25.0	12.2	0.0108	0.0432	Yes
Concerns	Reduced thinking and agency	Older adults 65+	251	19.5	25.0	0.0598	0.2088	No
Concerns	Reduced thinking and agency	Young women 18–35	306	27.8	23.6	0.1188	0.2088	No
Concerns	Reduced thinking and agency	Young men 18–35	188	30.3	23.6	0.0418	0.2088	No
Concerns	Manipulation and malicious use	Older adults 65+	251	24.3	18.9	0.0465	0.2088	No
Concerns	Inaccurate or fabricated outputs	Older adults 65+	251	19.1	14.9	0.0878	0.2088	No
Concerns	Job loss and automation	Young men 18–35	188	28.7	14.8	< 0.0001	< 0.0001	Yes
Concerns	Environmental and energy costs	Young women 18–35	306	13.7	4.0	< 0.0001	< 0.0001	Yes

Table A6: Demographic composition of the study sample (UAS 738) compared with the full UAS panel and with the AI-using eligible pool identified in the screening surveys (UAS 607 and UAS 689). The eligible pool comprises panel members who reported having used generative AI applications. Panel figures are from UAS 689 ($n = 10,733$). Weighted columns use each survey's own post-stratification weight; because those weights are study-specific, the weighted full panel matches U.S. Census marginals while the weighted study sample matches the AI-using population.

Characteristic	Category	Full UAS panel (un-weighted)	Full UAS panel (weighted)	AI-using eligible pool (weighted)	Study sample (un-weighted)	Study sample (weighted)
Age	18–34	16.2%	23.2%	35.7%	23.7%	36.2%
	35–49	29.1%	30.1%	36.5%	37.3%	37.2%
	50–64	28.0%	23.8%	18.2%	26.1%	18.3%
	65+	26.8%	22.9%	9.5%	12.8%	8.3%
Gender	Female	60.9%	51.3%	47.1%	54.5%	45.4%
	Male	39.1%	48.7%	52.9%	45.5%	54.6%
Education	HS or less	20.5%	37.7%	18.1%	8.1%	16.9%
	Some college, no BA	33.3%	25.9%	22.1%	20.6%	21.0%
	Bachelor's	26.1%	20.8%	31.3%	36.5%	33.0%
	Graduate degree	20.1%	15.7%	28.4%	34.8%	29.2%
Race	White	73.5%	74.5%	72.2%	69.6%	70.5%
	Black	12.4%	13.5%	9.8%	8.9%	9.9%
	Asian	7.7%	6.1%	10.9%	14.6%	12.9%
	Other/Multiple	6.3%	5.8%	7.1%	6.9%	6.7%
Household income	Under \$25k	18.1%	21.7%	11.3%	8.8%	11.5%
	\$25–49k	18.2%	19.0%	13.6%	10.9%	12.8%
	\$50–99k	29.7%	29.3%	27.4%	25.3%	24.2%
	\$100k+	34.0%	30.0%	47.8%	55.0%	51.4%