WORKING PAPER · NO. 2025-144

Social Dynamics of AI Adoption

Leonardo Bursztyn, Alex Imas, Rafael Jiménez-Durán, Aaron Leonard, and Christopher Roth

NOVEMBER 2025

THE UNIVERSITY OF CHICAGO
BECKER FRIEDMAN INSTITUTE FOR ECONOMICS

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

Social Dynamics of AI Adoption
Leonardo Bursztyn, Alex Imas, Rafael Jiménez-Durán, Aaron Leonard, and Christopher Roth
November 2025
JEL No. D83, D91, I20, O33

ABSTRACT

Anxiety about falling behind can drive people to embrace emerging technologies with uncertain consequences. We study how social forces shape demand for AI-based learning tools early in the education pipeline. In incentivized experiments with parents—key gatekeepers for children's AI adoption—we elicit their demand for unrestricted AI tools for teenagers' education. Parental demand rises with the share of other teenagers using the technology, with social forces increasing willingness to pay for AI by more than 60%. Providing information about potentially adverse effects of unstructured AI use negatively shifts beliefs about the merits of AI, but does not change individual demand. Instead, this information increases parents' preference for banning AI in schools. Follow-up experiments show that social information has little effect on beliefs about AI quality, perceived skill priorities, or support for bans, suggesting that effects operate through social pressure rather than social learning. Our evidence highlights social pressure driving individual technology adoption despite widespread support for restricting its use.

Leonardo Bursztyn
University of Chicago
Department of Economics
and NBER
bursztyn@uchicago.edu

Alex Imas
University of Chicago
Booth School of Business
and NBER
alex.imas@chicagobooth.edu

Rafael Jiménez-Durán
Bocconi University
rafael.jimenez@unibocconi.it

Aaron Leonard
University of Chicago
aaronleonard@uchicago.edu

Christopher Roth
University of Cologne
Department of Economics
cproth89@gmail.com

This project is pre-registered at AsPredicted #220139 (https://aspredicted.org/cggm-p9yp.pdf), #221155 (https://aspredicted.org/4x57-6r46.pdf), #231705 (https://aspredicted.org/qmb6-djsw.pdf), and #253820 (https://aspredicted.org/si5xq8.pdf).

1 Introduction

The fear of falling behind may cause individuals to adopt new technologies that enhance relative outcomes in the short-run, even in the face of potential long-term risks. These social forces might generate rat race dynamics that accelerate the unrestricted use of these new technologies. Consider the case of artificial intelligence in the context of education: while its integration holds immense potential to transform learning environments and immediate academic performance (Acar, 2024; Kazmi, 2024; De Simone et al., 2025), considerable uncertainty remains about its long-term effects on students' cognitive development and human capital outcomes (Bastani et al., 2025; Melumad and Yun, 2025; Kosmyna et al., 2025).

To understand how the trade-off between short-run relative performance gains and perceived long-run risks shapes decision making, we study the motivating factors that drive parents – the key gatekeepers for adolescent and teenage AI adoption – to adopt advanced AI tools for their children's education.[1] In particular, parents' choices raise a critical question: Do parental decisions to adopt educational AI tools for their kids reflect informed judgments about their potential risks for human capital formation, or are they primarily driven by social factors and anxiety about their children falling behind their peers? If adoption is driven mainly by the perceived necessity of "keeping up" with others in the short- term, then decisions to adopt AI in education may be insensitive to its risks, even if evidence about the long-term harms becomes more apparent.

To answer this question, our paper investigates rat race dynamics in the adoption of AI tools through incentivized, pre-registered experiments involving more than 2,000 parents of teenagers from the United States, Canada, and the United Kingdom. We measure parental demand for advanced AI tools by eliciting parents' willingness to pay (WTP) for a three-month subscription to a premium unrestricted plan for one of the leading generative AI models for their teenagers' education. Then, we explore two determinants of AI demand for teenagers. Motivated by literature on social externalities (Dixit, 2003; Schelling, 1978; Bursztyn et al., 2025a), we provide the first causal estimate for how adoption rates among teenagers' peers influence parental demand. Parents report their willingness to pay for AI tools under scenarios where varying proportions (20% to 80%) of other teenagers in their state already use AI tools.[2] Increasing proportions of peers' usage capture the social

[1] The risks associated with AI adoption may be particularly pronounced at earlier stages of the education pipeline, such as among teenagers.

[2] We implement the peer uptake scenario based on the actual rate of AI adoption that we find from our

pressure for their teenagers to avoid falling behind peers (Bottan and Perez-Truglia, 2022; Clark and Oswald, 1996; Cullen and Perez-Truglia, 2022; Luttmer, 2005; Perez-Truglia, 2020; Bursztyn et al., 2025c).

Second, we investigate how beliefs about AI's impact on cognitive skills shape demand for advanced AI tools. Parents are randomly assigned to either a control group, receiving information emphasizing AI's short-term educational benefits, or a treatment group that additionally highlights potential longer-term risks to quantitative reasoning associated with unrestricted reliance on AI systems in education. Notably, this "AI-risk" information is taken from an existing study that documents the risk of unrestricted use, i.e., AI without guardrails (Bastani et al., 2025). To further test the rat race hypothesis, we also analyze treatments where parents are made aware of the potential long-run costs of adoption, but not of the short-run benefits of AI adoption.

Our results show that social externalities are an important driver of AI adoption in education. First, parents' willingness to pay for AI tools increases by more than 60% as the proportion of their children's peers who use it increases from 20% to 80%. Second, information about the potential long-term AI-risk leads to a large negative shift (-0.43 standard deviations) in incentivized beliefs about the effects of AI on cognitive skills, but importantly, this does little to curb demand as we find that parents are still largely affected by teenage peer adoption. Third, consistent with rat race dynamics, this information increases parents' preference for banning AI in education for all students. To provide further evidence, we collect open-ended qualitative data showing that a substantial portion of parents justify allowing their child to use AI due to fears of them falling behind, despite supporting a ban.

We conduct additional experiments to rule out alternative mechanisms. In particular, we find that the preference to ban AI and perceived AI quality do not decrease with the proportion of others' use, which mitigates concerns about social learning as the driver of our findings. We also show that the importance of long-term skills does not significantly change as peer take-up increases. Instead, the perceived importance of quantitative reasoning and critical thinking remains high as AI adoption increases, consistent with parents trading off long-term human capital when acquiring AI for their children's education.

Taken together, these results demonstrate that AI adoption resembles rat race dynamics: peer adoption creates short-run externalities – in particular anxiety about falling behind – that accelerate adoption despite potential long-term drawbacks. More broadly,

survey, ensuring the WTP is incentivized for every social scenario.

our results suggest that when rat race dynamics drive the diffusion of new technologies with uncertain long-term risks, policy interventions that provide information alone will be insufficient. Instead, collective action, such as school-level adoption of structured AI tools or coordinated adoption decisions, may be required to achieve socially optimal outcomes.[3]

2 Framework: The AI Rat Race

We begin by developing a framework that models parents' demand for AI as a function of its short-run and long-run impact on their children. The formal model can be found in Supplementary Information; we present the intuition here. Parents care both about their child's relative academic performance in the short run and about their long-run human capital. When other teenagers adopt unrestricted AI tools, this lowers their own child's short-run relative performance, generating a negative externality. As a result, even parents who believe that unrestricted AI use erodes long-run cognitive skills may still choose to adopt it for their child when other children use it provided that the relative short-run gains are large enough. Further, as more teenagers use the tool, the worse their child's relative position, increasing their demand for the tool.

This strategic environment resembles a prisoner's dilemma: although parents collectively prefer a scenario where no child uses AI at school, individually each parent is incentivized to allow their own child to use AI if they anticipate that others will do the same. Our framework predicts that beliefs about long-run risks shape collective preferences; parents who believe AI harms long-run human capital are more likely to favor school-level bans. However, as long as short-run competition dominates parental concerns, individual demand for AI remains unchanged—even in the face of more pessimistic beliefs about long-run outcomes.

In our empirical analyses, we test two predictions from this model:

(i) Demand for advanced AI tools increases in response to peer take-up.

(ii) Information about long-run risks increases support for collective bans, without changing individual demand.

[3] Bastani et al. (2025) show that structured AI educational tools (GPT Tutor) which facilitate problem solving rather than providing solutions have positive effects while they are being used and little to no negative downstream impacts on learning.

3 Data and Design

3.1 Sample: Parents as AI Gatekeepers

To test our hypotheses, we conduct experiments with parents of teenagers aged 13 to 18—the key decision-makers and gatekeepers of adolescent technology use.[4] Their beliefs and preferences shape both household adoption of AI tools and broader norms around educational technology. The sample was recruited via Prolific, an online research platform, with eligibility limited to parents meeting the age criteria. For our first and main experiment we sampled residents of the United States. Our second and third experiment also include residents of the United Kingdom and Canada. As pre-specified, we only include respondents who correctly complete a comprehension check that is explained in more detail below.

The final sample of our two main experiments consists of 1,992 parents with at least one teenage child aged 13–18. The average age of respondents is 48.2 years (SD = 8.8), with a roughly balanced gender distribution: 53% female and 47% male. Respondents report an average of 1.34 children in the 13–18 age range. Self-reports indicate that 92.3% of parents and 92.7% of teenagers have been exposed to AI tools, suggesting near-universal knowledge about AI across both groups (see Supplementary Information for a detailed breakdown).

3.2 Experimental Design

Overview. Parents report demographic characteristics, beliefs about AI adoption among peers, and their own and their teenager's AI usage. We then measure how willingness to pay (WTP) for unrestricted use of an advanced AI tool (premium plan for advanced AI model) varies in response to (i) randomized exposure to information about both AI's potential educational benefits and potential long-term risks, and (ii) under different teenager usage rates.

Background characteristics. At the outset, we collect demographic information including the respondent's year of birth, gender, race/ethnicity, and the number and age of their children. We restrict the analysis to parents with at least one child aged 13–18. Respondents report both their own and their teenager's exposure to AI tools (e.g., ChatGPT, Claude, Gemini, Llama), as well as their beliefs about AI adoption rates among

[4]For example, ChatGPT requires users to be at least 13 years of age, and parental consent is required for users 13-18 years old (*Is ChatGPT safe for all ages?*, n.d.).

other teenagers.

Pre-treatment beliefs about the effects of using AI tools. We quantitatively measure respondents' beliefs about the impact of unrestricted AI use on cognitive skills and human capital with an incentivized question asking them to guess the treatment effect from a real study on unrestricted AI use in education.[5] Our participants receive the following instructions:

> A field experiment with high school students tested how access to AI (GPT) for learning affected their **quantitative reasoning skills.** Students who had access to GPT throughout the semester later took a test on quantitative reasoning without AI.
>
> **By what percentage** do you think their quantitative reasoning scores increased or decreased compared to students who never used GPT?

To incentivize this elicitation, we truthfully tell the parents that they will receive a monetary bonus if their estimate is close to the correct answer. We keep the instructions about incentives deliberately simple to avoid participant confusion (Danz et al., 2022).

Information treatment. To assess how information about the effects of unrestricted AI on quantitative skills influences beliefs and valuations, we randomly assign participants to one of two information conditions. In the control group, respondents receive information about recent studies that found improvements in writing and language performance when students use AI tools during learning. In particular, respondents are told:

> Recent research has shown that using advanced AI models helps improve performance on writing and language tasks **while it is being used.**

In the treatment group, participants receive the same message, followed by an additional sentence reporting the results of a field experiment in which high school students with semester-long access to unrestricted AI experienced a nearly 20% decline in quantitative reasoning scores when tested without the tool.[6] In particular, respondents are additionally

[5]For the pre-treatment elicitation of beliefs, we use a field experiment conducted with nearly 1,000 Turkish high school students (Bastani et al., 2025). The authors found a 17% reduction in math test scores for students who used unrestricted AI during the study, after access to the tool was removed.

[6]The eventual productivity effects of AI are still highly uncertain and subject to considerable academic debate (Acemoglu, 2024). Our analysis is focused on the potential for rat-race dynamics conditional on beliefs about productivity effects.

told:

> On the other hand, a field experiment with high school students found that access to GPT for learning led to a **nearly 20% decline** in quantitative reasoning scores when students were later tested without AI.

To enhance credibility, parents are informed that declining test scores may reflect students' overreliance on AI tools like GPT, resulting in reduced independent problem-solving, diminished mental effort, less argument structuring practice, and decreased cognitive engagement. The goal of this treatment is to estimate the sensitivity of demand for AI tools to information about their risks when social externalities are potentially at play.

Willingness to pay elicitation. To measure parents' valuation of AI access for their children, we implement an incentive-compatible mechanism to elicit willingness to pay (WTP) for a three-month premium AI subscription, worth $60 in total. The incentive compatibility of the mechanism mitigates concerns about social desirability bias (Bursztyn et al., 2025b). Specifically, parents are asked for their maximum willingness to pay for a premium unrestricted AI subscription for their child. If their stated willingness to pay exceeds a randomly drawn bonus, they receive the subscription; otherwise, they receive the money. This mechanism ensures that it is optimal for participants to truthfully report their valuation (Becker et al., 1964). To ensure high data quality, parents need to pass a comprehension question on how the mechanism works.

We focus on a premium plan for a leading AI model because it is a real commercial product with a nontrivial price tag, enabling us to measure meaningful trade-offs in parental valuation.[7]

Social scenarios. We assess how parental demand depends on AI adoption of their children's peers. Parents are presented with a series of scenarios in which different proportions of teenagers (20%, 40%, 60%, or 80%) in their state use AI tools. For each scenario, parents report how much they would be willing to pay to obtain the premium AI subscription. To ensure incentive compatibility, parents are truthfully told that we will use the empirically closest scenario to their actual peer environment to determine which of the choices is implemented. Extended Data Fig. 1 shows that respondents attach a high likelihood to each

[7] The premium plan increases usage by a substantial factor over the free plan, which is what the vast majority of participants have access to.

of these scenarios occurring, indicating that they deem them as consequential. While one may be concerned about belief updating, we conduct a related version of this experiment, elaborated on in Section 5 to explicitly rule out social learning contributing to our observed effects.[8]

Post-treatment beliefs about the effects of using AI tools. To quantify the extent to which information moves beliefs about the effects of AI on cognitive skills, we elicit incentivized post-treatment beliefs similarly to pre-treatment beliefs, but using a different underlying study by Essel et al. (2024). In particular, parents see the following instructions:

> A different field experiment with college undergraduates tested how access to AI (GPT) for learning affected their **critical thinking skills.** Students who had access to GPT throughout the semester later took a test on critical thinking without AI.
>
> **By what percentage** do you think their critical thinking scores increased or decreased compared to students who never used GPT?

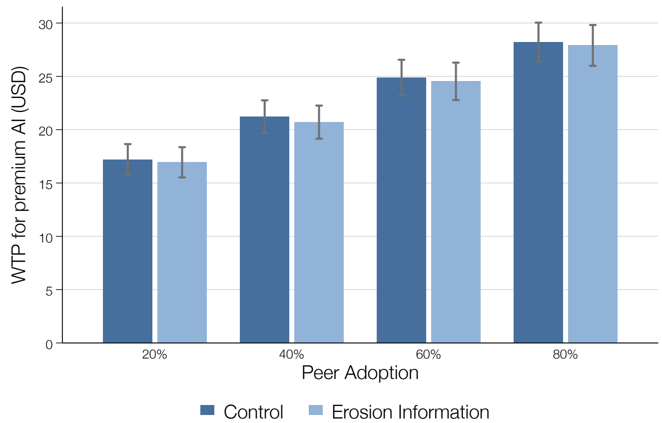
4 Results

Social forces dominate. Figure 1 shows that parents are highly responsive to social context. When asked to state their WTP for premium AI under different scenarios of peer AI adoption (20%, 40%, 60%, or 80%), average WTP increases as take-up rises. In a pooled linear regression across scenarios, we estimate that WTP rises by $1.83 for a 10 percentage point increase in peer usage. When peer take-up increases from 20% to 80%, WTP increases by more than 60%. This effect is both economically large and highly statistically significant ($p < 0.001$). Extended Data Table 1 describes these results in regression format. Further, respondents viewed each of these social scenarios as consequential, as shown in Extended Data Fig. 1.

Information on AI-risk shifts beliefs but not individual demand. We next turn to the effects of information about AI-risk. We first establish that the treatment successfully shifts beliefs about the impact of AI on quantitative reasoning skills. Figure 2 shows that both treatment and control group respondents pre-treatment believe in positive effects

[8]Social learning in this context would refer to parents updating on the quality of AI based on the social scenarios revealing information on the private valuations of others.

Figure 1: WTP by information treatment and peer adoption.



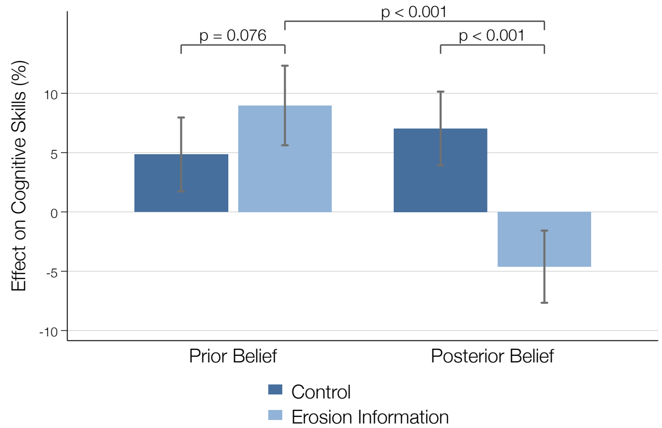
Notes: This figure displays the mean willingness to pay for a three-month subscription for premium AI (worth $60), conditional on peer adoption of AI tools (the percentage of other students who use AI) and the information treatment. The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. Bars in dark blue represent the control condition (n=511); bars in light blue represent the erosion information condition (n=482). Error bars indicate 95% confidence intervals for the group means.

of AI on quantitative skills. Post-treatment, control group respondents also still believe in average positive impacts, while treatment group respondents, on average, think that using unrestricted AI tools *decreases* quantitative skills in the long run. These differences are economically sizable (0.44 standard deviations) and represent a qualitative shift in perceptions about the long-term risk of AI use in education ($p < 0.001$, two-sided t-test). Extended Data Figure 2 displays the distribution of beliefs. These patterns showcase that parents internalize the message about long-term AI-risks.

Despite the qualitative change in beliefs, exposure to information about the risks of AI has almost no effect on demand for AI tools with an effect size close to zero ($p = 0.747$). Moreover, the slope of WTP with respect to peer take-up remains unchanged across treatment arms: we find no evidence that risk information reduces susceptibility to peer adoption.

A preference for banning unrestricted AI in education. To measure preferences regarding AI adoption in educational settings, we ask participants whether they would prefer to live in a world where no students were allowed to use AI tools. Figure 3 shows that a substantial share expresses such a preference—desiring a return to a pre-AI educational era. Strikingly, 57% of parents in the erosion information group prefer a world where no

Figure 2: Average prior and posterior beliefs



Notes: This figure shows and compares the group means of respondents' prior and posterior beliefs, and reports p-values from two-sided t-tests. The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. Bars in dark blue represent the control condition (n=511); bars in light blue represent the erosion information condition (n=482). Error bars indicate 95% confidence intervals.

students use AI, compared to 44% of parents in the control group. This difference is sizable and statistically significant ($p < 0.001$) and underscores the effectiveness of the information treatment in shaping preferences over societal outcomes, highlighting a fundamental conflict between individual and social welfare. In addressing this conflict, our research connects to a long-standing literature on rat races and positional externalities, which shows how competition for relative status can lead to inefficient outcomes (Frank, 1985, 2005; Imas and Madarász, 2023; Bursztyn et al., 2018; Hopkins and Kornienko, 2004).

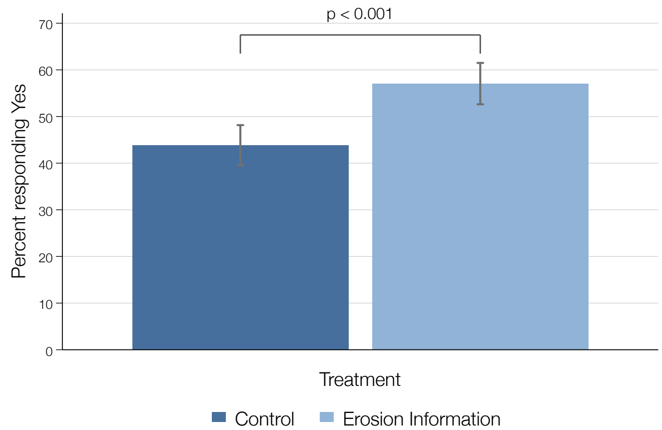
5 Mechanisms

Rat race dynamics. Our data suggest a tension between parental preferences regarding collective AI adoption and the decisions they make for their own children: they allow their children to use AI while also believing that it would be better if nobody used AI in the educational context. We provide three pieces of evidence in support of rat race dynamics as an important driver of these results.

First, we ask parents why they support a ban of AI tools at school even though they allow their child to use them.[9] A systematic content analysis of these responses reveals that

[9]The exact wording of the question was: "You indicated that you would prefer if students were not allowed to use AI. You also indicated that you let your teenager use AI at least sometimes. Why is that?"

Figure 3: Preference for all students learning without AI.



Notes: This figure displays the percentage of respondents who answer Yes to the following question: "Would you prefer if no student was allowed to use AI tools and instead learned the way they used to before AI tools were introduced?" The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. Bars in dark blue represent the control condition (n=511); bars in light blue represent the erosion information condition (n=482). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

approximately 20% of parents mention factors related to social externalities in the form of non-users falling behind, the most frequent category of responses (see Extended Data Fig. 3). Given measurement error in open-ended responses, we view these estimates as lower bounds for the actual importance of these motives.[10] These quantitative estimates, together with qualitative comments such as "I let my teenager use AI because I don't want them to fall behind their peers" and "We let him use AI so he can keep up with his peers," reinforce the interpretation that social externalities in the form of non-users falling behind—rather than purely positive beliefs about AI—drive parental demand. This pattern aligns closely with our experimental finding that willingness to pay increases with higher peer adoption rates. Our work thereby confirms prior work on how perceived social externalities and competitive pressures shape parents' educational choices (Doepke et al., 2019; Bursztyn and Jensen, 2015; Bursztyn et al., 2019).

Second, a large fraction of parents explicitly agree with key aspects of the rat race interpretation when asked. Extended Data Figure 4 shows that a majority of parents in our main experiment (77.6%) agree with the statement that "I think AI tools will help my child now, but I'm worried about the long-term impact using AI tools will have on my child." Notably, agreement is above 70% even for the control group, which was not

[10] For a detailed description of our coding scheme, see Methods.

exposed to AI-risk information. In experiment 3 (Extended Data Fig. 5) we find additional support, including that a majority of parents believe not using AI puts their child at a disadvantage compared with their peers, and would not allow their child to use AI if their peers did not.

Third, we ran a second experiment that is nearly identical to the main design, but where information about the short-run benefits of AI use was not made salient. We use this experiment to establish the robustness of the first study. It replicates the finding that social information strongly shapes WTP for AI learning tools and also confirms that the effects of AI-risk information are relatively muted for demand, though somewhat larger in this experiment. This is consistent with the idea that the anxiety about falling behind is less top-of-mind when the short-run benefits are not mentioned across treatments (see Extended Data Table 2).

Additionally, this second experiment allows us to compare parents' susceptibility to social information when the risks and short-run benefits are either made salient or not. Extended Data Table 3 provides suggestive evidence that parents become more susceptible to information about peer take-up when the short-run information is made salient. This further supports the interpretation that anxiety about their child falling behind in the short-run plays an important role in driving the demand for AI tools.

Alternative mechanism: social learning. An alternative potential explanation for the strong response to peer take-up is social learning: parents might positively revise their beliefs about the benefits of AI tools based on others' adoption.

To explore the social learning channel, we conduct two experiments that measure how parents' preference to ban AI use in schools depends on the fraction of teenagers adopting AI. If the social learning mechanism is quantitatively important, then the positive shift in beliefs should decrease parents' preference to ban AI as more students adopt it.

To test this mechanism, we conducted a third pre-registered experiment. This experiment elicits parents' support for banning AI tools in schools using the same question as in the first two experiments, but it elicits this support in one of two different scenarios of AI take-up among teenagers in the parent's state (20%, or 80%). Extended Data Figure 8 shows that support for the ban does not significantly differ across these two scenarios with a point estimate close to zero ($p = 0.747$). If anything, the point estimate is positive—the opposite of what the social learning mechanism would predict. This response contrasts sharply with the large effects of AI-risk information on the preference to ban, underscoring that support for the ban is an elastic outcome to begin with.[11]

Our fourth pre-registered experiment tests the same hypothesis but using an information treatment design that is more well-powered and mitigates concerns about the hypothetical nature of the previous experiment. In this experiment, we provide an information treatment on the percent of teenagers that use AI in their state based on a sample of responses from previous surveys we conducted. We truthfully randomize this percent to either 40% or 80%, as discussed in Section 6. This approach has the advantage of credibly providing information to respondents on the private valuations of others, which could change their view on the quality and general importance of AI if social learning is an important channel. In Extended Data Figure 7, we show that the information treatment meaningfully shifts participants' posterior beliefs on teenage AI take-up in their state, validating the first stage from our treatment. However, as shown in Extended Data Figure 8, we find no statistically significant difference in either the preference to ban AI or the perceived quality of the AI product, another metric for social learning.[12] Taken together, our evidence suggests that social learning about AI quality does not seem to be the primary mechanism that explains the increase in WTP from Experiments 1 and 2.

Alternative mechanism: future skill requirements. Another candidate explanation is that peer uptake affects how parents prioritize long-term human capital development. In particular, it is possible that as more teenagers use AI, parents may believe that AI literacy is more important to human capital relative to quantitative reasoning and critical thinking skills.

We test this explanation in Experiment 4 by eliciting the perceived importance of various skills for the long-term development of their teenagers. In Extended Data Figure 9, we show that there is not a significant effect on the perceived importance of AI literacy between the 40% and 80% social scenario information treatment. We also do not find a significant effect on the importance of other skills, such as quantitative reasoning, as shown in Extended Data Figure 10.[13] In fact, in terms of levels, critical thinking and

[11] The muted difference in support for a ban across the social scenarios provides suggestive evidence that our main results are not driven by experimenter demand effects. Specifically, it suggests that the substantial differences in WTP documented across social scenarios are unlikely to result from participants responding to perceived expectations of the experimenters (de Quidt et al., 2018).

[12] If social learning was the main mechanism for our results, we would expect that people would update on the quality of the AI product, explaining the WTP increase. We find a small positive coefficient that is not statistically significant.

[13] We only find a significant effect on critical thinking, but in the opposite direction of a future skill

quantitative reasoning are both viewed as highly valued skills by parents. These findings are inconsistent with a mechanism that WTP increases in Experiment 1 and 2 because people update on the importance of AI skills relative to cognitive skills for long-term human capital accumulation. Instead, our results – that WTP increases as peer uptake increases and is inelastic to information about risks – are consistent with the rat race dynamic of prioritizing short-term success over long-term human capital levels.

6 Discussion

Our study reveals how social pressure rather than trust can drive widespread adoption of educational technology. Our findings highlight a critical dilemma inherent in technology adoption: competitive social pressures, rather than cost-benefit tradeoffs, may accelerate the diffusion of emerging technologies before their implications are fully clear. This rat race dynamic, consistent with research on how perceptions of others' behavior guide individual decisions (Tankard and Paluck, 2016), implies that demand will be insensitive to information about long-term risk. Recognizing and mitigating these social dynamics is vital for designing effective policy interventions and educational frameworks, particularly as societies increasingly grapple with powerful yet insufficiently understood technological innovations. Coordinated interventions—such as the combination of tailored and structured AI tools and school-level restrictions—may be necessary to overcome the pressure from social externalities in driving demand for unstructured AI tools. This is particularly important in light of evidence suggesting that *structured* AI environments in the form of digital tutors are quite effective in improving both short- and long-run learning outcomes (Bastani et al., 2025; De Simone et al., 2025).

More broadly, the mechanisms we document mirror the competitive dynamics currently unfolding in the global race to employ general-purpose AI systems. Here too, actors—whether firms, countries, or institutions—face pressure to adopt quickly for fear of falling behind, even as they express deep concern about long-term risks. In the case of firms, rat race dynamics can generate over-investment in AI technology such as data centers due to a similar strategic process outlined in Section 2: conditional on other firms investing, the fear of falling behind will lead the firm to invest even if the individual investment itself will not be profitable. This can generate a potential bubble in the industry.[14] Understanding how

mechanism since we find that critical thinking skills—which unrestricted AI use may erode—are even more valued in the high peer adoption condition.

[14]The prospect of an AI bubble is an active topic of discussion; see, for example, `https://www.wired.`

rat race dynamics interact with systemic technological change is thus not only essential for educational policy, but also for managing broader technological transitions that pose complex trade-offs between short-run gains and long-run consequences.

com/story/ai-bubble-will-burst.

References

Acar, O. A., "With generative AI we can reimagine education—and the sky is the limit," World Economic Forum 2024. Accessed 28 April 2025.

Acemoglu, Daron, "The simple macroeconomics of AI," *Economic Policy*, 08 2024, *40* (121), 13–58.

Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman, "Generative AI without guardrails can harm learning: evidence from high school mathematics," *Proceedings of the National Academy of Sciences*, 2025, *122* (26), e2422633122.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral Science*, 1964, *9* (3), 226–232.

Bottan, Nicolas L. and Ricardo Perez-Truglia, "Choosing your pond: location choices and relative income," *The Review of Economics and Statistics*, 09 2022, *104* (5), 1010–1027.

Bursztyn, Leonardo and Robert Jensen, "How does peer pressure affect educational investments?," *The Quarterly Journal of Economics*, 2015, *130* (3), 1329–1367.

_ , Benjamin R Handel, Rafael Jiménez-Durán, and Christopher Roth, "When product markets become collective traps: the case of social media," *American Economic Review*, 2025.

_ , Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao, "Status goods: experimental evidence from platinum credit cards," *The Quarterly Journal of Economics*, 2018, *133* (3), 1561–1595.

_ , Georgy Egorov, and Robert Jensen, "Cool to be smart or smart to be cool? Understanding peer pressure in education," *The Review of Economic Studies*, 2019, *86* (4), 1487–1526.

_ , Ingar K Haaland, Nicolas Röver, and Christopher Roth, "The social desirability atlas," Preprint at `https://www.nber.org/papers/w33920` 2025.

_ , Rafael Jiménez-Durán, Aaron Leonard, Filip Milojević, and Christopher Roth, "Non-user utility and market power: the case of smartphones," Preprint at `https://www.nber.org/papers/w33642` 2025.

Clark, Andrew E. and Andrew J. Oswald, "Satisfaction and comparison income," *Journal of Public Economics*, 1996, *61* (3), 359–381.

Cullen, Zoë and Ricardo Perez-Truglia, "How much does your boss make? The effects of salary comparisons," *Journal of Political Economy*, 2022, *130* (3), 766–822.

Danz, David, Lise Vesterlund, and Alistair J. Wilson, "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 2022, *112* (9), 2851–2883.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and bounding experimenter demand," *American Economic Review*, November 2018, *108* (11), 3266–3302.

Dixit, Avinash, "Clubs with entrapment," *American Economic Review*, 2003, *93* (5), 1824–1829.

Doepke, Matthias, Giuseppe Sorrenti, and Fabrizio Zilibotti, "The economics of parenting," *Annual Review of Economics*, 2019, *11* (1), 55–84.

Essel, Harry Barton, Dimitrios Vlachopoulos, Albert Benjamin Essuman, and John Opuni Amankwa, "ChatGPT effects on cognitive skills of undergraduate students: receiving instant responses from AI-based conversational large language models (LLMs)," *Computers and Education: Artificial Intelligence*, 2024, *6*, 100198.

Frank, Robert H., *Choosing the Right Pond: Human Behavior and the Quest for Status*, Oxford University Press, 1985.

_ , "Positional externalities cause large and preventable welfare losses," *American Economic Review*, 2005, *95* (2), 137–141.

Hopkins, Ed and Tatiana Kornienko, "Running to keep in the same place: consumer choice as a game of status," *American Economic Review*, 2004, *94* (4), 1085–1107.

Imas, Alex and Kristof Madarász, "Superiority-seeking and the preference for exclusion," *The Review of Economic Studies*, 08 2023, *91* (4), 2347–2386.
Is ChatGPT safe for all ages?

Is ChatGPT safe for all ages?*, OpenAI Help Center. Accessed 19 June 2025.*

Kazmi, Asiya*, "Can AI transform education?," Bill & Melinda Gates Foundation 2024. Accessed 28 April 2025.*

Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes*, "Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task," Preprint at* `https://arxiv.org/abs/2506.08872` *2025.*

Luttmer, Erzo F. P.*, "Neighbors as negatives: relative earnings and well-being," The Quarterly Journal of Economics, 08 2005, 120 (3), 963–1002.*

Melumad, S. and J.-H. Yun, *"Experimental evidence of the effects of large language models versus web search on depth of learning," Preprint at* `https://ssrn.com/abstract=5104064` *2025.*

Perez-Truglia, Ricardo, *"The effects of income transparency on well-being: evidence from a natural experiment," American Economic Review, April 2020, 110 (4), 1019–54.*

Schelling, Thomas C., *Micromotives and Macrobehavior, W. W. Norton & Company, 1978.*

Simone, Martin Elias De, Federico Hernan Tiberti, Maria Rebeca Barron Rodriguez, Federico Alfredo Manolio, Wuraola Mosuro, and Eliot Jolomi Dikoru, *"From chalkboards to chatbots: evaluating the impact of generative AI on learning outcomes in Nigeria," Tech. Rep. WPS11125, The World Bank 2025.*

Tankard, Margaret E. and Elizabeth Levy Paluck, *"Norm perception as a vehicle for social change," Social Issues and Policy Review, 2016, 10 (1), 181–211.*

Methods

The human subjects research described in this article was approved by the University of Chicago Social and Behavioral Sciences Institutional Review Board (IRB25-0487). The methods were carried out in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants.

Samples

Recruitment for all four experiments took place on Prolific. Experiments 1 and 2 were fielded between March 31 and April 8, 2025, with a target sample size of n=1,000 each. The actual sample sizes are 993 and 999, respectively. Experiment 3 was fielded on June 3-4, 2025, with a target and actual sample size of n=200. Experiment 4 was fielded between October 24-28, 2025. We collected 500 responses as well as 174 pilot responses that we include to increase statistical power. Screeners were used to limit the sample to adults with children. Custom in-survey screening was used to additionally limit the sample to parents of at least one teenager (ages 13-18) and to people who correctly answered comprehension checks, when applicable. Our main experiment and Experiment 4 were fielded with U.S. residents only, while Experiments 2 and 3 additionally include residents of the United Kingdom and Canada. For detailed sample descriptives including demographics, see Supplementary Information.

Pre-registration

We pre-registered all four data collections. The pre-registrations include the experimental design, hypotheses, analysis, sample sizes, and exclusion criteria.

- Experiment 1 (main): `https://aspredicted.org/cggm-p9yp.pdf`
- Experiment 2: `https://aspredicted.org/4x57-6r46.pdf`
- Experiment 3: `https://aspredicted.org/qmb6-djsw.pdf`
- Experiment 4: `https://aspredicted.org/si5xq8.pdf`

Details of Experiments 1 and 2

Experiments 1 and 2 are nearly identical to each other. An initial demographics battery is used to screen out respondents who do not have children in the target age group. Next are questions about the respondent's own self-reported use of AI tools, as well as their teenage child's, measured on a five-item Likert scale from "Never" to "Very frequently".

This is followed by the first of two nearly identical belief elicitations. Respondents are presented with a brief description of a field experiment which tested how AI-supported learning affected the test scores of students. To allow independent elicitation of pre-treatment and post-treatment beliefs, two distinct field experiments are referenced: Bastani et al. (2025); Essel et al. (2024). After reading the brief description, respondents are asked to estimate the percentage by which AI use increased or decreased test scores relative to a control group, by moving a slider from -100 to 100. To incentivize higher effort, respondents are truthfully told that they will receive a bonus payment if they come close to the correct answer. Respondents who came within ± 2 percentage points of the correct answer received a bonus payment of $1.

The next part of the survey is composed of an incentive-compatible mechanism designed to elicit parents' willingness-to-pay (WTP) for AI tools for their children. Respondents are informed that one in every 100 participants will be randomly selected to receive either the premium AI subscription (valued at $60) or a monetary bonus. They are told that the subscription is intended for their teenage child's use, e.g. to help with homework, grammar, or research (although we have no way of enforcing this).

Whether they receive the premium AI subscription or a bonus depends on their stated willingness to pay. This is elicited as follows. Each selected participant's outcome is determined via a Becker–DeGroot–Marschak (BDM) procedure (Becker et al., 1964): if their stated WTP exceeds a randomly drawn bonus, they receive the subscription; otherwise, they receive the money. This mechanism ensures that it is optimal for participants to truthfully report their valuation. To ensure high data quality, respondents need to correctly answer a simple comprehension question.[15] As pre-registered, only respondents that pass this comprehension question (49.9%) are allowed to proceed with the experiment.

[15] The question was: "Which of the following statements is **true** about how the premium AI subscription or bonus will be allocated? (i) Entering a higher valuation for the premium AI subscription makes it **less likely** that I will receive the subscription and more likely that I will receive the bonus. (ii) The amount I am willing to forgo for the premium AI subscription **does not affect** whether I receive the subscription or the bonus. (iii) Entering a higher valuation for the premium AI subscription makes it **more likely** that I will receive the subscription and less likely that I will receive the bonus."

To allow elicitation of WTP conditional on AI adoption among other teenagers, willingness to pay is elicited in four different scenarios. Specifically, all respondents are given four answer fields to state their WTP, corresponding to four hypothetical scenarios in which the share of other teenagers using AI is 20%, 40%, 60% or 80%. They are told that their prize will be determined by their stated WTP in the scenario that comes closest to the actual share determined by our survey.[16] [17]

The difference between Experiments 1 and 2 lies in the information conditions (see Extended Data Table 4), which are shown on the same screen as the answer fields for WTP. While the provision of erosion information is randomized between treatment and control in both experiments, Experiment 1 additionally shows information about potential short-run benefits of AI use to both treatment and control groups (active control). Experiment 2 shows no such additional information.

As stating a WTP higher than \$60 is functionally equivalent to stating a WTP of \$60 (besides being arguably nonsensical), we top-code all values above \$60. This is in accordance with our pre-registrations for both experiments.

The remainder of the survey does not differ between the two experiments. Post-treatment beliefs about the cognitive effects of AI use are elicited as described above, again incentivized with a bonus payment. This is followed by a yes-or-no question whether the respondent would prefer students to be banned from using AI. If the answer is yes, *and* if the respondent self-reported earlier that their child uses AI tools at all (i.e., "Never" was not selected), they are prompted to explain the implied contradiction in an open-ended text box. A direct yes-or-no question asking essentially whether the respondent agrees with our hypothesized short-run vs. long-run trade-off is placed at the very end, as it would otherwise be likely to induce experimenter demand effects.

Details of Experiment 3

The purpose of Experiment 3 is to rule out social learning as an alternative mechanism driving the positive relationship between WTP for AI and peer adoption.

[16] We use IP addresses to determine the region of parents, and we define adoption as the share of respondents who indicate that their child uses AI at least "rarely". Several regions appear only once or twice in our data. Still, in practice, all regions we observe at all have empirical adoption rates above 70%, except South Dakota.

[17] Since the company behind the AI model currently offers no way of gifting subscriptions or vouchers, participants who won the subscription ultimately received a payment of \$60 (the cost of the subscription), a link to the AI model's website where they can create a premium account, and instructions to cancel their subscription on time if they do not wish to spend money in excess of their prize.

Experiment 3 starts with a demographics battery to screen out respondents who do not have teenage children. Then all participants receive short-run information and erosion information about the effects of AI use, identical to what the treatment group sees in Experiment 1. Participants are then told about a prospective survey which will estimate the share of teenagers in their country who use AI tools, and are informed that the next question will be hypothetical, based on possible outcomes of that survey.

At this point respondents are randomly assigned into one of two groups: high and low adoption. The high-adoption group receives the prompt

> Imagine the scenario that **80% of teenagers** use AI tools.

while for the low-adoption group, the prompt says "20% of teenagers" instead. Displayed below this prompt is a yes-or-no question whether the respondent would prefer students to be banned from using AI, with the same wording as the other experiments:

> Would you prefer if no student was allowed to use AI tools and instead learned the way they used to before AI tools were introduced?

In this experiment, support for an AI ban is the only outcome of interest. The experiment does not attempt to repeat the willingness-to-pay or belief elicitations from the previous two. We included three Likert-style questions, which were not pre-registered, for exploratory purposes at the end of the survey.

Details of Experiment 4

The purpose of Experiment 4 was to further rule out social learning using an alternative experiment that additionally collected information on perceived quality of AI products and the importance of long-term skills.

Experiment 4 also starts with a demographics battery to screen out respondents who do not have teenage children. Then all participants receive short-run information and erosion information about the effects of AI use, identical to what the treatment group sees in Experiment 1.

Next, we randomize participants into one of two information treatments about the take-up of AI use among teenagers in their state.[18] The high adoption information treatment receives:

[18] We only recruit participants for states in which the information treatment is true based on the responses from our previous survey. We then tell all respondents the share of teenage use of AI across the entire US to ensure all respondents receive information on the full distribution of our survey sample.

> We recently ran a survey on AI use. Among a sample of these respondents in your state, we found that approximately 80% of teenagers use AI tools.

while the low adoption information treatment says 40% instead of 80%. We then ask the same ban preference question as in Experiment 3. We additionally ask respondents to rate their perceived quality of Claude Pro on a Likert scale from 1 to 10, the AI product used in Experiment 1 and 2.

We also ask a set of questions about parents' perceived importance of long-term skills in the future. In particular, for critical thinking, writing and communication, creativity, quantitative reasoning, teamwork, and self-discipline, we ask:

> How important do you think the following skills will be to the success of your teenager over the next 10 years?

on a Likert scale from Not at all important to Extremely important. We also ask on a scale from 1 to 10:

> How important do you think AI literacy is to the long term success of your teenager?

Finally, we include the same final three Likert questions as in Experiment 3. We collected an additional 174 responses as part of a pilot. We opt to pool our data with the pilot responses in order to increase our statistical power. Our primary outcome variables remain not significant at conventional levels without the pilot responses.

Coding scheme for open-ended responses

Experiments 1 and 2 include an open-ended question to provide further insight into the reasons for adopting AI tools in spite of potential reservations. The question is shown to respondents who indicate that their children use AI at least rarely, *and* who also indicate that they would "prefer if no student was allowed to use AI tools and instead learned the way they used to before AI tools were introduced." The text of the question is

> You indicated that you would prefer if students were not allowed to use AI. You also indicated that you let your teenager use AI at least sometimes. Why is that?

We do not impose a minimum or maximum length on the open-text box. Most responses are two sentences or less. We find almost no instances of responses appearing to be LLM-generated.

Rationale The goal of this analysis is to find further evidence for a rat race dynamic: a situation in which consumers find it individually optimal to use AI because of social pressure, but would prefer if nobody used it. A person may demand AI while simultaneously wanting to ban it for everyone if the negative utility of social externalities given non-adoption outweigh the negative utility of adoption.

Specifically, teenagers compete with their peers in school for good grades. In the "steroids" scenario, AI may be harmful in the long run but boosts performance while it is being used. Thus, there is a pressure to use AI to keep up with AI-using peers in the short run: given that many others already use AI to boost their performance, an individual child is better off if they also use AI, despite potential future risks. The framework in the next section formally outlines these predictions.

A parent outlining this type of scenario is classified as citing "social pressure" as a reason for allowing their child to use AI despite preferring a ban. We aim to quantify the prevalence of such reasoning relative to other types of responses in our sample.

Hand-coding of responses We defined seven categories, which are described in the Supplementary Information for reasons of space. We assigned each response to the category that fits best. Categories are not mutually exclusive, but responses are only assigned multiple categories if they fit equally well.

Statistical tests

Peer adoption and WTP This analysis is based on Experiment 1. We compare respondents' willingness to pay for premium AI across four different scenarios of peer AI adoption. Formally, we use ordinary-least-squares (OLS) to estimate the following linear regression model:

$$WTP_{is} = \alpha + \beta X_s + \varepsilon_{is}, \tag{1}$$

where WTP_{is} is respondent i's willingness to pay for premium AI in scenario s, and X_s is the share (in percent) of peers using AI in scenario s. No additional covariates are included. Extended Data Table 1 shows that the results are robust to the inclusion of the treatment indicator and that the slope with respect to peer adoption does not vary across treatment conditions.

The sample consists of 993 respondents. For each respondent we obtain four measurements of WTP, one for each scenario, for a total of 3,972 observations. We use clustered standard errors at the individual level to account for the non-independence of observations of the same respondent. No other corrections or assumption tests were performed.

The estimated coefficient is $\beta = 0.183$ ($SE = 0.009$), 95% CI [0.167, 0.200]. A two-sided t-test with cluster-robust standard errors, using 992 degrees of freedom, yields a test statistic of $t = 21.44$ ($p < 0.001$). Figure 1 visualizes this result (disaggregated across treatment arms).

Information and beliefs This analysis is based on Experiment 1. We use a basic difference-in-differences framework to estimate the effect of erosion information on respondents' beliefs. Formally, we use OLS to estimate the following linear regression model:

$$B_{it} = \alpha + \beta Treat_i + \gamma Post_t + \delta Treat_i \times Post_t + \varepsilon_{it}, \tag{2}$$

where B_{it} is respondent i's belief at time t (pre- or post-treatment) about the effect (in percent) of AI use on students' test scores, $Treat_i$ indicates assignment to treatment or control group, and $Post_t$ indicates the timing of the observation. No additional covariates are included. We are interested in the coefficient δ on the interaction term.

The sample consists of 993 respondents. The control group consists of 511 respondents and the treatment group consists of 482 respondents. For each respondent we measure beliefs twice, once before treatment and once after, for a total of 1,986 observations. We use clustered standard errors at the individual level to account for the non-independence of observations of the same respondent. No other corrections or assumption tests were performed.

The estimated coefficient of interest is $\delta = -15.775$ ($SE = 2.170$), 95% CI [-20.033, -11.517].[19] A two-sided t-test with cluster-robust standard errors, using 992 degrees of freedom, yields a test statistic of $t = -7.27$ ($p < 0.001$). Figure 2 visualizes this result.

Information and WTP This analysis is based on Experiment 1. We compare respondents' willingness to pay for premium AI across the two treatment conditions (pooled across peer adoption scenarios). Formally, we use OLS to estimate the following linear regression model:

$$WTP_{is} = \alpha + \beta Treat_i + \varepsilon_{is}, \tag{3}$$

[19] In the main text of the paper, we report this coefficient in terms of standard deviations of the posterior belief, where $SD = 36.696$.

where WTP_{is} is respondent i's willingness to pay for premium AI in scenario s, and $Treat_i$ indicates assignment to treatment or control group. No additional covariates are included. Extended Data Table 1 presents alternate specifications which control for differences in level and slope across scenarios, with near-identical results.

The sample consists of 993 respondents. The control group consists of 511 respondents and the treatment group consists of 482 respondents. For each respondent we obtain four measurements of WTP, one for each scenario, for a total of 3,972 observations. We use clustered standard errors at the individual level to account for the non-independence of observations of the same respondent. No other corrections or assumption tests were performed.

The estimated coefficient is $\beta = -0.359$ ($SE = 1.114$), 95% CI [-2.546, 1.828]. A two-sided t-test with cluster-robust standard errors, using 992 degrees of freedom, yields a test statistic of $t = -0.32$ ($p = 0.747$). Figure 1 visualizes this result (disaggregated across scenarios).

Information and willingness to ban This analysis is based on Experiment 1. We compare respondents' support for banning AI in education across the two treatment conditions. Formally, we use OLS to estimate the following linear regression model:

$$Y_i = \alpha + \beta Treat_i + \varepsilon_i, \tag{4}$$

where Y_i is a dummy that takes the value 1 if respondent i responds Yes to our ban question. No additional covariates are included.

The sample consists of 993 respondents. The control group consists of 511 respondents and the treatment group consists of 482 respondents. Support for a ban is measured only once per respondent, so it is not necessary to cluster standard errors. No other corrections or assumption tests were performed.

The estimated coefficient is $\beta = 0.132$ ($SE = 0.032$), 95% CI [0.070, 0.194]. A two-sided t-test with 991 degrees of freedom yields a test statistic of $t = 4.20$ ($p < 0.001$). Figure 3 visualizes this result.

The p-value in Extended Data Figure 4 is computed analogously. The estimated coefficient is $\beta = 0.043$ ($SE = 0.026$), 95% CI [-0.009, 0.095] and the t-statistic is $t = 1.64$ ($p = 0.101$).

Social learning in Experiment 3 This analysis is based on Experiment 3. Recall that this experiment randomly assigns respondents to one of two hypothetical peer adoption scenarios (20% and 80%), holding information constant. We compare respondents' support for banning AI across the two treatment conditions (i.e., across scenarios). Formally, we use OLS to estimate the following linear regression model:

$$Y_i = \alpha + \beta S_i + \varepsilon_i, \tag{5}$$

where Y_i indicates support for banning AI in education, and S_i is a binary dummy indicating the peer adoption scenario to which respondent i was assigned, where $S_i = 1$ corresponds to the 80% scenario.

The sample consists of 200 respondents. The 20% adoption group consists of 101 respondents and the 80% adoption group consists of 99 respondents. Support for a ban is measured only once per respondent, so it is not necessary to cluster standard errors. No other corrections or assumption tests were performed.

The estimated coefficient is $\beta = 0.022$ ($SE = 0.069$), 95% CI [-0.114, 0.159]. A two-sided t-test with 198 degrees of freedom yields a test statistic of $t = 0.32$ ($p = 0.747$). Extended Data Figure 8 visualizes this result.

Social learning and skills in Experiment 4 This analysis is based on Experiment 4. Recall that this experiment randomly assigns respondents to one of two information treatments (40% and 80%). We compare respondents' answers for several outcomes variables across the two treatment conditions (i.e., across scenarios). Formally, for each outcome variable, we use OLS to estimate the following linear regression model:

$$Y_i = \alpha + \beta S_i + \varepsilon_i, \tag{6}$$

where Y_i indicates outcome variable, and S_i is a binary dummy indicating the information treatment to which respondent i was assigned, where $S_i = 1$ corresponds to the 80% scenario.

The sample consists of 674 respondents. The 40% adoption group consists of 336 respondents and the 80% adoption group consists of 338 respondents.

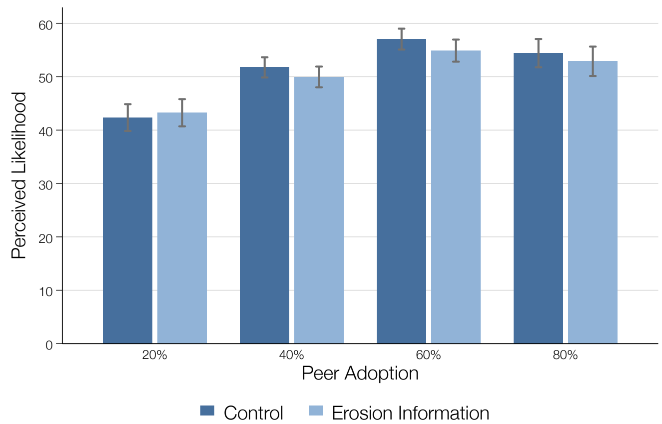
For the perceived quality of Claude Pro, the estimated coefficient is $\beta = 0.180$ ($SE = 0.186$), 95% CI [-0.186, 0.546]. A two-sided t-test with 672 degrees of freedom yields a test statistic of $t = 0.97$ ($p = 0.334$). Extended Data Figure 8 visualizes this result.

For the support for the ban outcome, the estimated coefficient is $\beta = -0.028$ ($SE = 0.037$), 95% CI [-0.1, 0.045]. A two-sided t-test with 672 degrees of freedom yields a test statistic of $t = -0.75$ ($p = 0.453$). Extended Data Figure 8 visualizes this result.

For the posterior beliefs, the estimated coefficient is $\beta = 15.04$ ($SE = 1.30$), 95% CI [12.482, 17.593]. A two-sided t-test with 672 degrees of freedom yields a test statistic of $t = 11.56$ ($p < 0.001$). Extended Data Figure 7 visualizes this result.

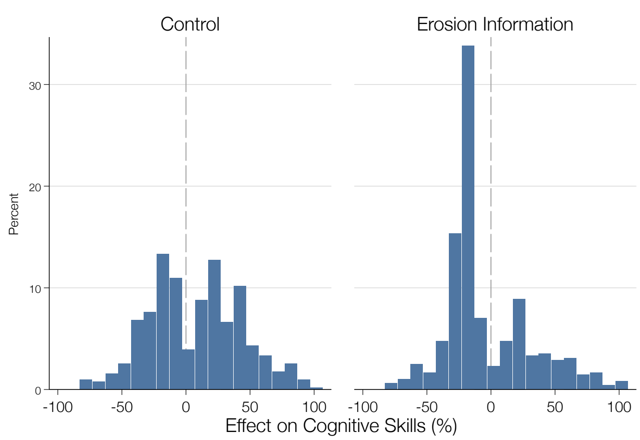
Extended Data

Extended Data Fig. 1: Beliefs about likelihood of different adoption levels.



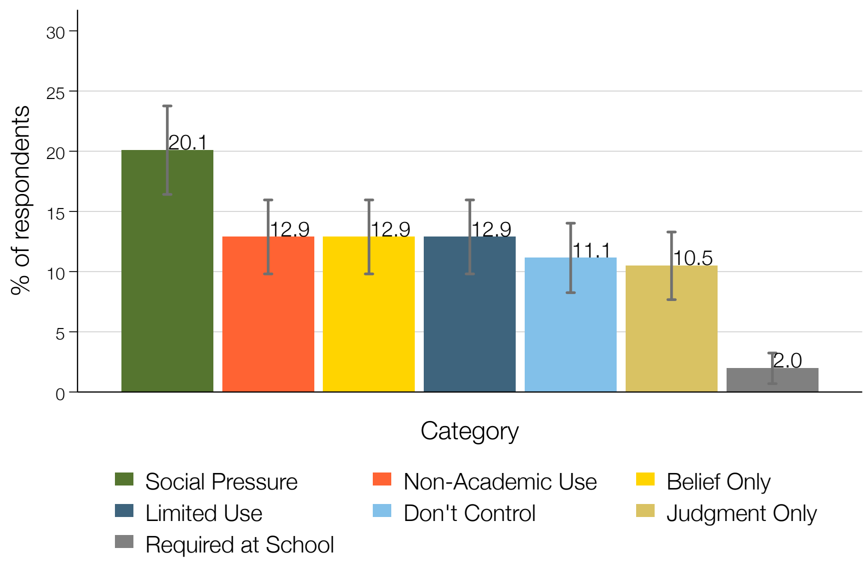
Notes: This figure shows the mean stated likelihood of each social scenario as perceived by respondents. These perceived likelihoods were elicited immediately following the willingness-to-pay elicitation. The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. Bars in dark blue represent the control condition (n=511); bars in light blue represent the erosion information condition (n=482). Error bars indicate 95% confidence intervals for the group means.

Extended Data Fig. 2: Distribution of posterior beliefs.



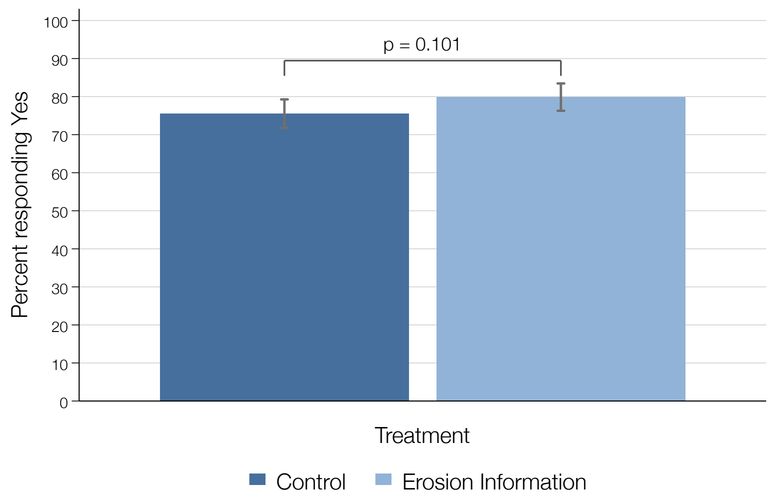
Notes: This figure shows the distribution of posterior beliefs by treatment group. The left panel shows the belief distribution for control group respondents, while the right panel shows the belief distribution for respondents in the erosion information group. The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. The bins of each histogram are 10 percentage points wide. A vertical line is drawn at 0 as a visual reference.

Extended Data Fig. 3: Reasons behind preference for no student being allowed to use AI.



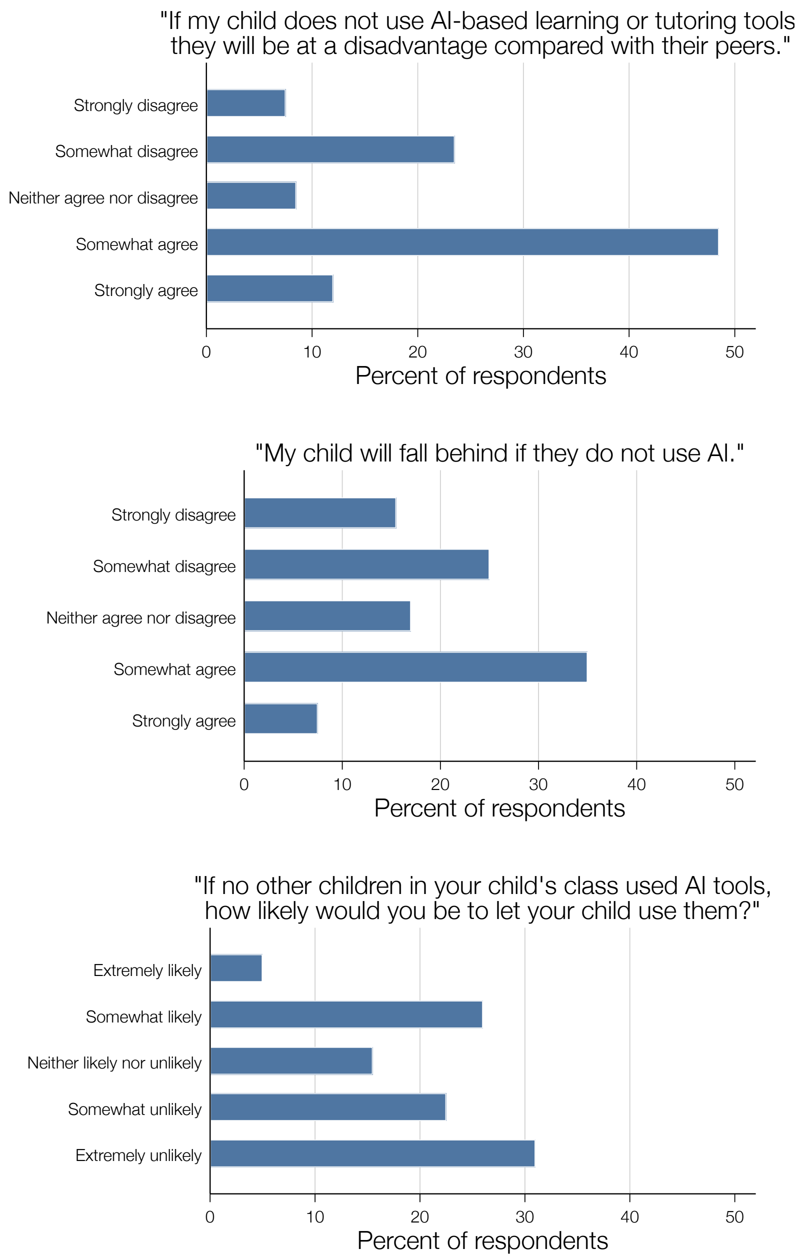
Notes: This figure shows the results of our open-ended analysis for the question "You indicated that you would prefer if students were not allowed to use AI. You also indicated that you let your teenager use AI at least sometimes. Why is that?". The analysis is based on the subset of respondents from Experiment 1 to whom this question was shown because of their previous answers (n=458). Responses were hand-coded by a research assistant. *Category definitions:* Social Pressure: allows use to avoid peer disadvantage; Non-Academic Use: allows only outside schoolwork; Belief Only: belief stated without justification; Limited Use: restricted or supervised access; Don't Control: no parental oversight; Judgment Only: normative view without rationale; Required at School: mandated by school.

Extended Data Fig. 4: Agreement with AI helps now, but worried about long-term impact.



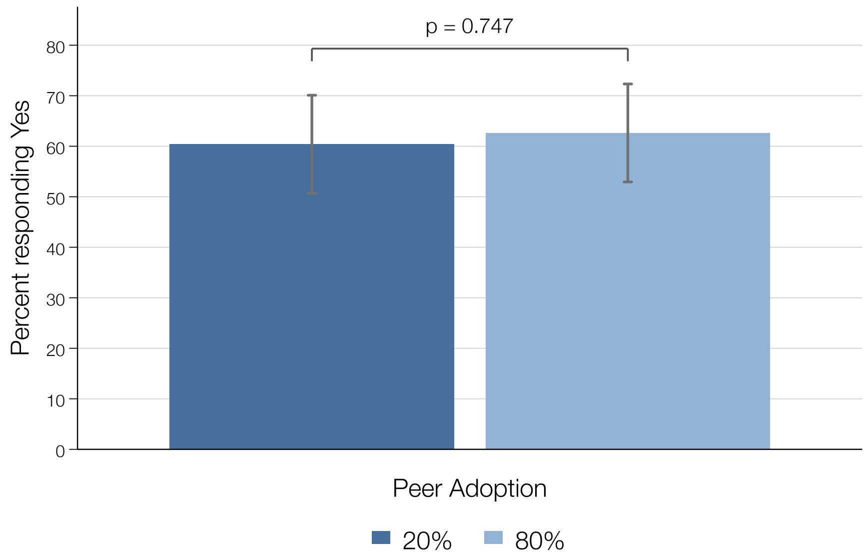
Notes: This figure displays the percentage of respondents who indicate agreeing with the following statement: "I think AI tools will help my child now, but I'm worried about the long-term impact using AI tools will have on my child." The analysis uses data from Experiment 1, as defined in Methods, consisting of 993 respondents. Bars in dark blue represent the control condition (n=511); bars in light blue represent the erosion information condition (n=482). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

Extended Data Fig. 5: More evidence for mechanisms.



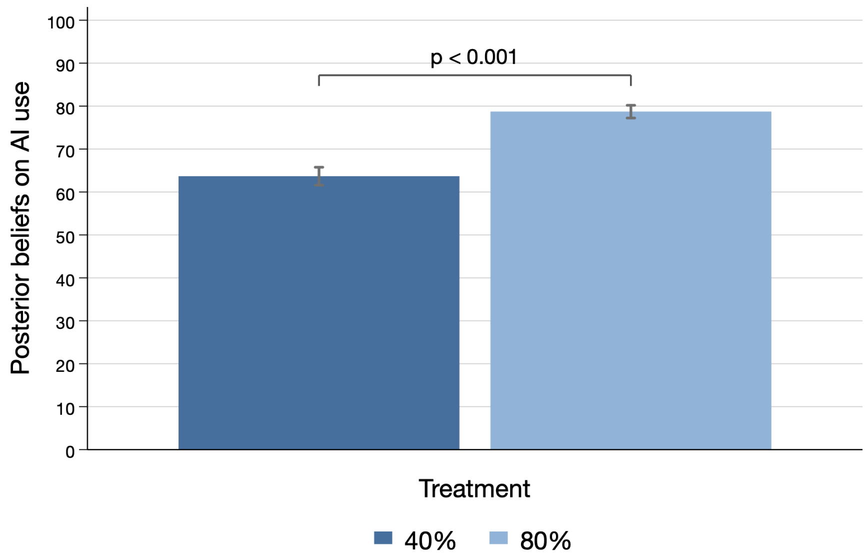
Notes: This figure displays results of three additional questions that were asked at the end of Experiment 3.

Extended Data Fig. 6: No evidence for social learning.



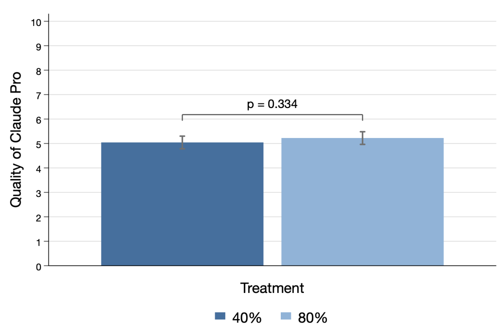
Notes: This figure displays the results of our third pre-registered experiment. After seeing the erosion information, respondents were asked to consider one of two hypothetical scenarios in which adoption of AI among teenagers is randomized between 20% and 80%. We then asked respondents (using language identical to the other experiments) whether they would support an AI ban in this scenario. The figure displays the percentage of respondents answering Yes, i.e. supporting a ban. The analysis uses data from Experiment 3, as defined in Methods, consisting of 200 respondents. Bars in dark blue represent the low-adoption condition (n=101); bars in light blue represent the high-adoption condition (n=99). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

Extended Data Fig. 7: Difference in posterior beliefs.

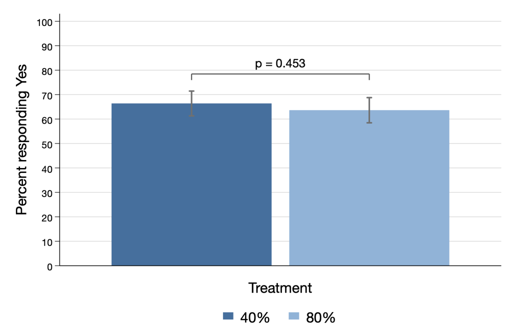


Notes: This figure displays the posterior beliefs regarding AI use from our fourth pre-registered experiment. After seeing the erosion information, respondents were shown one of two information treatments in which adoption of AI among teenagers is either 40% or 80%. At the end of the survey we ask participants about their belief on the percent of teenagers who use AI in their state to validate that our information treatment was effective. The analysis uses data from Experiment 4, as defined in Methods, consisting of 674 respondents. Bars in dark blue represent the low-adoption (40%) condition (n=336); bars in light blue represent the high-adoption (80%) condition (n=338). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

Extended Data Fig. 8: Minimal evidence of quality updating and social learning



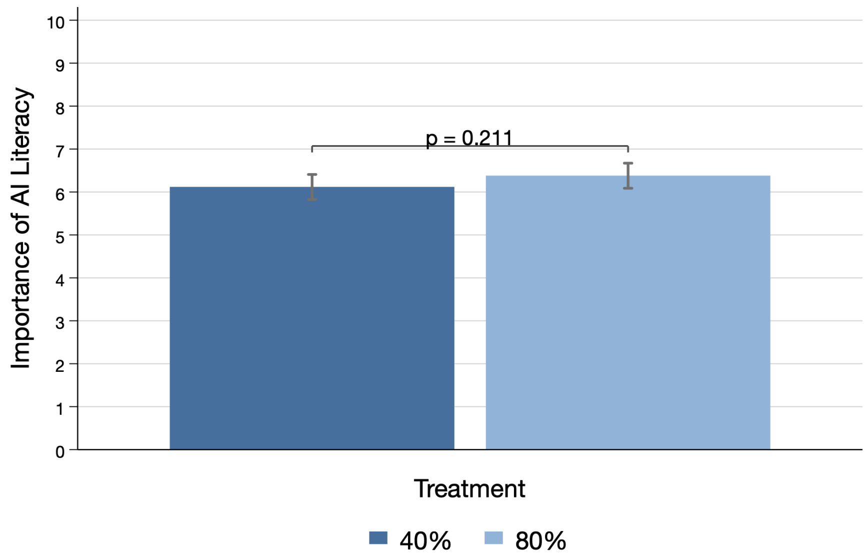
(a) Perceived Quality of Claude Pro



(b) Support for an AI Ban

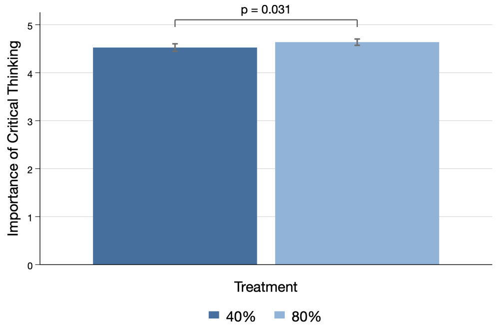
Notes: This figure displays results from of our fourth pre-registered experiment. Figure 8 Panel (a) shows average perceived quality of Claude Pro under two information treatments in which AI adoption among teenagers is either 40% or 80%. Panel (b) shows results for the same question about support for an AI ban from Experiment 3. The analysis uses data from Experiment 4, as defined in *Methods*, consisting of 674 respondents. Bars in dark blue represent the low-adoption (40%) condition ($n = 336$); bars in light blue represent the high-adoption (80%) condition ($n = 338$). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

Extended Data Fig. 9: Minimal evidence of AI skill updating

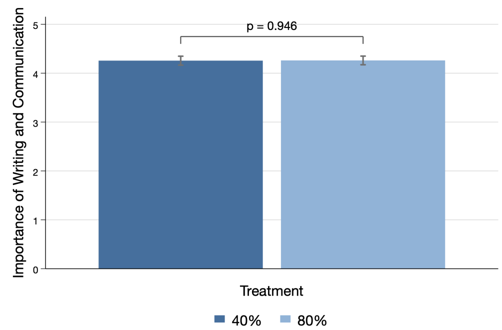


Notes: This figure displays results from of our fourth pre-registered experiment. After seeing the erosion information, respondents were shown one of two information treatments in which adoption of AI among teenagers is either 40% or 80%. We collect the perceived importance of AI literacy. The analysis uses data from Experiment 4, as defined in Methods, consisting of 674 respondents. Bars in dark blue represent the low-adoption (40%) condition (n=336); bars in light blue represent the high-adoption (80%) condition (n=338). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

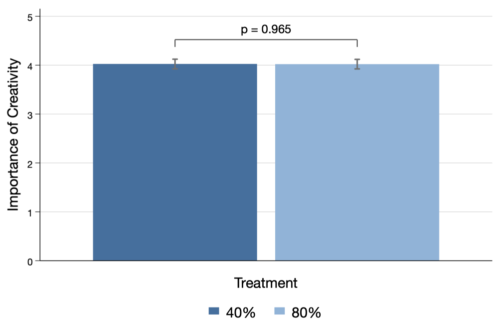
Extended Data Fig. 10: Minimal evidence of general skill updating



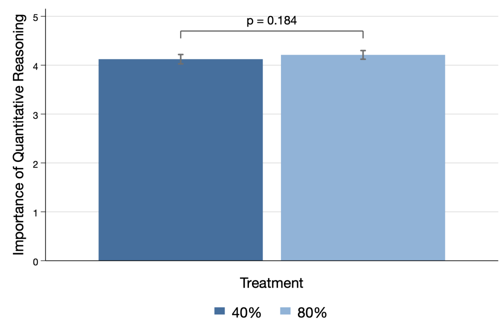
(a) Critical thinking



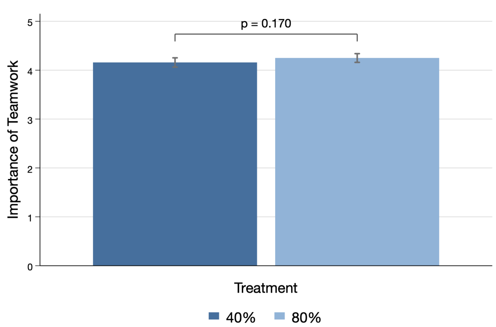
(b) Writing and communication



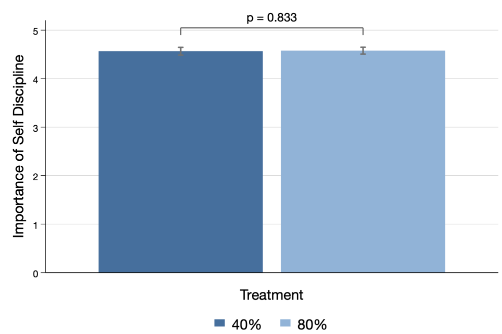
(c) Creativity



(d) Quantitative reasoning



(e) Teamwork



(f) Self discipline

Notes: This figure displays results from of our fourth pre-registered experiment. Panels (a)–(f) display the perceived importance of long-term skills under two information treatments (low vs. high peer adoption). The analysis uses data from Experiment 4, as defined in *Methods*, consisting of 674 respondents. Bars in dark blue represent the low-adoption (40%) condition ($n = 336$); bars in light blue represent the high-adoption (80%) condition ($n = 338$). The p-value is from a two-sided t-test. Error bars indicate 95% confidence intervals for the group means.

Extended Data Table 1: Effect on willingness to pay: main experiment.

	(1)	(2)	(3)
Erosion Information	-0.359	-0.359	-0.367
	(1.114)	(1.114)	(1.146)
Peer Adoption		0.183***	0.183***
		(0.009)	(0.012)
Erosion Information × Peer Adoption			0.000
			(0.017)
Observations	3,972	3,972	3,972
R-squared	0.00	0.05	0.05
Control mean	22.88	22.88	22.88

Extended Data Table 2: Boundary conditions for social effect.

	(1)	(2)	(3)
Erosion information	-4.217***	-4.217***	-0.834
	(1.063)	(1.063)	(1.094)
Peer Adoption		0.178***	0.209***
		(0.009)	(0.012)
Erosion information × Peer Adoption			-0.068***
			(0.017)
Observations	3,996	3,996	3,996
R-squared	0.01	0.06	0.06
Control mean	22.49	22.49	22.49

Extended Data Table 3: Effect of short-run information about benefits of AI.

	(1)
Peer Adoption	0.142***
	(0.012)
Short-Run Info	2.159*
	(1.113)
Short-Run Info × Peer Adoption	0.042**
	(0.017)
Observations	3,812
R-squared	0.05
Sample	Pooled

Extended Data Table 4: Treatment conditions in Experiments 1 and 2.

	Treatment	Control
Exp. 1	Short-Run Info + Erosion Info	Short-Run Info
Exp. 2	Erosion Info	No Info

Supplementary Information

Conceptual Framework

There is a continuum of parents, $\mathcal{I}$. Each parent $i \in \mathcal{I}$ can choose whether their kids use an AI tool or not, which we denote by the indicator variable $x_i \in \{0, 1\}$. Let X denote the aggregate use of AI among students. Each parent has two sources of utility: a short-run component that depends on the kid's current academic performance and a long-run component that depends on their future human capital. This framework is intentionally stylized; its purpose is to illustrate the interaction between short-run social externalities and long-run potential risk in driving current demand. We allow both of these components to depend on others' AI consumption decisions, to capture important spillovers in this setting.

Concretely, i's utility is:

$$v_i(x_i, X, \delta) = u_i^{sr}(x_i, X) + \beta u_i^{lr}(h_i - \delta x_i, H - \delta X),$$

where u_i^{sr} is the short-run component and u_i^{lr} is the long-run one. The term h_i denotes individual i's long-run human capital in the absence of AI use and δ represents any depreciation caused by AI use. Likewise, H denotes the aggregate human capital of all students. Note that short-run utility depends on other's AI use, to capture, for example, that students who do not use AI might fall behind as the usage of others is higher. Similarly, long-run utility depends on the human capital of all students, to capture potential complementarities due to knowledge spillovers.

For convenience, define the short-run and long-run gains of using AI (relative to not using it) as $\Delta u_i^{sr}(X) = u_i^{sr}(1, X) - u_i^{sr}(0, X)$ and $\Delta u_i^{lr}(X, \delta) = u_i^{lr}(h_i - \delta, H - \delta X) - u_i^{lr}(h_i, H - \delta X)$, respectively. We make the following assumptions about these utilities:

(i) Parents believe there is a short-run benefit of allowing their kids to use AI. That is, u_i^{sr} is increasing in their kids' AI use, or there are short run gains from AI, $\Delta u_i^{sr}(X) > 0$.

(ii) In the short run, parents exhibit negative non-user externalities; for example, parents' believe that their kids might fall behind when they do not use AI and others do. That is, $\frac{\partial u_i^{sr}(0,X)}{\partial X} < 0$. Moreover, the negative externalities are worse among non-AI users than among AI users, $\frac{\partial u_i^{sr}(1,X)}{\partial X} > \frac{\partial u_i^{sr}(0,X)}{\partial X}$.

(iii) If AI use depreciates human capital ($\delta > 0$), there is a long-term cost of allowing their kids to use AI. That is, u_i^{lr} is increasing in their kids' human capital, or there are long-run losses from AI $\Delta u_i^{lr}(X, \delta) < 0$.

(iv) There are human capital complementarities between individual and aggregate human capital, $\frac{\partial^2 u_i^{sr}}{\partial h_i \partial H} > 0$.

The parents' willingness to pay to give their kids access to AI is:

$$WTP_i(X, \delta) = v_i(1, X, \delta) - v_i(0, X, \delta) = \Delta u_i^{sr}(X) + \beta \Delta u_i^{lr}(X, \delta).$$

For a given market price of AI p and an aggregate level of AI use X, parents choose to give their kids access to AI if $WTP_i(X, \delta) \geq p$.

We also introduce notation to denote the parents' willingness to ban AI, given by:

$$WTB_i(X, \delta, p) = v_i(0, 0, \delta) - v_i(1, X, \delta) + p.$$

AI rat race We say that parent i is in an *AI rat race* when the parent is willing to pay to give AI to their child ($WTP_i(X, \delta) \geq p$) but would simultaneously prefer AI to not exist ($WTB_i(X, \delta, p) \geq 0$). Both inequalities can be rewritten as bounds on the short-run gains from AI: $LB_i \leq \Delta u_i^{sr}(X) \leq UB_i$, where:

$$LB_i = \underbrace{\beta |\Delta u_i^{lr}(X, \delta)|}_{\text{LR losses}} + p$$

$$UB_i = \underbrace{\beta |\Delta u_i^{lr}(X, \delta)|}_{\text{LR losses}} + p + \underbrace{[u_i^{sr}(0, 0) - u_i^{sr}(0, X)]}_{\text{FOMO avoidance}} + \underbrace{\beta \left[u_i^{lr}(h, H) - u_i^{lr}(h, H - \delta) \right]}_{\text{Human capital externalities}}$$

These expressions say that parent i is in a rat race when the short-run gains from using AI compensate the long-term losses in human capital (and the price of AI). At the same time, the rat race requires these short run gains to be lower than the gains from banning AI, given by the sum of three positive terms: 1) the reduction in long-term losses in human capital, 2) the avoidance of short-run anxiety that their kid falls behind, and 3) the externalities from an increased aggregate human capital.

Positive network effects The parents' AI decisions exhibit positive network effects when their WTP is increasing in X. This occurs when: $\frac{\partial \Delta u_i^{sr}(X)}{\partial X} \geq -\beta \frac{\partial \Delta u_i^{lr}(X, \delta)}{\partial X}$. This expression is equivalent to the following inequality:

$$\frac{\partial u_i^{sr}(1,X)}{\partial X} - \frac{\partial u_i^{sr}(0,X)}{\partial X} \geq -\beta\delta \left| \frac{\partial u_i^{lr}(h_i-\delta, H-\delta X)}{\partial H} - \frac{\partial u_i^{lr}(h_i, H-\delta X)}{\partial H} \right|,$$

which holds when AI depreciates long-run human capital ($\delta > 0$) since the left-hand side is positive due to our assumption (2).

Impact of information about long-run skill decay Lastly, we are interested in understanding the impact of providing information about the long-run skill decay due to AI, which we interpret as an increase in δ (where δ is negative).

First, the impact on the parents' WTP for AI is:

$$\frac{\partial WTP_i(X,\delta)}{\partial \delta} = -\beta \frac{\partial u_i^{lr}(h_i-\delta, H-\delta X)}{\partial h} + \beta X \left| \frac{\partial u_i^{lr}(h_i-\delta, H-\delta X)}{\partial H} - \frac{\partial u_i^{lr}(h_i, H-\delta X)}{\partial H} \right|.$$

Thus, the effect has a sign of ambiguous effect since it is the sum of two components with opposite sign. The first component results from a higher perceived loss in the kids' human capital and the second component results from the fact that the damage from the aggregate loss in human capital is now lower due to the knowledge spillovers. On net, both effects can cancel out, as we observe empirically.

The impact on the parent's willingness to ban AI is:

$$\frac{\partial WTB_i(X,\delta)}{\partial \delta} = \beta \left[\frac{\partial u_i^{lr}(h_i-\delta, H-\delta X)}{\partial h} + X \frac{\partial u_i^{lr}(h_i-\delta, H-\delta X)}{\partial H} \right].$$

This impact is unambiguosuly positive since a higher long-run skill decay decreases both the kids' human capital and aggregate human capital.

Supplementary Tables

Supplementary Table 1: Characteristics of sample for Experiment 1.

	Mean	Std. dev.	Median	Min	Max	Obs.
Panel A: Demographics						
Age	47.88	9.08	47	26	82	993
% Female	52.9	49.9	100	0	100	993
Panel A.1: Race and ethnicity						
% White	77.6	41.7	100	0	100	993
% Black or African American	16.0	36.7	0	0	100	993
% Asian	3.5	18.4	0	0	100	993
% Hispanic or Latino	5.9	23.7	0	0	100	993
Panel A.2: Family						
Number of children 13-18	1.35	0.59	1	1	3	993
Panel B: Parents' AI use						
Panel B.1: Tools						
% using ChatGPT	83.6	37.1	100	0	100	993
% using Gemini	47.2	49.9	0	0	100	993
% using Llama	4.7	21.2	0	0	100	993
% using Claude	16.8	37.4	0	0	100	993
% using other AI tools	24.4	43.0	0	0	100	993
Panel B.2: Overall use						
% using AI tools	94.5	22.9	100	0	100	993
% using AI tools frequently	47.5	50.0	0	0	100	993
Panel C: Teenagers' AI use						
Panel C.1: Tools						
% using ChatGPT	79.9	40.1	100	0	100	993
% using Gemini	39.9	49.0	0	0	100	993
% using Llama	3.1	17.4	0	0	100	993
% using Claude	9.9	29.8	0	0	100	993
% using other AI tools	11.8	32.3	0	0	100	993
Panel C.2: Overall use						
% using AI tools	94.5	22.9	100	0	100	993
% using AI tools frequently	39.0	48.8	0	0	100	993

Supplementary Table 2: Characteristics of sample for Experiment 2.

	Mean	Std. dev.	Median	Min	Max	Obs.
Panel A: Demographics						
Age	48.44	8.44	48	15	80	999
% Female	52.3	50.0	100	0	100	999
Panel A.1: Race and ethnicity						
% White	82.5	38.0	100	0	100	999
% Black or African American	8.7	28.2	0	0	100	999
% Asian	6.2	24.1	0	0	100	999
% Hispanic or Latino	2.5	15.6	0	0	100	999
Panel A.2: Family						
Number of children 13-18	1.32	0.54	1	1	3	999
Panel B: Parents' AI use						
Panel B.1: Tools						
% using ChatGPT	77.0	42.1	100	0	100	999
% using Gemini	36.4	48.1	0	0	100	999
% using Llama	3.0	17.1	0	0	100	999
% using Claude	7.3	26.0	0	0	100	999
% using other AI tools	21.2	40.9	0	0	100	999
Panel B.2: Overall use						
% using AI tools	90.2	29.8	100	0	100	999
% using AI tools frequently	39.8	49.0	0	0	100	999
Panel C: Teenagers' AI use						
Panel C.1: Tools						
% using ChatGPT	70.9	45.5	100	0	100	999
% using Gemini	29.1	45.5	0	0	100	999
% using Llama	3.0	17.1	0	0	100	999
% using Claude	6.3	24.3	0	0	100	999
% using other AI tools	11.3	31.7	0	0	100	999
Panel C.2: Overall use						
% using AI tools	90.9	28.8	100	0	100	999
% using AI tools frequently	34.4	47.5	0	0	100	999

Supplementary Table 3: Characteristics of sample for Experiment 3.

	Mean	Std. dev.	Median	Min	Max	Obs.
Panel A: Demographics						
Age	48.32	9.35	48	22	75	200
% Female	48.5	50.1	0	0	100	200
Panel A.1: Race and ethnicity						
% White	80.5	39.7	100	0	100	200
% Black or African American	11.0	31.4	0	0	100	200
% Asian	4.0	19.6	0	0	100	200
% Hispanic or Latino	5.0	21.8	0	0	100	200
Panel A.2: Family						
Number of children 13-18	1.33	0.56	1	1	3	200

Supplementary Table 4: Characteristics of sample for Experiment 4.

	Mean	Std. dev.	Median	Min	Max	Obs.
Panel A: Demographics						
Age	46.76	8.07	48	26	75	674
% Female	59.8					674
Panel A.1: Race and ethnicity						
% White	77.9					674
% Black or African American	14.2					674
% Asian	4.9					
% Hispanic or Latino	7.4					674
Panel A.2: Family						
Number of children 13-18	1.36	0.57	1	1	3	674

Supplementary Table 5: Overview of coding scheme for open-ended responses.

Category	Definition	Examples
Belief Only	The response is a positive statement (e.g. a belief or assertion about facts), but does not explain why the respondent allows their teenager to use AI.	"I think it improves their thinking skills to not use AI" *or* "AI is helpful but it has also caused students to become lazy."
Social Pressure	The respondent allows AI use only because, regardless of their feelings about AI, these tools exist and others will use them, and thus it is individually optimal for their child to use them, too. The response is consistent with concerns that the existence or widespread use of AI imposes a negative externality on non-users, e.g. being at a disadvantage in school or later on the job market.	"Since AI is so common now, I think my teenager should know how to use it" *or* "They feel that all the other students are using it, so they need to use it to keep up" *or* "Because I know that in this day and age AI tools are socially acceptable and I feel that there is some instances where it's almost required in the same age."
Don't Control	The respondent says they cannot or do not control their child's AI use, due to age, independence, or context.	"She wants to" *or* "It is here and there are areas out of my control that allow its use, like school" *or* "I don't let her use it, she uses it without my knowledge at school."
Judgment Only	The response is a normative statement (e.g. a preference, imperative or value judgment), but does not explain why the respondent allows their teenager to use AI.	"Students need to learn critical thinking skills" *or* "AI can be helpful at times. However, it should not be relied upon excessively for information."
Limited Use	The respondent allows AI use in a restricted way, under supervision, or for specific tasks, with the implication that this is supposed to be less harmful to the child than unfettered AI use.	"I let them use it when we are answering questions but I do not let them use it for actual school work being turned in" *or* "That doesn't mean that they can never use it just not for everything, everyday."
Non-Academic Uses	The respondent allows AI use for non-school tasks such as hobbies, creativity, or fun—but not for academic work.	"She uses it for resumes because she is job hunting" *or* "I do not allow them to use it for school. They use it for creative development outside of school."
Required at School	The respondent allows AI use because it is assigned or required by school, teachers, or curriculum.	"Sometimes his school has a project where they have them us AI on purpose" *or* "She has too for school I personally think it makes them dumber."