An inverted U-shape association between mind health and generative AI-use frequency

Jerzy Bala and Tara C. Thiagarajan

Sapien Labs, Arlington, Virginia, USA

Jerzy Bala https://orcid.org/0009-0003-9655-7136

Tara C. Thiagarajan https://orcid.org/0000-0001-5664-3281

CONTACT Jerzy Bala jerzy@sapienlabs.org Sapien Labs, Arlington, Virginia, USA

Generative AI chatbots such as ChatGPT and Claude have rapidly become part of how we work and live, yet there is little systematic understanding of what they do to the human mind. Here we use a global Internet-enabled sample of 176,198 adults aged 18+, collected in 2025-2026, to examine how frequency and type of generative AI use relate to mind health. Respondents completed a 47-item assessment of emotional, social, cognitive, and physical capacities and problems, aggregated into an overall mind health score, the MHQ, alongside questions on AI use, lifestyle, and life context. Mind health rises with low to moderate GenAI use and then falls sharply with daily use, an inverted U most pronounced among adults aged 18-34. The pattern holds for both cognitive and emotional purposes, although those using GenAI for emotional support or companionship have markedly lower MHQ at every frequency, reflecting higher levels of suicidal thoughts, confused or slow thinking, hallucinations and other problems linked to psychosis. These associations are cross-sectional and cannot be considered causal, but they

persist after adjustment for education, trauma, exercise, nature exposure, and sleep. As GenAI becomes part of daily life, these results suggest a downward trajectory of collective global mind health.

Keywords: generative AI; mental health; ChatGPT; dose-response; AI companionship; young adults

Introduction

Since its public release in late 2022, generative AI has achieved one of the fastest technology adoption curves on record, with ChatGPT alone reportedly reaching approximately 800 million weekly active users by 2025 (Bellan, 2025). These tools produce human-like language responses and images that are growing increasingly indistinguishable from real human responses and photographs or videos of real events. They are now freely available for everyone with an internet connection to use and with large investments made, are aggressively promoted. The range of tasks and queries that generative AI can respond to are vast. These include almost any kind of query and specialized tools are emerging to promote specific uses such as coding and content creation as well as therapy and other kinds of emotional support. Today most companies expect employees to integrate generative AI use as a regular part of their job across a range of tasks.

Thus there is a growing reliance on GenAI for carrying out many of the tasks that humans were once expected to master and navigate on their own. These include not just cognitive tasks but also social and emotional navigation of the world.

A number of studies have recently been published that suggest various negative mental effects of generative AI use. A recent study of 20,847 adults in the United States shows that daily use of AI was associated with 30% higher depressive symptoms (Perlis et al., 2026). Emerging evidence also points to diminished critical thinking and reduced neural engagement with frequent AI use for cognitive tasks. A survey of 319 knowledge workers found that higher confidence in generative AI tools was associated with less critical thinking, and that AI use shifted cognitive effort away from independent problem-solving and toward passive verification and oversight of AI-generated output, a pattern the authors termed a risk of "mechanized convergence" in thought (Lee et al., 2025). An EEG-based study found that those who used ChatGPT to assist in essay writing exhibited weaker brain connectivity compared to search-engine users and unassisted writers, had impaired recall of their own recently-written content, and reduced neural engagement that persisted even after AI assistance was withdrawn (Kosmyna et al., 2025). These findings suggest that the use of generative AI tools may produce broader and more persistent effects on cognitive engagement than earlier tools such as search engines or calculators.

Similarly, concerns have been raised about negative impact of using generative AI tools for emotional support. A joint OpenAI–MIT Media Lab investigation, combining analysis of approximately 40 million real-world ChatGPT interactions with a four-week randomized controlled trial, found that higher daily usage was associated with greater loneliness, emotional dependence, and problematic use, and with reduced offline socialization, effects concentrated among a subset of heavy users with stronger emotional attachment to the chatbot (Fang et al., 2025; OpenAI, 2025). Meanwhile, national survey data indicate that a growing share of young people are turning to AI chatbots specifically for emotional support, largely without disclosing this use to parents or clinicians (McBain et al., 2026), and clinicians increasingly report patients

substituting AI engagement for professional mental health care (American Psychological Association, 2026). At the most severe end of this spectrum, an emerging body of clinical case reports has described a phenomenon informally termed "AI psychosis," in which sustained, emotionally immersive chatbot engagement appears to coincide with the onset or exacerbation of delusional thinking, including in individuals without prior psychiatric history (Morrin et al., 2025; Pierre et al., 2025).

Thus, while rapid adoption of generative AI chatbots for both cognitive tasks and emotional support has already begun to raise concerns about its consequences for human cognition and mental health, there is not a systematic comparative understanding of how AI-use frequency and type of use affect overall mind health and its specific components. Here we define mind health as the range of cognitive, emotional and social capacities that enable us to navigate life's stresses and challenges and function productively. In this study, we use a global sample of 176,198 respondents who have completed a comprehensive assessment of mind health alongside questions on GenAI use frequency and types of use and various aspects of lifestyle and life circumstance to provide a large-scale and in-depth view on the association between GenAI use and various aspects of mind health.

Methods

Data Acquisition

The data was acquired by the Global Mind Project between June 16, 2025, when questions on AI use were added to the assessment, and April 12, 2026 (Newson & Thiagarajan, 2020; Newson et al., 2024). The Global Mind Project acquires data from the Internet-enabled population in 85+ countries using an assessment called the MHQ, which includes 47 items as well as numerous

demographics, life context, and lifestyle questions. These 47 items include 27 mental capacities and 20 problem items (Newson & Thiagarajan, 2020; Newson et al., 2022)

Here we use data from 176,198 people who answered questions on AI use in addition to the MHQ. These are a subset of respondents during this same period (66.7%) as they exclude individuals who indicated that they do not use generative AI chatbots. Respondents were retained if they answered the AI-use frequency question, completed the MHQ assessment, and reported their age band and biological sex (female or male); the 87,887 respondents who selected "I have never used an AI assistant" were excluded. Non-English response options were mapped to their English equivalents, and the top frequency label available only in non-English surveys ("All of the time", 1.9% of users) was pooled with "Constantly".

In addition to the AI questions, numerous other demographic and lifestyle and life circumstance factors were simultaneously queried in the Global Mind Project including age, biological sex, education, lifetime trauma experience, exercise frequency, sleep quality and frequency, ultra-processed food consumption and exposure to nature.

AI Questions

AI questions include the following: 1) how often they use generative AI assistants such as ChatGPT or Claude ("I have never used an AI assistant", "Rarely", "A few times a month", "Several days a week", "Several times a day", "Constantly"); 2) selection of which of 10 cognitive purposes they use AI for (finding quick facts or information; solving problems or making decisions; learning new concepts or skills; analyzing data or doing calculations; writing support; translating languages; summarizing long content; creating content such as images, videos, music or stories; planning trips, schedules, or events; coding or debugging); and 3) which

of three emotional purposes (companionship; emotional support, life coaching or counseling; romantic or intimate conversation) they use AI for.

Lifestyle and Life Circumstance Questions

Lifestyle and life circumstance questions in the Global Mind Project used for this analysis were on frequency of getting a good night's sleep, time spent in nature in the past week, frequency of exercising 30 mins or more, and the experience of traumas in both childhood and adulthood that included abuse and assault, parental neglect, death and illness, financial adversity and various other traumatic experiences. With respect to traumas, here we sum up the number of traumatic experiences reported by the respondent as a lifetime trauma load.

Analysis

Analysis includes metric averages computed for different age-use type groups. We also computed statistical significance using two-sided Welch (unequal-variance) tests on each difference of group means, reporting 95% confidence intervals on the difference, to examine differences in peak MHQ between low or moderate AI-use frequency and constant AI use. The Welch test compares two group means without assuming that the two groups have equal variances or equal sizes (Welch, 1947), which is the appropriate default here because the AI-use frequency groups differ in both. The construction, scoring and calibration of the MHQ, including the score ranges used to classify respondents, are described in the Global Mind Project's published methodology and annual reporting (Newson et al., 2022; Newson et al., 2024; Sapien Labs, 2025).

Because the samples analyzed here are large, almost any difference between groups will reach statistical significance, so a P value on its own says little about whether a difference is large

enough to matter. To confirm that the differences we report are not simply a consequence of the large sample, we repeated each test on random subsamples drawn from each group, 2,000 subsamples at each size. With 200 respondents per group, a size typical of a primary study, the difference in MHQ between the moderate-use plateau and constant use among those 18-34 reaches $P<.05$ in 92% of subsamples and the difference between cognitive-only and emotional use in 99%, so both would be detected by a conventionally sized study (Table S14 in Multimedia Appendix 1). Differences on individual items are smaller and require larger samples. Because 47 item-level tests are reported, we also applied a Bonferroni correction for multiple comparisons, giving a threshold of $P=.001$. Forty-five of the 47 differences between rare and constant use remain significant under that threshold, most of them by several orders of magnitude. The two exceptions are also the two smallest differences on the raw 1-9 scale: sensory sensitivity, which differs by 0.04 points ($P=.048$), and sleep quality, which differs by 0.08 points ($P=.006$). Neither is of a magnitude a respondent would be expected to notice, and we do not treat either as a finding.

In addition, we report Cohen's d for each of the principal comparisons. Cohen's d expresses a difference between two group means in units of the pooled standard deviation of the two groups, so, unlike a P value, it does not grow as the sample grows: it describes how large a difference is rather than how confidently it can be told apart from zero. The original convention treats 0.2 as small, 0.5 as medium and 0.8 as large (Cohen, 1988), but those benchmarks were offered as rules of thumb for controlled experiments and translate poorly to a population survey, where the denominator is the full spread of mind health across an entire population rather than the residual variation within a controlled comparison. The MHQ has a standard deviation of about 71 points in this sample, so a d of 0.3 corresponds to a shift of roughly 21 MHQ points in a group average.

Across the comparisons reported here, d for the difference between the moderate-use plateau and constant use falls steadily with age, from −0.37 among those 18-24 to −0.13 among those 65 and older, with the pooled 18-34 comparison at −0.33; the difference between emotional and cognitive-only use is +0.44, and the largest single value is −0.60, for those reporting romantic or intimate conversation with an AI. Item-level differences are an order of magnitude smaller, as is expected of single items on a 1-9 scale. Cohen's d and the 95% confidence interval for every principal comparison are reported together in Table S13 in Multimedia Appendix 1.

Finally, we fit a multivariate gradient-boosted tree model (LightGBM; Ke et al., 2017) predicting MHQ scores from 22 features jointly: the AI-use frequency, the thirteen types of use queried, demographics (age band, sex, education and survey language) and four lifestyle and life circumstance factors (restful sleep, exercise frequency, time spent in nature and lifetime trauma load). We then used SHAP (SHapley Additive exPlanations; Lundberg & Lee, 2017) to decompose the model's predictions into additive per-feature contributions. These are the same five covariates examined one level at a time in Figure 7, so the model adjusts for all of them jointly. Every feature, with the survey response options behind it and the numeric coding applied, is listed in Table S15 in Multimedia Appendix 1, and the method and detailed results are given there.

These four approaches answer four different questions about the same set of differences. The Welch test asks whether a difference can be told apart from chance. The subsampling asks whether it would still be detected in a study of conventional size. Cohen's d asks how large it is relative to the spread of mind health in the population. The SHAP decomposition asks how much of it remains once the other characteristics of the people at each level of use are accounted for. We report all four because no one of them settles the question on its own.

Results

AI adoption and frequency of use

Across a total of 264,085 global Internet-enabled respondents, 176,198 (66.7%) reported using generative AI chatbots such as ChatGPT and Claude (from here on referred to as GenAI). This is dominated by young adults age 18-34, where 89.0% of Internet-enabled respondents are active users of GenAI compared to only 55.4% of those above 35. 41.3% of 18-34 age users engage with it several times a day or constantly compared to 18.2% among those age 35 or older (Figure 1). Differences between males and females were negligible in both age groups. The full distributions of GenAI-use frequency by age and biological sex are shown in Figure 1.

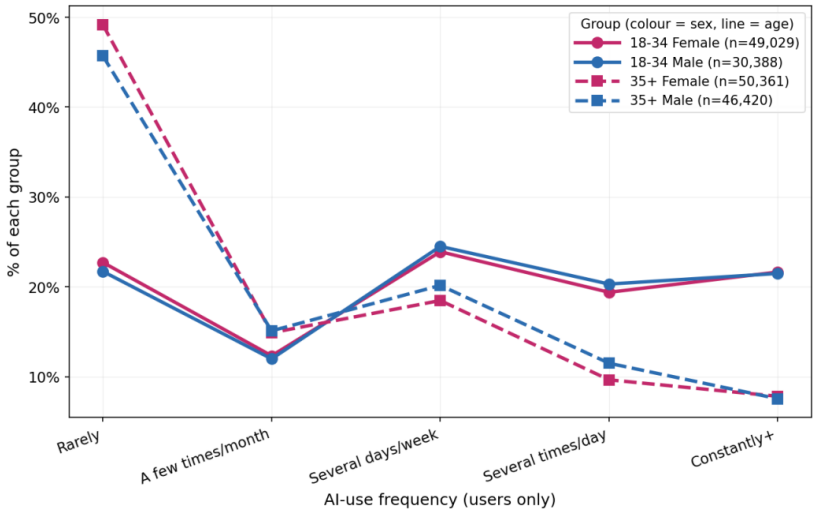


Figure 1. *AI-use frequency by broad age groups and sex shown as the % of each group reporting each frequency of use.*

An inverted-U shape relationship between mind health and AI use frequency

For those 18-34 who are the most aggressive users of GenAI, there is a prominent inverted U-shape relationship between mind health and GenAI use frequency (Figure 2A). Here, MHQ scores are higher among those who use GenAI several days a week compared to those who rarely use GenAI, suggesting mental health benefits, but then sharply decline among those who use GenAI every day to even lower levels. While older age groups above 35 show a similar pattern, it is substantially muted with smaller differences across GenAI use frequencies. The difference in MHQ from lower frequency peak values to constant GenAI use shrinks steadily with age, from 25.1 MHQ points among those 18-24 and 17.4 among those 25-34 to 7.5 among those 65 and older (Welch 95% CIs: 23.4–26.7, 14.5–20.2 and 3.5–11.5 points respectively; all P<.001). Overall, each older age group has better mind health at all AI use frequencies, consistent with a decline over the past two decades that has now been reported in numerous studies (Sapien Labs, 2024, 2025; Newson et al., 2024; Twenge & Blanchflower, 2025; Blanchflower & Bryson, 2025).

Figure 2B shows a similar inverted U-shape for specific dimensions of mind health for the 18-34 age group, where the decline between the several days a week and constant use is sharpest for Cognition and Social Self. The decline is 21.2 MHQ points for Cognition, 21.0 for Social Self and 20.9 for Mood & Outlook, compared with 19.0 for Adaptability & Resilience, 17.6 for Drive & Motivation and 16.4 for Mind-Body Connection (Welch 95% CIs span ±1.3 to ±1.7 points around each dimension estimate; all P<.001).

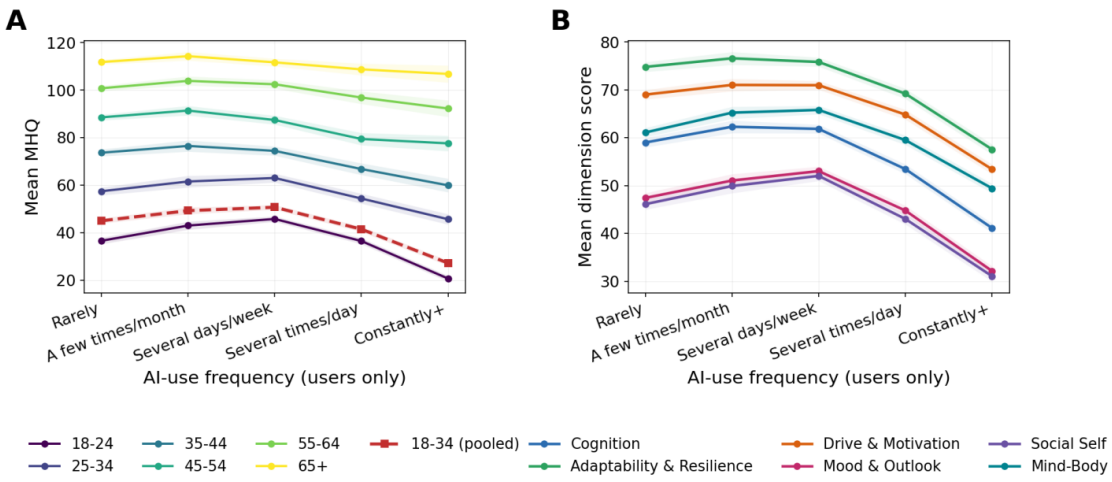


Figure 2. (A) Mean MHQ by AI-use frequency for each age band and (B) for the six MHQ dimensions among those 18-34.

The difference between moderate-use and constant use is in how groups are distributed across the scale, not a uniform shift in MHQ (Figure 3). MHQ is bimodal in this population, with one mode just below zero (those with clinical level mental health problems in the ranges called Distressed and Struggling) and a second near 100 (the expected mean of a normal population). Among those 18-34, relative to the moderate-use group, the constant-use group has more of its members in the Distressed/Struggling range and fewer in the normal range. 32.3% among those using GenAI several days a week are in the Distressed and Struggling groups compared to 46.7% among those using it constantly. The share in the Distressed range alone more than doubles, from 6.1% to 13.1%.

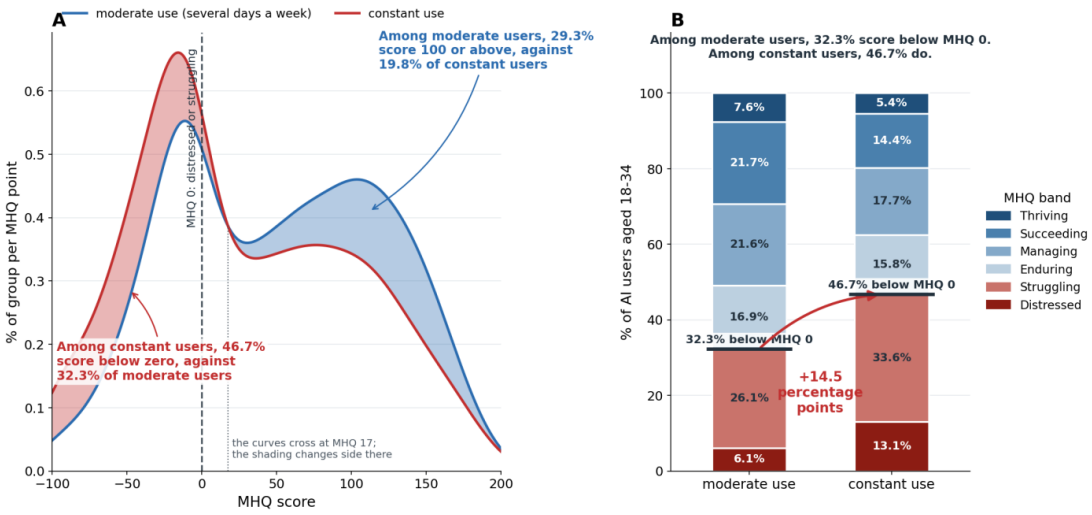


Figure 3*. (A) MHQ distribution for moderate and constant AI use, ages 18-34. Shading marks where each group has the greater share of its members (B) Distributions shown in the six MHQ score ranges (Distressed: MHQ -100 to -50, Struggling: -51-0, Enduring: 1-50, Managing: 51-100, Succeeding: 101-150, Thriving: 151-200.*

Individual capacities and problems with different levels of GenAI-use

At the level of individual capacities queried, ratings of 22 out of 27 capacities were higher with low levels of GenAI use and then decreased, while five did not increase. Focus and concentration stood out from the rest, decreasing prominently with each level of GenAI use (Figure 4A). The sharpest declines between the MHQ at peak and constant daily use were for focus and concentration (-0.58 points on the raw 1-9 scale between the Rarely group and constant use, Cohen's d = −0.24), self-image (-0.43), emotional control (-0.41), emotional stability and

calmness (-0.38) (Welch 95% CIs span approximately ±0.05 points around each estimate; all P<.001). In contrast, sleep quality (+0.13 raw points at its peak) and adaptability to change (+0.09) increased modestly but significantly (sleep quality 95% CI +0.07 to +0.18; adaptability to change +0.04 to +0.15; both P<.001) for most levels of AI use except constant.

Similarly, out of 20 problems, all except one followed the U-shape pattern except for addictions (Figure 4B). The sharpest increases in problems were for addictions and cravings (+0.68 raw points at constant use relative to the Rarely group, Cohen's d = +0.23), confusion and slowed thinking (+0.60), unwanted and obsessive thoughts (+0.58), and sadness and hopelessness (+0.55). The problems that decreased the most for moderate levels of AI use were hallucinations (−0.29 raw points at several days a week of use) and suicidal thoughts (−0.23). All of these item-level changes are statistically significant (Welch 95% CIs span approximately ±0.06 points around each estimate; all P<.001).

Thus overall, while moderate levels of AI use were associated with benefits to several aspects of mind health, it was nonetheless associated with reduced focus and concentration and higher levels of addiction. However daily use was associated with worse mind health on every dimension spanning 41 out of 47 items queried for several times a day and all 47 items for constant daily use.

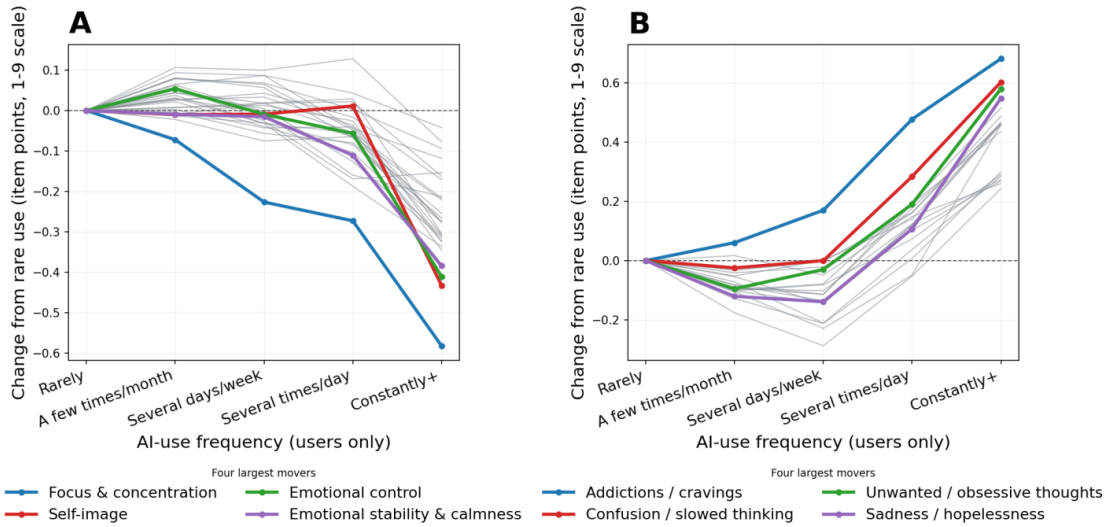


Figure 4. *Item-level dose response: change relative to the Rarely group at each use frequency for the 27 capacities (A) and the 20 problems (B), raw 1-9 scale; problem items are not reverse-coded.*

Type of use and the U-shape dose response

To further explore what aspects of GenAI-use might drive this association we classified GenAI use into two broad categories: cognitive and emotional. Cognitive uses queried included learning new concepts, planning trips and schedules, information seeking, translating, writing code, writing support, creating content, analysis and calculations, summarizing text and problem solving & decision making. Emotional uses included emotional advice and support, companionship and romantic or intimate engagement. Here we focused on the 18-34 age group where the relationship to GenAI-use frequency was most prominent. Almost 100% of respondents used GenAI for cognitive purposes while 36.7% used it for both emotional and cognitive purposes. However, the proportion of those using GenAI for emotional purposes was

significantly higher in the daily and constant categories (Figure 5A), rising from 19.9% of those who use GenAI rarely to 55.1% of those who use it constantly. Thus, one possibility is that the inverted U-shape is due to lower mental health associated with emotional use of GenAI. However, the inverted U-shape remained equally prominent for both types of use although those using it for emotional + cognitive purposes had lower MHQ at all levels of use (26 to 32 MHQ points) than those using it only for cognitive purposes (Figure 5B). Thus the inverted U-shape cannot be attributed to this broad type of use.

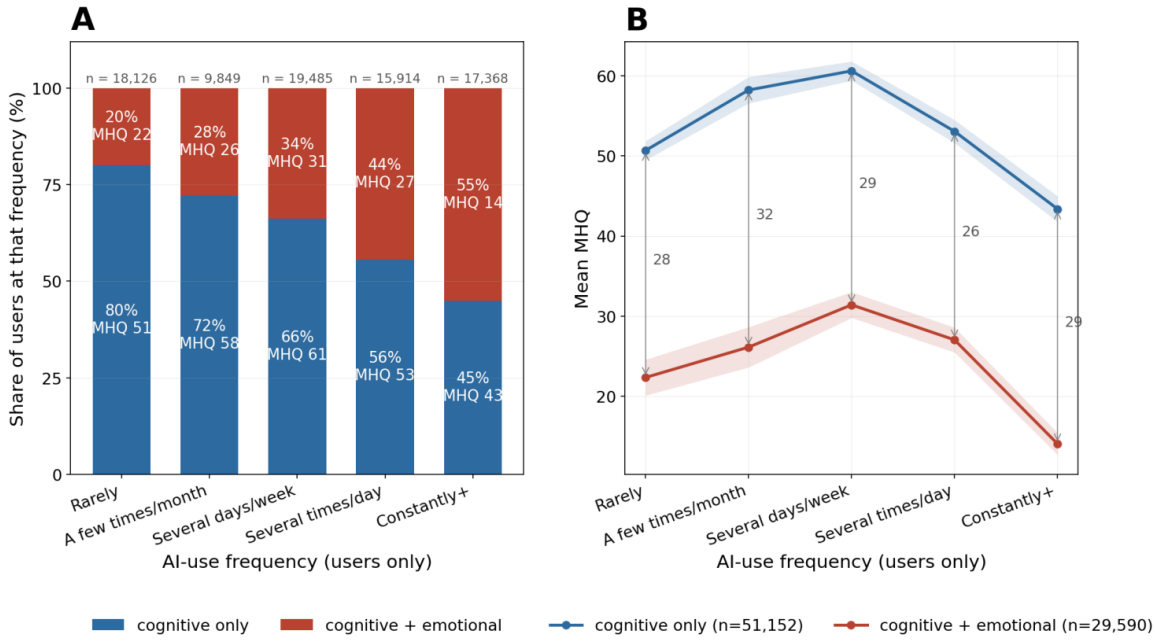


Figure 5*. The share of users also using AI for emotional purposes at each use frequency (A) and mean MHQ by frequency for the two types of use (B).*

Type of use and MHQ

We next compared the MHQ of those who were using it for each of the different types of cognitive and emotional purposes, noting that people typically use it for multiple purposes and that there is therefore an overlap in the groups (Figure 6). Those who were using it for planning, learning new concepts and for information had significantly higher MHQ scores compared to overall average (all ages and all use frequencies pooled; each comparison is the raw mean MHQ of those reporting the use minus that of those not reporting it, n=178,009; this analysis does not require reported sex, so n slightly exceeds the 176,198 of the Methods sample; Figure 6) - 8.6 points above for planning of trips and schedules, 6.9 for learning new concepts and 5.9 for information seeking. Considered separately, the 18-34 age group shows the same ordering but a cleaner split (Figure 6A, colored bars): most cognitive uses sit at or above the 18-34 average; learning new concepts +18.2 MHQ points (Cohen's d = +0.26), planning +10.2 (+0.14), information seeking +7.4 (+0.10), coding +4.0 (+0.06); with only problem solving & decision making clearly below it (−16.7, d = −0.24), while the three emotional uses remain far below (companionship −28.1, d = −0.40; emotional support −27.6, d = −0.39; romantic or intimate engagement −29.7, d = −0.42). Welch 95% confidence intervals put each of these differences to within about ±1 MHQ point for the purposes most respondents report, widening to about ±2 points for those only a small minority report, such as planning and romantic or intimate engagement, and every difference is significant at P<.001. Cohen's d is given after each difference; the emotional uses reach roughly twice the effect size of the largest cognitive one.

In contrast, those who were using it for other cognitive tasks such as creating content, analysis, writing support, coding and problem solving/decision making had MHQ scores below the average. The deficits were 3.7 MHQ points for creating content (Cohen's d = −0.05), 3.9 for writing support (−0.05), 5.0 for analysis and calculations (−0.07), 9.2 for summarizing (−0.13),

13.4 for coding (−0.18) and 24.2 for problem solving and decision making (−0.34). The worst MHQ scores were for those engaging in emotional tasks and support with AI with the lowest scores for those romantically engaged with an AI. Those reporting AI companionship use scored 33.7 MHQ points below the mean (d = −0.47), emotional support 39.5 points below (−0.56) and romantic or intimate engagement 43.0 points below (−0.60). Welch 95% confidence intervals put each of these differences to within about ±0.8 MHQ points for the purposes most respondents report, widening to about ±1.6 points for those only a small minority report, and every difference is significant at P<.001. These are the largest effect sizes reported anywhere in this manuscript, and they are listed with the other principal comparisons in Table S13 in Multimedia Appendix 1.

A further SHAP analysis, which allocates each respondent's predicted MHQ across use frequency, the thirteen purposes, and background factors simultaneously, confirms this (Table S16 in Multimedia Appendix 1), where the emotional-use deficit of 28-30 raw points drops to 11-14 attributed points.

The top 10 items that were significantly worse among those using GenAI for emotional purposes compared to cognitive purposes, ranked by the size of the gap on the raw 1-9 scale, (Figure 6B, see Figure S1 in Multimedia Appendix 1 for the full list and statistics) were: suicidal thoughts (1.11), sadness and hopelessness (0.96), traumatic flashbacks (0.87), confused or slowed thinking (0.87), feeling detached from reality (0.85), repetitive actions (0.82), unwanted obsessive thoughts (0.78), guilt and blame (0.77), hallucinations (0.72) and avoidance and withdrawal (0.71). We note that five of these top 10 items are considered psychosis symptoms.

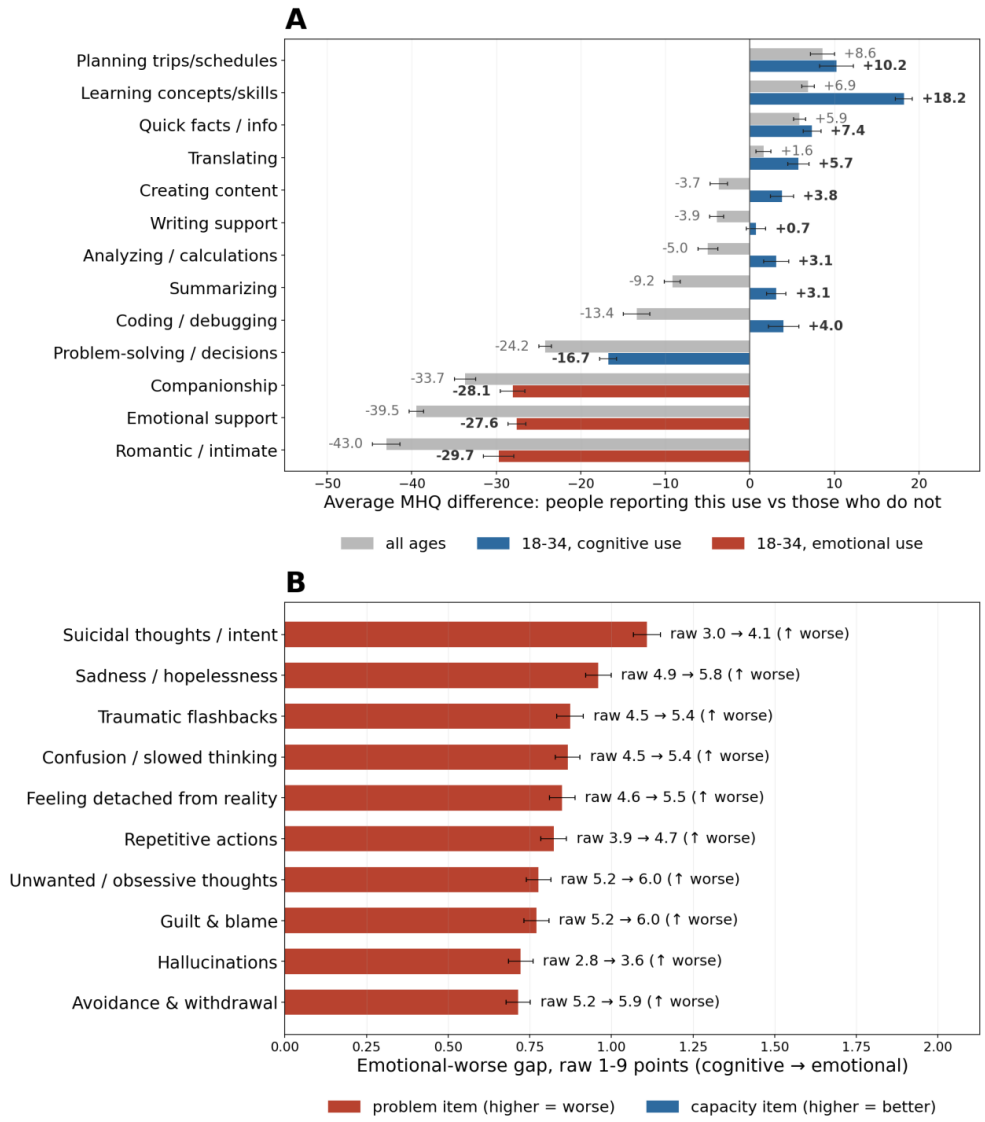


Figure 6. *Mean MHQ difference between those reporting and not reporting each purpose of use (A), and the ten MHQ items with the largest gap between those using AI for emotional purposes and those using it for cognitive purposes only (B). In (A), grey bars are all ages pooled, and colored bars are users 18-34 (blue cognitive, red emotional uses).*

The inverted U-shape dose response cannot be attributed to lifestyle or trauma

We next looked to see if the inverted U-shape dose response could be attributed to lifestyle factors that may be associated with constant AI use, such as education levels, lack of exercise, poor sleep, exposure to nature, or experience of trauma (Figure 7). However, while those with better education, higher levels of exercise, better sleep or less trauma had higher MHQ scores overall the inverted U-shape persisted for all education levels, all exercise levels, sleep levels, time in nature, and for all levels of past traumatic experience (Figure S2 in Multimedia Appendix 1). Overall, the decrease in MHQ points from its peak at moderate AI use frequency to constant use was similar across the majority of lifestyle and life experience conditions (20-25 points in most cases; Figure 7; Table S9 in Multimedia Appendix 1). One surprising aspect was that the decrease was greatest for those who slept well (34.1 points) compared to those who rarely slept well (15.9 points). This was explained by MHQ scores being closer to the floor overall in people who rarely slept well such that the fall was likely an underestimate due to a floor effect. Altogether, the inverted U-shape cannot be explained by an indirect effect of lifestyle factors or experiences that might be impacted by or contribute to excessive time spent online.

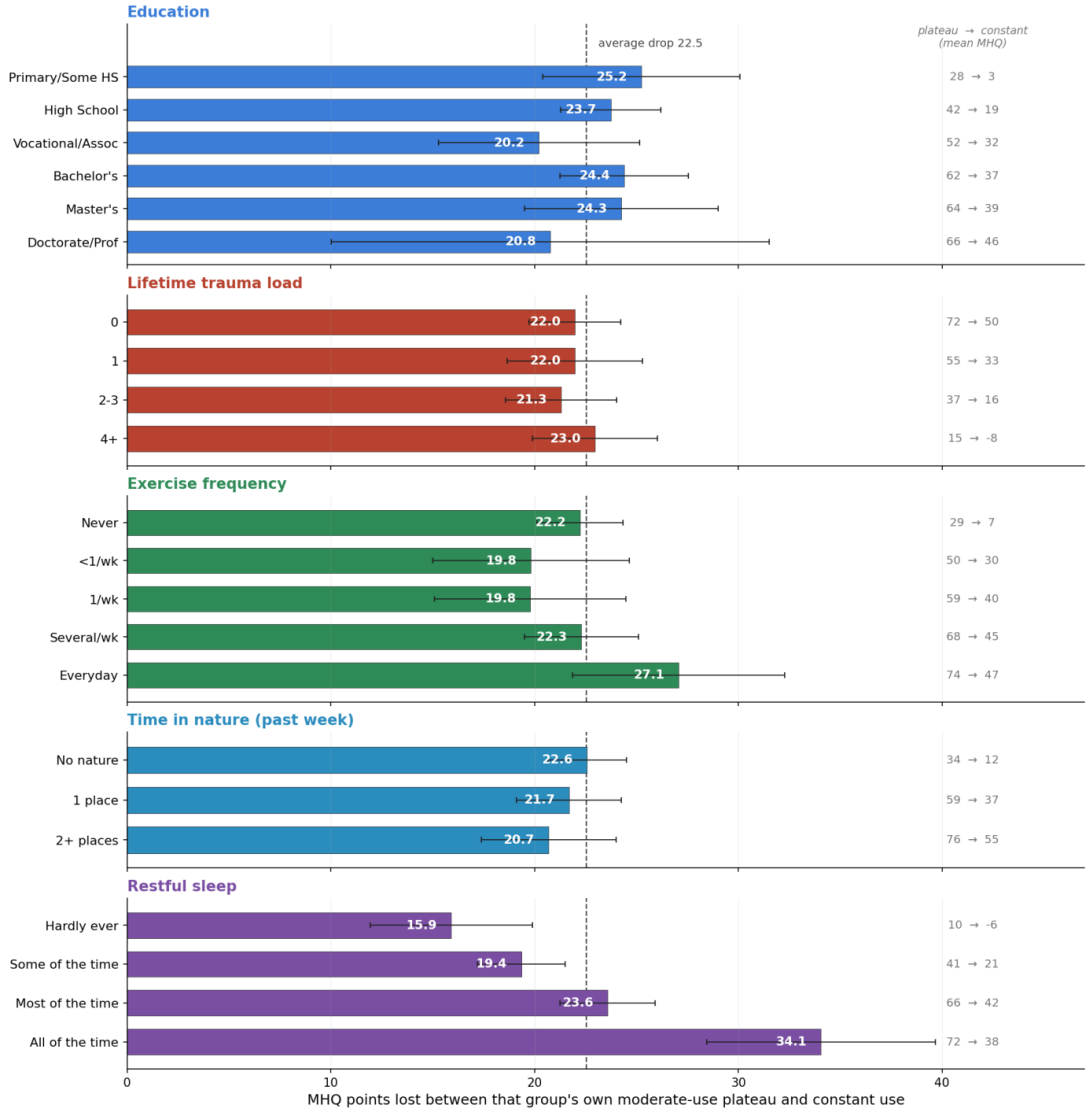


Figure 7. *The drop from each group's own moderate-use plateau to constant use, at every level of education, trauma, exercise, time in nature and sleep.*

The inverted U-shape emerges from a multivariate attribution analysis

As a further check that the dose-response is not a byproduct of the characteristics of who occupies each use-frequency level, we fit a gradient-boosted model predicting MHQ for the 18-34 age group sample from all GenAI-use variables and demographic and life factors jointly (Multimedia Appendix 1) and decomposed its predictions with SHAP. The SHAP contribution of use frequency, the share of each person's predicted MHQ attributable to how often they use GenAI, after the model has assigned credit to trauma, sleep, exercise, time in nature, education, age, sex and purpose of use, traces the same inverted U as the raw group means, peaking at several days a week (+5.3 MHQ points) and falling to −8.7 at constant use (Figure 8). The peak-to-constant swing is 13.9 points against 23.5 in the raw means: about 60% of the descriptive drop is attributed to use frequency itself rather than to co-varying characteristics. Exercise and time in nature are plausibly downstream of heavy use as well as correlated with it, so adjusting for them may remove part of the effect of frequency itself and this figure should be read as a lower bound. The emotional purposes carry the largest negative GenAI-related contributions (−11 to −14 points when reported, against raw deficits of 28-30 points), and the frequency contribution follows the same curve whether or not GenAI is also used emotionally (Figure S5 in Multimedia Appendix 1). The same inverted U emerges when use frequency is supplied to the model as five separate binary indicators carrying no ordering information (Figure S6 in Multimedia Appendix 1), indicating that the shape is recovered from the data rather than imposed by the coding of the frequency variable. As with all analyses in this paper, these attributions describe cross-sectional associations, not causal effects.

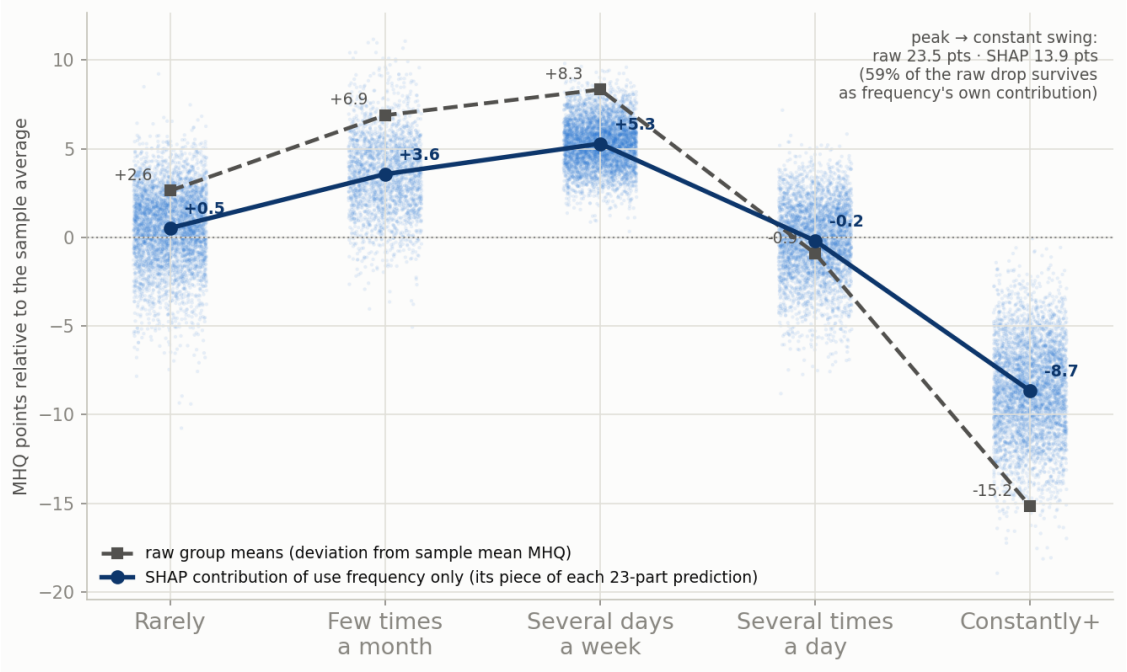


Figure 8. *The inverted U from two constructions: raw mean MHQ per use-frequency level, expressed as deviation from the sample mean (grey, dashed), and the mean SHAP contribution of use frequency from the multivariate model (dark blue line). Recentering only puts the two on a common axis, since SHAP contributions are zero-centered by construction; it does not affect shape or amplitude. Each light-blue dot is one respondent's frequency contribution (horizontal jitter added for visibility); the vertical spread within a level reflects interactions with the other features. 18-34 AI users; SHAP values from the held-out test set.*

Discussion

We show here two key results: First, an inverted U-shape association between higher levels of generative AI use and mind health that is most pronounced in young adults 18-34, arises irrespective of type of GenAI use and cannot be explained by indirect effects of lifestyle or trauma exposure. Second, we show that those using GenAI for emotional purposes have significantly worse mind health at all levels of use compared to those who use it for cognitive purposes only. Finally, this worse mind health is characterized by significantly higher levels of various problems including suicidal thoughts, confused thinking, hallucinations and obsessive thoughts that are associated with psychosis.

Magnitude of difference

Changes in MHQ between moderate and constant generative AI use are on the order of 20 points on the MHQ scale which ranges from -100 to +200. To put these in context, the MHQ has been shown to relate linearly to function such that a 12-point decrease in the MHQ corresponds to a loss of productivity of approximately one day per month (Newson et al., 2024). Thus while 20 points is only a drop down 7% of the scale, the societal implications are substantial. In addition, a doubling of the population with serious mental health distress (from 6.1% to 13% in the Distressed range) is a substantial shift in the mental health burden.

Direction of causation and implications

As this is a cross-sectional study, the direction of causation cannot be unequivocally ascertained from this data. Three possibilities exist: 1) Daily GenAI use degrades mind health; 2) Those who have poor mind health are more likely to use GenAI daily; and 3) Both the above are true, and there is a positive feedback loop between them. Two aspects of the data argue for a degradative effect of GenAI on mind health. First, the U-shape is not dependent on the type of

use. Thus, even though those who use AI constantly are disproportionately those using it for emotional purposes and have poorer mind health overall, this cannot explain the U-shape. Second, GenAI is now a ubiquitous and growing part of daily work, and regular use of GenAI for cognitive purposes is now heavily promoted to working adults and particularly young adults. With over half of young adults using GenAI daily, and the U-shape not dependent on type of use, it is unlikely that only those with characteristics such as poor focus and addictions disproportionately use GenAI more frequently. Second, the inverted U-shape cannot be explained by other indirect possibilities such as exercise, sleep, or time in nature, which may be displaced by spending excessive time online on activities that do not engage GenAI. This displacement explanation, where constant GenAI use lowers mind health by crowding out sleep, exercise or time outdoors, is further ruled out by the fact that the declines are sharper among those who exercise regularly and sleep well.

On the other hand, we cannot rule out all possibilities and favor the third possibility: that causality goes in both directions and effects may be exacerbated by a positive feedback loop. With past trauma and displacement of beneficial activities such as sleep, exercise and nature exposure ruled out as a primary explanation, these results may reflect more direct explanations. For example handing over of tasks to GenAI results in a degradation of one's own capacities; essentially that when you stop using your capacities you lose them, which in turn may create emotional problems.

Age differences in prominence of U-shaped impact of AI-use and mind health

One question raised by this data is why the inverted U-shape is more pronounced among younger generations. While we have no definitive answer, we offer a few possibilities. Several studies have shown that younger generations have increasingly worse mental health, which is linked to

various factors that result in a more vulnerable and less resilient brain. This can both impacts their decision-making online and makes them more susceptible to GenAI exposures. These factors include early smartphone exposure in childhood (Thiagarajan et al., 2025), less in-person peer interaction (Twenge et al., 2019), and less stable families (Sapien Labs, 2023) that may make them more gullible or vulnerable to GenAI suggestions. They also include higher exposure to pesticides and microplastics (Newson et al., 2025), and ultra-processed food (Bala et al., 2025) starting from birth that can have a degradative effect on brain function. For example, UPF consumption is associated with a decreased ability for emotional and cognitive control, which would also impact how younger adults respond to GenAI.

Mind health and AI-use for emotional support

The data showed unequivocally that those who use GenAI for emotional purposes have significantly worse mind health than those who do not, characterized by greater suicidal thoughts, confusion and slowed thinking, hallucinations and obsessive thoughts. This is consistent with a study in Japan (Nakagomi et al., 2026) where 24.7% of 291 users of AI-companion apps had psychological distress compared to 14.5% among non-companion AI users and 9.7% for nonusers. They also reported being more lonely. Thus it is possible that the lower MHQ scores for emotional use of AI in this data is a consequence of more lonely, socially isolated people using it for this purpose. In this data emotional use of GenAI is most pronounced in younger adults 18-34 (36.7% compared to 16% in those over 35), a group that has reported greater social isolation even prior to the advent of AI (Twenge et al., 2019). On the other hand, another study of 342 people (Novozhilova et al., 2026) reports that those using AI for companionship have similar self-esteem and friend networks. Moreover, studies have shown that

groups asked specifically to use GenAI for companionship developed significant increases in perceived attachment to AI and motivation to use GenAI for entertainment (Chandra et al., 2025), which means that this pattern of use can also be promoted and developed. Thus it is unlikely that the lower scores can be wholly attributed to the specific profiles of the people who choose to use GenAI for emotional purposes. Nonetheless, regardless of the explanation, one thing is clear, using AI for this purpose does not bring mind health into a strong, positive realm. This is an important consideration given large interest in the use of generative AI for therapy and other forms of emotional support.

Conclusions

Almost everything humans consume and need has an inverted-U shape association with physical health. For example, water is crucial for human survival, and insufficient water causes dehydration and, below certain limits, death. However, too much water can also lead to seizures, kidney failure, and death. So also, water is a solute that dissolves numerous metals and chemicals and therefore is never consumed in a 'pure' form. Many of these can be beneficial at trace levels while many can be toxic. Thus, a statement of whether water is good or bad for you must be qualified by how much, and what else is in it. The same can perhaps be said of GenAI. GenAI has many beneficial uses, and the inverted U-shape suggests that moderate use of GenAI in the right way can, in the aggregate, enhance human mind health. However, too much appears destructive, and consumption in certain forms is more destructive than others.

Regardless the direction of causation, the arithmetic of exposure is stark: 41% of AI users aged 18-34 already use it several times a day or more, the fall from moderate to constant use is 16 to 34 MHQ points within every population subgroup examined, and the fastest-growing type of use is the one associated with the worst mind health. With daily use of GenAI increasingly promoted

for both regular cognitive and emotional support, in the absence of strong guardrails, this suggests bleak outcomes for the collective mind of society.

Data availability statement

Data are available via the online data repository Brainbase, where data from the ongoing Global Mind Project are dynamically updated as respondents complete the MHQ. The anonymized data set, together with supporting information, is freely available for use for not-for-profit purposes and access can be requested at https://sapienlabs.org/global-mind-project/researcher-hub/.

Ethics statement

This study involves human participants. The data used in this study received ethics approval from the Health Media Lab Institutional Review Board (Office for Human Research Protections Institutional Review Board #00001211, Federal Wide Assurance #00001102, IORG #0000850). The study was conducted in accordance with local legislation and institutional requirements. Participants provided written informed consent to participate by clicking a start button after reading a detailed privacy policy.

Author contributions

JB: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. TT: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing - review & editing. Both authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work. TT is responsible for the overall content as guarantor.

Funding

The authors declare that financial support was received for the research and/or publication of this article. This work was supported by funding from Sapien Labs. It did not receive any specific award or grant from funding agencies in the public, commercial or not-for-profit sectors.

Acknowledgments

We thank members of the Sapien Labs team for their assistance with recruitment and data management. We are grateful to all survey respondents for their participation in the Global Mind Project.

Conflict of interest

None.

Generative AI statement

Claude (Anthropic), accessed through the Claude desktop application during 2026, was used in the preparation of this manuscript. It was used to partially write and check the Python code for the statistical analyses, generate figures and supplementary tables from the authors' data, edit figure and table captions, and copy-edit, format, and cross-reference. The authors specified every analysis, verified all outputs against the underlying data, wrote and revised the scientific content, and took full responsibility for the article as submitted.

References

American Psychological Association. (2026, June 16). Psychologists say patients are turning to chatbots as mental health professionals [Press release]. https://www.apa.org/news/press/releases/2026/06/patients-chatbots-mental-health

Bala, J., Sukhoi, O., Newson, J. J., Machado, P. P., Lawrence, M., & Thiagarajan, T. C. (2025). Estimation of the nature and magnitude of mental distress in the population associated with ultra-processed food consumption. Frontiers in Nutrition, 12, 1562286. https://doi.org/10.3389/fnut.2025.1562286

Bellan, R. (2025, October 6). Sam Altman says ChatGPT has hit 800M weekly active users. TechCrunch. https://techcrunch.com/2025/10/06/sam-altman-says-chatgpt-has-hit-800m-weekly-active-users/

Blanchflower, D. G., & Bryson, A. (2025). The declining mental health of the young and the global disappearance of the unhappiness hump shape in age. PLOS ONE, 20(8), e0327858. https://doi.org/10.1371/journal.pone.0327858

Chandra, M., Hernandez, J., Ramos, G., Ershadi, M., Bhattacharjee, A., Amores, J., Okoli, E., Paradiso, A., Warreth, S., & Suh, J. (2025). Longitudinal study on social and emotional use of AI conversational agent. arXiv preprint arXiv:2504.14112. https://arxiv.org/abs/2504.14112

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

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). How AI and human behaviors shape

psychosocial effects of extended chatbot use: A longitudinal controlled study. MIT Media Lab.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30, 3146–3154.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprint arXiv:2506.08872. https://arxiv.org/abs/2506.08872

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI ’25), Article 1121. https://doi.org/10.1145/3706598.3713778

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

McBain, R. K., Cantor, J. H., Breslau, J., Diliberti, M. K., Zhang, L. A., Zhang, F., Kofner, A., Stein, B. D., Burnett, A., Pataranutaporn, P., Mehrotra, A., & Rader, B. (2026). AI chatbot use and disclosure for mental health among US adolescents and young adults. JAMA Pediatrics. Advance online publication. https://doi.org/10.1001/jamapediatrics.2026.2015

Morrin, H., Nicholls, L., & Levin, M. (2025). Delusions by design? How everyday AIs might be fuelling psychosis (and what can be done about it) [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/cmy7n_v6

Nakagomi, A., Akutsu, Y., Yasuoka, M., Abe, N., Ihara, S., Teroh, T., & Tabuchi, T. (2026). AI companions and subjective well-being: Moderation by social connectedness and loneliness. Technology in Society, 85, 103229. https://doi.org/10.1016/j.techsoc.2026.103229

Newson, J. J., Marinova, Z., & Thiagarajan, T. C. (2025). Are the growing levels of neurotoxic and neuro-disruptive chemicals in our food and drink contributing to the youth mental health crisis? A narrative review. Neuroscience & Biobehavioral Reviews, 176, 106290. https://doi.org/10.1016/j.neubiorev.2025.106290

Newson, J. J., Pastukh, V., & Thiagarajan, T. C. (2022). Assessment of population well-being with the Mental Health Quotient: Validation study. JMIR Mental Health, 9(4), e34105. https://doi.org/10.2196/34105

Newson, J. J., Sukhoi, O., & Thiagarajan, T. C. (2024). MHQ: Constructing an aggregate metric of population mental wellbeing. Population Health Metrics, 22, 16. https://doi.org/10.1186/s12963-024-00336-y

Newson, J. J., & Thiagarajan, T. C. (2020). Assessment of population well-being with the Mental Health Quotient (MHQ): Development and usability study. JMIR Mental Health, 7(7), e17935. https://doi.org/10.2196/17935

Novozhilova, E., Vu, C., & Katz, J. (2026). From moral panic to normalization: comparing users and non-users of AI companionship apps. AI & Society, 41, 4057–4075. https://doi.org/10.1007/s00146-025-02751-7

OpenAI. (2025). Investigating affective use and emotional well-being on ChatGPT [Research report]. https://openai.com/index/affective-use-study/

Perlis, R. H., Gunning, F. M., Uslu, A. A., Santillana, M., Baum, M. A., Druckman, J. N., Ognyanova, K., & Lazer, D. (2026). Generative AI use and depressive symptoms among US adults. JAMA Network Open, 9(1), e2554820. https://doi.org/10.1001/jamanetworkopen.2025.54820

Pierre, J. M., Gaeta, B., Raghavan, G., & Sarma, K. (2025). “You’re not crazy”: A case of new-onset AI-associated psychosis. Innovations in Clinical Neuroscience, 22, 11–13.

Sapien Labs. (2023). The Mental State of the World in 2022. Global Mind Project. https://mentalstateoftheworld.report/wp-content/uploads/2023/02/Mental-State-of-the-World-2022.pdf

Sapien Labs. (2024). Mental State of the World Report 2023 (4th Annual Report). Global Mind Project. https://mentalstateoftheworld.report/

Sapien Labs. (2025). Mental State of the World Report 2024 (5th Annual Report). Global Mind Project. https://mentalstateoftheworld.report/

Thiagarajan, T. C., Newson, J. J., & Swaminathan, S. (2025). Protecting the developing mind in a digital age: A global policy imperative. Journal of Human Development and Capabilities, 26(3), 493–504. https://doi.org/10.1080/19452829.2025.2518313

Twenge, J. M., & Blanchflower, D. G. (2025). Declining life satisfaction and happiness among young adults in six English-speaking countries (NBER Working Paper No. 33490). National Bureau of Economic Research. http://www.nber.org/papers/w33490

Twenge, J. M., Spitzberg, B. H., & Campbell, W. K. (2019). Less in-person social interaction with peers among U.S. adolescents in the 21st century and links to loneliness. Journal of Social and Personal Relationships, 36(6), 1892–1913. https://doi.org/10.1177/0265407519836170

Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. Biometrika, 34(1/2), 28–35.

Multimedia Appendix 1

Supplementary figures and tables. Figures are numbered S1 to S6 and tables S1 to S16, in two separate series, as required for a combined Multimedia Appendix.

A. Supplementary figures

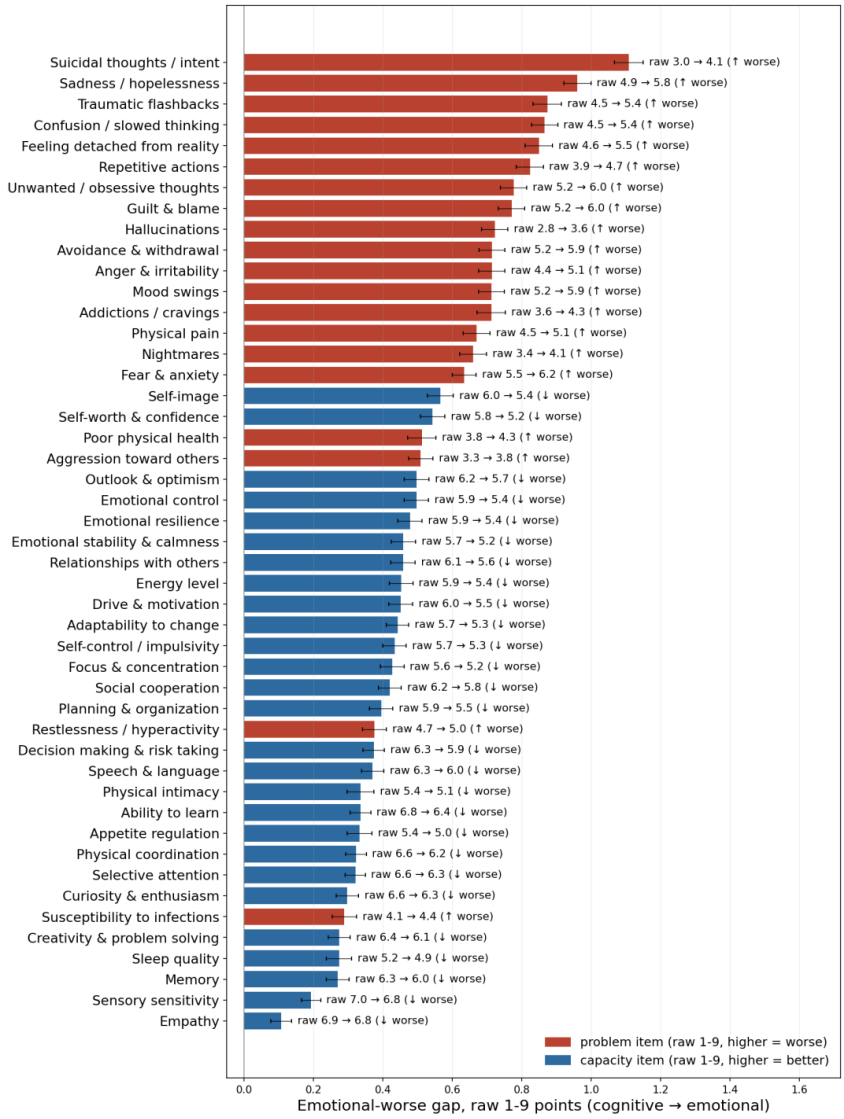


Figure S1. All 47 MHQ items ranked by the gap between those using AI for emotional as well as cognitive purposes and those using it for cognitive purposes only.

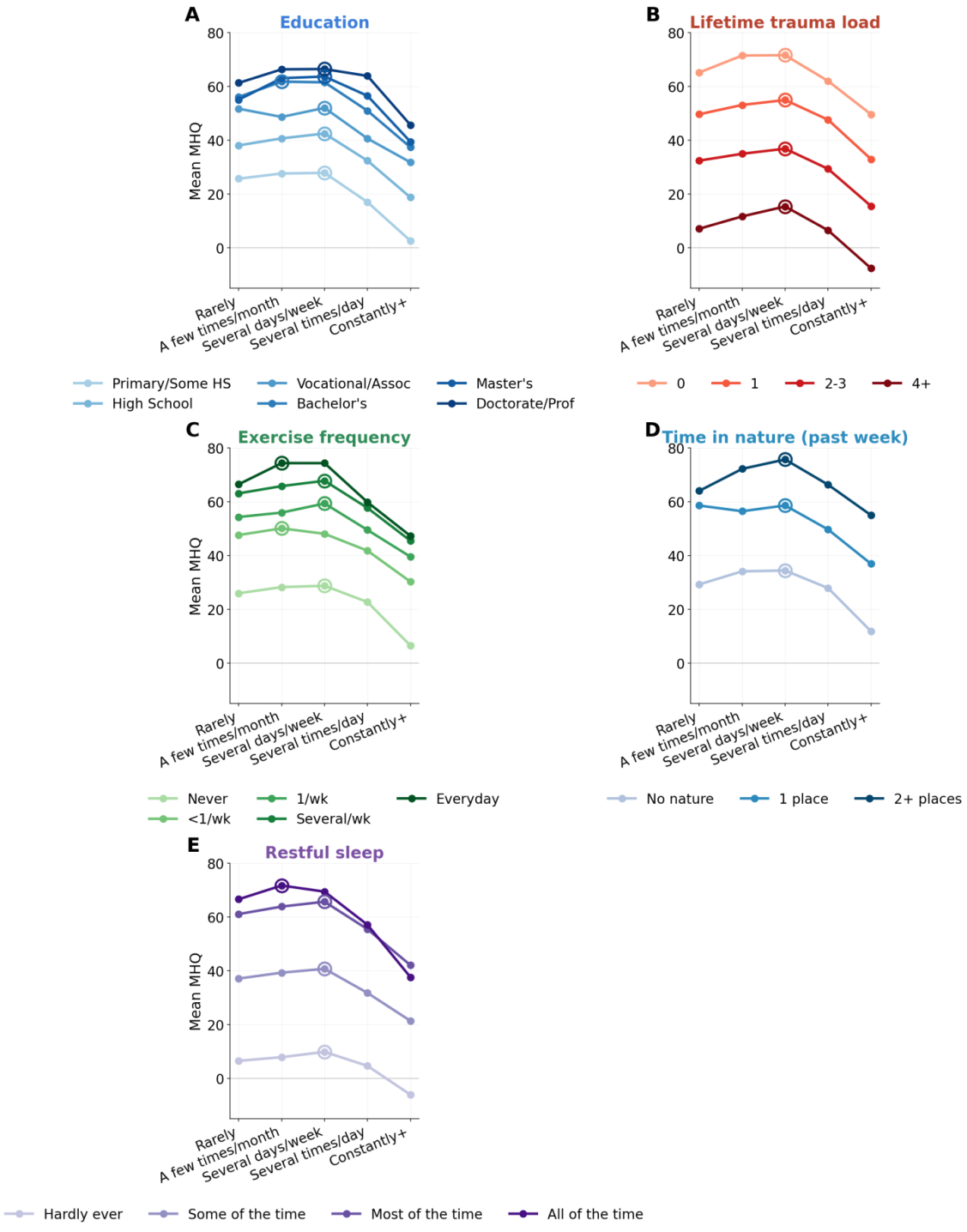


Figure S2. The dose-response curves behind Figure 7, one line per covariate level.

B. AI-use frequency and mind health

Table S1. Percentage of users at each AI-use frequency, by age group and sex. Each column sums to 100%.

AI-use frequency	18-34 Female	18-34 Male	35+ Female	35+ Male
Rarely	22.7	21.7	49.2	45.7
A few times/month	12.3	12	14.9	15.1
Several days/week	23.9	24.5	18.5	20.2
Several times/day	19.4	20.3	9.7	11.5
Constantly+	21.6	21.5	7.8	7.5

Table S2. Mean MHQ by AI-use frequency and age band (18-34 pooled shown in the last column).

AI-use frequency	18-24	25-34	35-44	45-54	55-64	65+	18-34 (pooled)
Rarely	36.6	57.5	73.7	88.6	100.8	111.9	45
A few times/month	43	61.6	76.6	91.4	103.9	114.3	49.3
Several days/week	45.8	63.1	74.5	87.5	102.5	111.7	50.7
Several times/day	36.5	54.4	66.8	79.5	96.9	108.8	41.5
Constantly+	20.7	45.7	60	77.6	92.3	106.9	27.2

Table S3. Mean wellbeing-dimension score by AI-use frequency (18-34).

AI-use frequency	Cognition	Adaptability	Drive	Mood & Outlook	Social Self	Mind-Body
Rarely	59	74.8	69	47.4	46.1	61.1
A few times/month	62.3	76.6	71	51	49.9	65.2
Several days/week	61.8	75.8	71	53	52	65.8
Several times/day	53.4	69.2	64.8	44.8	43	59.5
Constantly+	41.1	57.6	53.4	32.1	31	49.4

Table S4. All 27 capacity items: raw mean at rare use and the change at each higher AI-use frequency (18-34), ranked by the change at Constantly+. The last three columns give each item's own peak use level, the change from that peak to constant use, and Cohen's d for that change. Colored rows are the four shown in Figure 4A.

MHQ item	Rarely (raw 1-9)	Δ at A few times/month	Δ at Several days/week	Δ at Several times/day	Δ at Constantly+	Own peak level	Δ peak to Constantly+	Cohen d
Focus & concentration	5.72	−0.07	−0.23	−0.27	−0.58	Rarely	−0.58	−0.242
Self-image	5.89	−0.01	−0.01	+0.01	−0.43	Several times/day	−0.44	−0.171
Emotional control	5.79	+0.05	−0.01	−0.06	−0.41	A few times/month	−0.46	−0.188
Emotional stability & calmness	5.61	−0.01	−0.02	−0.11	−0.38	Rarely	−0.38	−0.153
Self-control / impulsivity	5.63	+0.08	+0.06	−0.05	−0.35	A few times/month	−0.43	−0.178
Speech & language	6.32	+0.03	−0.03	−0.19	−0.34	A few times/month	−0.37	−0.161
Energy level	5.79	−0.02	−0.08	−0.06	−0.32	Rarely	−0.32	−0.133
Self-worth & confidence	5.65	+0.01	+0.01	−0.03	−0.32	Several days/week	−0.34	−0.137
Social cooperation	6.12	+0.04	+0.00	−0.13	−0.31	A few times/month	−0.36	−0.154
Ability to learn	6.74	+0.02	−0.04	−0.10	−0.31	A few times/month	−0.33	−0.158
Planning & organization	5.87	+0.00	−0.04	−0.04	−0.31	A few times/month	−0.31	−0.130
Outlook & optimism	6.10	−0.01	+0.02	+0.03	−0.30	Several times/day	−0.33	−0.133

Decision making & risk taking	6.25	−0.01	−0.03	−0.08	−0.28	Rarely	−0.28	−0.125
Emotional resilience	5.80	+0.03	+0.03	−0.04	−0.28	Several days/week	−0.31	−0.127
Appetite regulation	5.31	+0.04	−0.03	−0.06	−0.27	A few times/month	−0.31	−0.125
Drive & motivation	5.92	−0.00	−0.06	−0.08	−0.27	Rarely	−0.27	−0.109
Relationships with others	5.94	+0.08	+0.06	−0.07	−0.25	A few times/month	−0.34	−0.136
Memory	6.21	+0.06	−0.00	−0.06	−0.22	A few times/month	−0.28	−0.124
Selective attention	6.57	+0.04	+0.02	−0.05	−0.22	A few times/month	−0.26	−0.128
Physical coordination	6.52	+0.03	−0.01	−0.16	−0.21	A few times/month	−0.24	−0.115
Creativity & problem solving	6.31	+0.03	+0.04	−0.02	−0.17	Several days/week	−0.21	−0.100
Curiosity & enthusiasm	6.48	+0.01	+0.02	+0.02	−0.16	Several times/day	−0.18	−0.081
Physical intimacy	5.37	+0.06	−0.03	−0.17	−0.15	A few times/month	−0.21	−0.077
Adaptability to change	5.56	+0.09	+0.09	−0.04	−0.12	A few times/month	−0.21	−0.093
Empathy	6.84	+0.08	+0.07	+0.00	−0.09	A few times/month	−0.17	−0.081
Sleep quality	5.02	+0.11	+0.10	+0.13	−0.08	Several times/day	−0.21	−0.079
Sensory sensitivity	6.89	+0.06	+0.09	+0.04	−0.04	Several days/week	−0.13	−0.067

Table S5. All 20 problem items, raw (not reverse-coded): raw mean at rare use and the change at each higher AI-use frequency (18-34), ranked by the change at Constantly+. The last three columns give each item's own best use level, the one at which its raw score is lowest, the change from that level to constant use, and Cohen's d for that change. Colored rows are the four shown in Figure 4B.

MHQ item	Rarely (raw 1-9)	Δ at A few times/month	Δ at Several days/week	Δ at Several times/day	Δ at Constantly+	Own best level	Δ best to Constantly+	Cohen d
Hallucinations	3.16	−0.18	−0.29	−0.05	+0.24	Several days/week	+0.53	0.205
Susceptibility to infections	4.12	+0.02	−0.05	+0.15	+0.26	Several days/week	+0.31	0.124
Aggression toward others	3.42	−0.05	−0.15	+0.10	+0.27	Several days/week	+0.41	0.170
Poor physical health	3.93	−0.09	−0.10	+0.12	+0.27	Several days/week	+0.38	0.132
Nightmares	3.66	−0.07	−0.21	+0.04	+0.28	Several days/week	+0.50	0.186
Restlessness / hyperactivity	4.75	−0.10	−0.08	+0.07	+0.29	A few times/month	+0.39	0.160
Physical pain	4.70	−0.13	−0.21	+0.01	+0.30	Several days/week	+0.51	0.188
Traumatic flashbacks	4.73	−0.09	−0.11	+0.19	+0.44	Several days/week	+0.55	0.191
Suicidal thoughts / intent	3.35	−0.11	−0.23	−0.05	+0.46	Several days/week	+0.69	0.239
Avoidance & withdrawal	5.36	−0.05	+0.00	+0.14	+0.46	A few times/month	+0.51	0.192
Feeling detached from reality	4.85	−0.10	−0.14	+0.12	+0.46	Several days/week	+0.60	0.216
Anger & irritability	4.52	−0.08	−0.11	+0.18	+0.46	Several days/week	+0.57	0.216
Fear & anxiety	5.65	−0.03	−0.00	+0.16	+0.47	A few times/month	+0.50	0.204

Mood swings	5.35	−0.08	−0.14	+0.12	+0.47	Several days/week	+0.60	0.229
Guilt & blame	5.35	−0.04	−0.02	+0.16	+0.49	A few times/month	+0.53	0.194
Repetitive actions	4.09	−0.10	−0.08	+0.19	+0.52	A few times/month	+0.61	0.221
Sadness / hopelessness	5.13	−0.12	−0.14	+0.11	+0.55	Several days/week	+0.69	0.246
Unwanted / obsessive thoughts	5.39	−0.10	−0.03	+0.19	+0.58	A few times/month	+0.68	0.251
Confusion / slowed thinking	4.62	−0.02	−0.00	+0.28	+0.60	A few times/month	+0.63	0.230
Addictions / cravings	3.60	+0.06	+0.17	+0.48	+0.68	Rarely	+0.68	0.234

C. Purpose of use

Table S6. Purpose-segment composition by AI-use frequency (18-34); each row's two percentages sum to 100%.

AI-use frequency	n cognitive only	n emotional	n total	% cognitive only	% emotional
Rarely	14,513	3,613	18,126	80.1	19.9
A few times/month	7,108	2,741	9,849	72.2	27.8
Several days/week	12,897	6,588	19,485	66.2	33.8
Several times/day	8,835	7,079	15,914	55.5	44.5
Constantly+	7,799	9,569	17,368	44.9	55.1

Table S7. Mean MHQ by AI-use frequency and purpose segment (18-34); the last column is the gap between the two segments at each use frequency.

AI-use frequency	MHQ cognitive only	MHQ emotional	Gap
Rarely	50.7	22.4	28.3
A few times/month	58.2	26.1	32.1
Several days/week	60.6	31.4	29.2
Several times/day	53.1	27	26
Constantly+	43.4	14.1	29.3

Table S8. All 13 AI-use aspects: prevalence, mean MHQ among people who report the use and among those who do not, and the raw difference (all AI users with a known age band, n = 178,009; reported sex not required), ranked from the most positive to the most negative difference.

Use type	Category	Prevalence (%)	MHQ if reported	MHQ if not	Difference	95% CI
Planning trips/schedules	cognitive	5.9	78.3	69.7	8.6	7.2, 10.0
Learning concepts/skills	cognitive	30.5	75.0	68.1	6.9	6.2, 7.6
Quick facts / info	cognitive	63.9	72.3	66.5	5.9	5.2, 6.6
Translating	cognitive	16.6	71.6	69.9	1.6	0.7, 2.5
Creating content	cognitive	12.6	67.0	70.7	−3.7	−4.7, −2.7
Writing support	cognitive	20.8	67.1	71.0	−3.9	−4.7, −3.1
Analyzing / calculations	cognitive	9.4	65.7	70.7	−5.0	−6.1, −3.8
Summarizing	cognitive	15.3	62.4	71.6	−9.2	−10.1, −8.2
Coding / debugging	cognitive	5.0	57.5	70.9	−13.4	−14.9, −11.8
Problem-solving / decisions	cognitive	26.8	52.5	76.7	−24.2	−25.0, −23.5
Companionship	emotional	8.3	39.3	73.0	−33.7	−35.0, −32.4
Emotional support	emotional	18.3	38.0	77.4	−39.5	−40.3, −38.6
Romantic / intimate	emotional	4.2	29.0	72.0	−43.0	−44.7, −41.4

D. Lifestyle and life circumstance

Table S9. The moderate-use plateau, the "Constantly+" mean, and the drop between them, for every level of the five covariates (AI users aged 18-34). Each level's plateau is its own peak across the four lower use frequencies, so no level is measured against another; cells with fewer than 100 people are excluded.

Covariate	Level	Plateau level	Plateau MHQ	Constantly+ MHQ	Drop	95% CI	n
Education	Primary/Some HS	Several days/week	27.9	2.6	25.2	20.4, 30.1	7,066
Education	High School	Several days/week	42.5	18.7	23.7	21.3, 26.2	25,402
Education	Vocational/Assoc	Several days/week	52.0	31.8	20.2	15.3, 25.2	7,154
Education	Bachelor's	A few times/month	61.8	37.4	24.4	21.3, 27.5	25,850
Education	Master's	Several days/week	63.7	39.4	24.3	19.5, 29.0	7,190
Education	Doctorate/Prof	Several days/week	66.4	45.7	20.8	10.0, 31.5	1,581
Lifetime trauma load	0	Several days/week	71.6	49.6	22.0	19.7, 24.2	32,729
Lifetime trauma load	1	Several days/week	54.9	33.0	22.0	18.7, 25.3	14,321
Lifetime trauma load	2-3	Several days/week	36.9	15.5	21.3	18.6, 24.0	19,479
Lifetime trauma load	4+	Several days/week	15.4	−7.6	23.0	19.9, 26.0	14,223
Exercise frequency	Never	Several days/week	28.8	6.5	22.2	20.1, 24.3	32,413
Exercise frequency	<1/wk	A few times/month	50.1	30.3	19.8	15.0, 24.7	9,667
Exercise frequency	1/wk	Several days/week	59.3	39.6	19.8	15.1, 24.5	7,446
Exercise frequency	Several/wk	Several days/week	67.8	45.5	22.3	19.5, 25.1	20,734
Exercise frequency	Everyday	A few times/month	74.3	47.3	27.1	21.9, 32.3	10,454
Time in nature (past week)	No nature	Several days/week	34.4	11.9	22.6	20.6, 24.5	39,952
Time in nature (past week)	1 place	Several days/week	58.6	36.9	21.7	19.1, 24.3	25,506
Time in nature (past week)	2+ places	Several days/week	75.7	55.0	20.7	17.4, 24.0	15,294
Restful sleep	Hardly ever	Several days/week	9.9	−6.0	15.9	11.9, 19.9	8,878
Restful sleep	Some of the time	Several days/week	40.8	21.4	19.4	17.2, 21.5	32,321
Restful sleep	Most of the time	Several days/week	65.7	42.1	23.6	21.2, 25.9	30,043
Restful sleep	All of the time	A few times/month	71.7	37.7	34.1	28.4, 39.7	9,510

Table S10. Mean MHQ at every AI-use frequency for every covariate level, with the number of people in each cell (AI users aged 18-34; cells with fewer than 100 people are shown as dashes).

Covariate	Level	Rarely	A few times/month	Several days/week	Several times/day	Constantly+
Education	Primary/Some HS	25.7 (2,007)	27.7 (842)	27.9 (1,540)	17.0 (1,359)	2.6 (1,318)
Education	High School	38.1 (5,826)	40.7 (3,096)	42.5 (6,219)	32.4 (4,830)	18.7 (5,431)
Education	Vocational/Assoc	51.7 (1,865)	48.6 (957)	52.0 (1,616)	40.6 (1,170)	31.8 (1,546)
Education	Bachelor's	56.1 (5,142)	61.8 (3,113)	61.5 (6,468)	51.0 (5,405)	37.4 (5,722)
Education	Master's	55.0 (1,358)	63.1 (856)	63.7 (1,777)	56.5 (1,634)	39.4 (1,565)
Education	Doctorate/Prof	61.3 (324)	66.3 (203)	66.4 (366)	63.9 (359)	45.7 (329)
Lifetime trauma load	0	65.2 (7,636)	71.5 (4,012)	71.6 (8,014)	62.0 (6,329)	49.6 (6,738)
Lifetime trauma	1	49.7 (3,248)	53.1 (1,812)	54.9 (3,538)	47.6 (2,789)	33.0 (2,934)

load						
Lifetime trauma load	2-3	32.4 (4,191)	35.0 (2,381)	36.9 (4,588)	29.4 (3,979)	15.5 (4,340)
Lifetime trauma load	4+	7.1 (3,056)	11.7 (1,644)	15.4 (3,346)	6.5 (2,819)	−7.6 (3,358)
Exercise frequency	Never	26.0 (7,829)	28.3 (3,811)	28.8 (7,238)	22.7 (6,253)	6.5 (7,282)
Exercise frequency	<1/wk	47.6 (1,993)	50.1 (1,255)	48.1 (2,483)	41.8 (1,987)	30.3 (1,949)
Exercise frequency	1/wk	54.3 (1,566)	56.0 (998)	59.3 (1,938)	49.5 (1,489)	39.6 (1,455)
Exercise frequency	Several/wk	63.1 (4,427)	65.8 (2,622)	67.8 (5,512)	57.7 (4,024)	45.5 (4,149)
Exercise frequency	Everyday	66.5 (2,304)	74.3 (1,155)	74.3 (2,309)	59.8 (2,157)	47.3 (2,529)
Time in nature (past week)	No nature	29.3 (8,980)	34.1 (4,568)	34.4 (9,306)	27.9 (8,174)	11.9 (8,924)
Time in nature (past week)	1 place	58.6 (5,948)	56.5 (3,291)	58.6 (5,995)	49.7 (4,885)	36.9 (5,387)
Time in nature (past week)	2+ places	64.0 (3,203)	72.2 (1,990)	75.7 (4,185)	66.3 (2,857)	55.0 (3,059)
Restful sleep	Hardly ever	6.6 (2,225)	7.9 (972)	9.9 (1,904)	4.7 (1,557)	−6.0 (2,220)
Restful sleep	Some of the time	37.1 (7,594)	39.3 (3,975)	40.8 (7,738)	31.8 (6,194)	21.4 (6,820)
Restful sleep	Most of the time	61.1 (6,165)	63.9 (3,848)	65.7 (7,814)	55.5 (6,257)	42.1 (5,959)
Restful sleep	All of the time	66.6 (2,147)	71.7 (1,054)	69.4 (2,030)	57.1 (1,908)	37.7 (2,371)

E. Item-level differences between the cognitive-only and emotional segments

Table S11. The 20 problem items, ranked by how far they separate the cognitive-only and emotional use segments (18-34); raw 1-9 means, higher = worse, not reverse-coded. The top rows are among those drawn in Figure 6B.

MHQ item	Cognitive only	Emotional	Gap (emotional worse)	95% CI	Cohen d	Magnitude
Suicidal thoughts / intent	2.97	4.08	1.11	1.07, 1.15	0.395	small
Sadness / hopelessness	4.87	5.83	0.96	0.92, 1.00	0.347	small
Traumatic flashbacks	4.50	5.38	0.87	0.83, 0.91	0.306	small
Confusion / slowed thinking	4.49	5.35	0.87	0.83, 0.90	0.325	small
Feeling detached from reality	4.61	5.46	0.85	0.81, 0.89	0.307	small
Repetitive actions	3.91	4.73	0.82	0.78, 0.86	0.305	small
Unwanted / obsessive thoughts	5.25	6.02	0.78	0.74, 0.81	0.290	small
Guilt & blame	5.19	5.96	0.77	0.73, 0.81	0.285	small
Hallucinations	2.85	3.57	0.72	0.68, 0.76	0.282	small
Avoidance & withdrawal	5.22	5.93	0.71	0.68, 0.75	0.271	small
Anger & irritability	4.36	5.07	0.71	0.68, 0.75	0.271	small
Mood swings	5.17	5.88	0.71	0.68, 0.75	0.271	small
Addictions / cravings	3.62	4.34	0.71	0.67, 0.75	0.249	small
Physical pain	4.45	5.12	0.67	0.63, 0.71	0.248	small
Nightmares	3.42	4.08	0.66	0.62, 0.70	0.249	small
Fear & anxiety	5.55	6.18	0.63	0.60, 0.67	0.263	small
Poor physical health	3.79	4.30	0.51	0.47, 0.55	0.182	very small
Aggression toward others	3.27	3.78	0.51	0.47, 0.54	0.209	small
Restlessness / hyperactivity	4.66	5.03	0.38	0.34, 0.41	0.156	very small
Susceptibility to infections	4.09	4.38	0.29	0.25, 0.32	0.117	very small

Table S12. The 27 capacity items, ranked the same way (18-34); raw 1-9 means, higher = better.

MHQ item	Cognitive only	Emotional	Gap (emotional worse)	95% CI	Cohen d	Magnitude
Self-image	6.01	5.44	0.57	−0.60, −0.53	−0.223	small
Self-worth & confidence	5.78	5.23	0.54	−0.58, −0.51	−0.223	small
Outlook & optimism	6.23	5.73	0.50	−0.53, −0.46	−0.204	small
Emotional control	5.88	5.38	0.50	−0.53, −0.46	−0.207	small
Emotional resilience	5.92	5.44	0.48	−0.51, −0.44	−0.198	very small
Emotional stability & calmness	5.67	5.21	0.46	−0.49, −0.42	−0.188	very small
Relationships with others	6.07	5.61	0.46	−0.49, −0.42	−0.189	very small
Energy level	5.86	5.40	0.45	−0.49, −0.42	−0.193	very small
Drive & motivation	5.99	5.54	0.45	−0.49, −0.42	−0.190	very small
Adaptability to change	5.72	5.28	0.44	−0.47, −0.41	−0.197	very small
Self-control / impulsivity	5.73	5.30	0.43	−0.47, −0.40	−0.186	very small
Focus & concentration	5.63	5.20	0.43	−0.46, −0.39	−0.181	very small
Social cooperation	6.19	5.77	0.42	−0.45, −0.39	−0.185	very small
Planning & organization	5.93	5.53	0.40	−0.43, −0.36	−0.171	very small
Decision making & risk taking	6.30	5.93	0.37	−0.40, −0.34	−0.176	very small
Speech & language	6.35	5.98	0.37	−0.40, −0.34	−0.166	very small
Physical intimacy	5.43	5.10	0.34	−0.37, −0.30	−0.124	very small
Ability to learn	6.77	6.43	0.34	−0.37, −0.31	−0.164	very small
Appetite regulation	5.35	5.02	0.33	−0.37, −0.30	−0.136	very small
Physical coordination	6.57	6.24	0.32	−0.35, −0.29	−0.156	very small
Selective attention	6.64	6.32	0.32	−0.35, −0.29	−0.162	very small
Curiosity & enthusiasm	6.56	6.26	0.30	−0.33, −0.26	−0.133	very small
Creativity & problem solving	6.38	6.11	0.27	−0.31, −0.24	−0.128	very small
Sleep quality	5.16	4.89	0.27	−0.31, −0.24	−0.107	very small
Memory	6.26	5.99	0.27	−0.30, −0.24	−0.120	very small
Sensory sensitivity	6.99	6.80	0.19	−0.22, −0.17	−0.100	very small
Empathy	6.88	6.78	0.11	−0.14, −0.08	−0.051	negligible

F. Statistical robustness

Table S13. Cohen's d for the principal comparisons reported in this manuscript, shown alongside the difference in means and its 95% confidence interval. AI users aged 18-34 unless stated. Each group's moderate-use plateau is its own highest mean across the four lower use frequencies. Cohen's d is the difference divided by the pooled standard deviation of the two groups. Every comparison here has a Welch P value below .001, which is why effect sizes rather than P values are tabulated; red marks |d| at or above 0.30 and blue 0.15 to 0.30.

Comparison	Difference	95% CI	Cohen d	Magnitude
Overall MHQ: moderate-use plateau versus constant use, by age				
18-24	−25.10	−26.75, −23.45	−0.37	small
25-34	−17.37	−20.22, −14.52	−0.24	small
18-34 (pooled)	−23.51	−24.95, −22.07	−0.33	small
35-44	−16.59	−20.29, −12.90	−0.23	small
65+	−7.48	−11.49, −3.47	−0.13	very small
The six MHQ dimensions, ages 18-34, plateau versus constant use				
Cognition	−21.18	−22.82, −19.54	−0.32	small

Social Self	−20.97	−22.46, −19.48	−0.29	small
Mood & Outlook	−20.88	−22.26, −19.50	−0.31	small
Adaptability & Resilience	−19.04	−20.74, −17.33	−0.27	small
Drive & Motivation	−17.63	−19.27, −16.00	−0.26	small
Mind-Body	−16.39	−17.70, −15.09	−0.26	small
Type of use, ages 18-34				
emotional use versus cognitive-only use	30.36	29.37, 31.34	0.44	small
Purpose of use versus everyone else, all ages				
romantic or intimate conversation	−43.01	−44.65, −41.37	−0.60	medium
emotional support or life coaching	−39.47	−40.33, −38.61	−0.56	medium
companionship	−33.70	−34.97, −32.43	−0.47	small
problem solving or decisions	−24.22	−24.98, −23.46	−0.34	small
coding or debugging	−13.38	−14.95, −11.81	−0.18	very small
learning new concepts or skills	6.89	6.16, 7.62	0.09	negligible
planning trips or schedules	8.61	7.18, 10.04	0.12	very small

Table S14. How often the same Welch t-test reaches P<.05 when it is run on a smaller random sample. For each size, 2,000 subsamples were drawn, taking that many respondents from each of the two groups being compared, and the test was repeated. Each right-hand column gives the percentage of those 2,000 subsamples in which the test reached P<.05. AI users aged 18-34.

Contrast	Unit	Full difference	n group A	n group B	n=50	n=100	n=200	n=500	n=1000
Overall MHQ, moderate vs constant use (18-34)	MHQ points	−23.51	17,370	19,486	39.1	64.1	91.6	100	100
Overall MHQ, cognitive-only vs emotional use (18-34)	MHQ points	−30.36	29,590	51,152	57.8	87.4	99.2	100	100
Cognition dimension, moderate vs constant use	MHQ points	−20.73	17,370	19,486	34.8	59.3	88.2	99.8	100
Focus and concentration, rarely vs constant use	raw 1-9	−0.58	17,370	18,131	22.2	39.6	68.3	97.2	100
Sadness and hopelessness, rarely vs constant use	raw 1-9	0.55	17,370	18,131	17.8	26.4	49.5	85.6	98.5

G. Multivariate SHAP analysis

Table S15. Every variable carried into the multivariate model described in this section: the survey response options as they are recorded in the data, the numeric coding applied to each, and the share of the 18-34 analysis sample (n = 80,752) at each value. The ten cognitive and three emotional purposes are separate options of two multiple-select questions, so a respondent may report any number of them and the percentages within a block do not sum to 100. Response text is reproduced exactly as it is recorded in the survey instrument, including its British spellings. "Companionship" and "Emotional support, life coaching or counselling" are two distinct response options: "emotional use" means reporting any of the three options in that block, while "companionship" means only the single option of that name. Categories left missing are handled natively by the gradient-boosted model rather than imputed.

Variable	Response options	Coding	% of sample

OUTCOME			
Overall MHQ score	Composite of the 47 MHQ items (27 capacities, 20 problems), each rated on a 1-9 scale	Continuous, −100 to +200; population mean calibrated to 100	
AI USE — THE AI BLOCK OF THE MODEL (14 FEATURES)			
AI-use frequency	"Rarely" / "A few times a month" / "Several days a week" / "Several times a day" / "Constantly" (also recorded as "All of the time" and "Constantly (when awake)")	Ordinal 1-5 in the order shown. "I have never used an AI assistant" is coded 0 and excluded from the analysis sample. Sample split across the five levels: 22.5 / 12.2 / 24.1 / 19.7 / 21.5	
Quick facts / info	"Finding quick facts or information"	Binary 0/1	64.9
Learning concepts/skills	"Learning new concepts or skills"	Binary 0/1	37.1
Problem-solving / decisions	"Solving problems or making decisions"	Binary 0/1	35.6
Writing support	"Writing support"	Binary 0/1	23.5
Summarizing	"Summarising long content"	Binary 0/1	22.5
Translating	"Translating languages"	Binary 0/1	18.8
Creating content	"Creating content (e.g. images, videos, music, stories)"	Binary 0/1	15.1
Analyzing / calculations	"Analysing data or doing calculations"	Binary 0/1	12.8
Coding / debugging	"Coding or debugging"	Binary 0/1	8.3
Planning trips/schedules	"Planning trips, schedules, or events"	Binary 0/1	6.6
Emotional support (emotional)	"Emotional support, life coaching or counselling"	Binary 0/1	27.5
Companionship (emotional)	"Companionship"	Binary 0/1	11.9
Romantic / intimate (emotional)	"Romantic or intimate conversation"	Binary 0/1	6.9
DEMOGRAPHICS — THE CONTEXT BLOCK (4 FEATURES)			
Age 25-34 (vs 18-24)	Age band	Binary: 18-24 = 0, 25-34 = 1	31.2
Male (vs female)	Biological sex	Binary: Female = 0, Male = 1. "Other/Intersex" and "Prefer not to say" left missing	37.6
Education (0-5)	"Primary Education", "Some High School", "High School", "Vocational certification", "Associate's Degree", "Bachelor's Degree", "Master's Degree", "PhD", "M.D.", "J.D."	Ordinal 0-5: Primary/Some High School = 0; High School = 1; Vocational or Associate's = 2; Bachelor's = 3; Master's = 4; PhD/M.D./J.D. = 5. "Other" and "Prefer not to say" left missing	
English survey	Language in which the survey was completed	Binary 0/1	34.6
LIFE CIRCUMSTANCE — THE CONTEXT BLOCK (4 FEATURES)			
Lifetime trauma count	Count of traumatic experiences reported across the childhood and adulthood trauma questions (abuse and assault, parental neglect, death and illness, financial adversity and others)	Count, 0-19 observed; mean 1.7	
Restful sleep (0-3)	"Hardly ever" / "Rarely" / "Some of the time" / "Most of the time" / "All of the time"	Ordinal 0-3: Hardly ever and Rarely = 0; Some of the time = 1; Most of the time = 2; All of the time = 3	
Exercise frequency (0-4)	Frequency of exercising 30 minutes or more: "Rarely/Never" / "Less than once a week" / "Once a week" / "Several days a week" / "Everyday"	Ordinal 0-4 in the order shown. Sample split across the five levels: 40.1 / 12.0 / 9.2 / 25.7 / 13.0	
Nature places past week (0-8)	Places spent time in during the past week, multiple-select: "Park, playing field, playground or grassy area near your home"; "Garden (yours or someone else's)"; "Beach, seaside or ocean"; "Fields, farmland or countryside"; "Woods or forest"; "Rivers, lakes or canals"; "Paid-entry green spaces (e.g. botanical gardens, city farms)"; "Other natural or green space not listed here"; "I	Count of places reported, 0-8; the none-of-these option is coded 0. 50.5% report none, 31.4% one, 17.9% two or more; mean 0.8	

	haven't spent time in any of these places in the past week"		

Rationale and method

SHAP (SHapley Additive exPlanations) assigns each feature its Shapley value: the feature's average marginal contribution to a respondent's prediction, averaged over all possible orderings of the features. This yields an exact additive decomposition in MHQ points (prediction = sample-average baseline plus one contribution per feature) and serves here as a robustness check that the paper's descriptive findings survive joint multivariate treatment. SHAP explains the model, not the world: attributions are adjusted associations that split credit among correlated features and inherit the cross-sectional design; they are not causal effects. Two further limits bound the reading. Because the AI features are correlated with trauma, sleep, exercise, time in nature and education, SHAP splits credit among them in a way that depends on the fitted model, and it does not adjudicate confounding versus mediation. A decomposition is only as informative as the model behind it: at a held-out R^2 of 0.276 these attributions describe a real but partial signal.

Model, sample and performance

One example is one respondent, represented by a 22-number feature vector used to predict the target: overall MHQ. The AI block within this feature vector comprises use frequency (1-5 ordinal) and thirteen 0/1 purpose flags (ten cognitive, three emotional); the demographic/life context block comprises age band (25-34 vs 18-24), sex, education (0-5 ordinal), lifetime trauma count, restful sleep (0-3 ordinal), exercise frequency (0-4 ordinal), the number of natural or green spaces visited in the past week (0-8) and an English-survey indicator. Every covariate shown in Figure 7 is therefore also carried in the model. Full response options and coding for all 22 features are given in Table S15. None of the 47 MHQ items is a feature: the outcome is never predicted from its own ingredients. The sample is the 18-34 analysis population (n = 80,752), split once into 75% training and 25% held-out test data; every metric and every SHAP value reported here comes from the held-out set. The model used is a LightGBM gradient-boosted tree ensemble (600 trees, learning rate 0.05, 63 leaves, minimum 200 observations per leaf, with 90% row and column subsampling and a fixed random seed); exact SHAP values are computed with TreeExplainer, using the training set as the background distribution. Held-out performance: $R^2 = 0.276$, mean absolute error 50.4 MHQ points (outcome SD 71.1); a companion classifier for MHQ < 0 with the same features reaches AUC = 0.740. The four lifestyle and life circumstance measures carry the largest contributions overall, as expected; the AI-use features rank behind them (Figure S3). In mean absolute contribution these are lifetime trauma 15.5, exercise frequency 11.7, restful sleep 10.6 and time in nature 8.4 MHQ points, followed by learning new concepts 6.6, age band 6.4, emotional-support use 5.8, education 5.1 and use frequency 4.3. Adding exercise and time in nature to the 20-feature model reported previously raises held-out R^2 from 0.226 to 0.276 and reduces the mean absolute contribution of use frequency from 4.9 to 4.3 points, while leaving the shape of its dose response unchanged. Because both are plausibly displaced by heavy AI use rather than merely correlated with it, part of that reduction may be mediation rather than confounding, and the adjusted frequency contribution should be read as a lower bound.

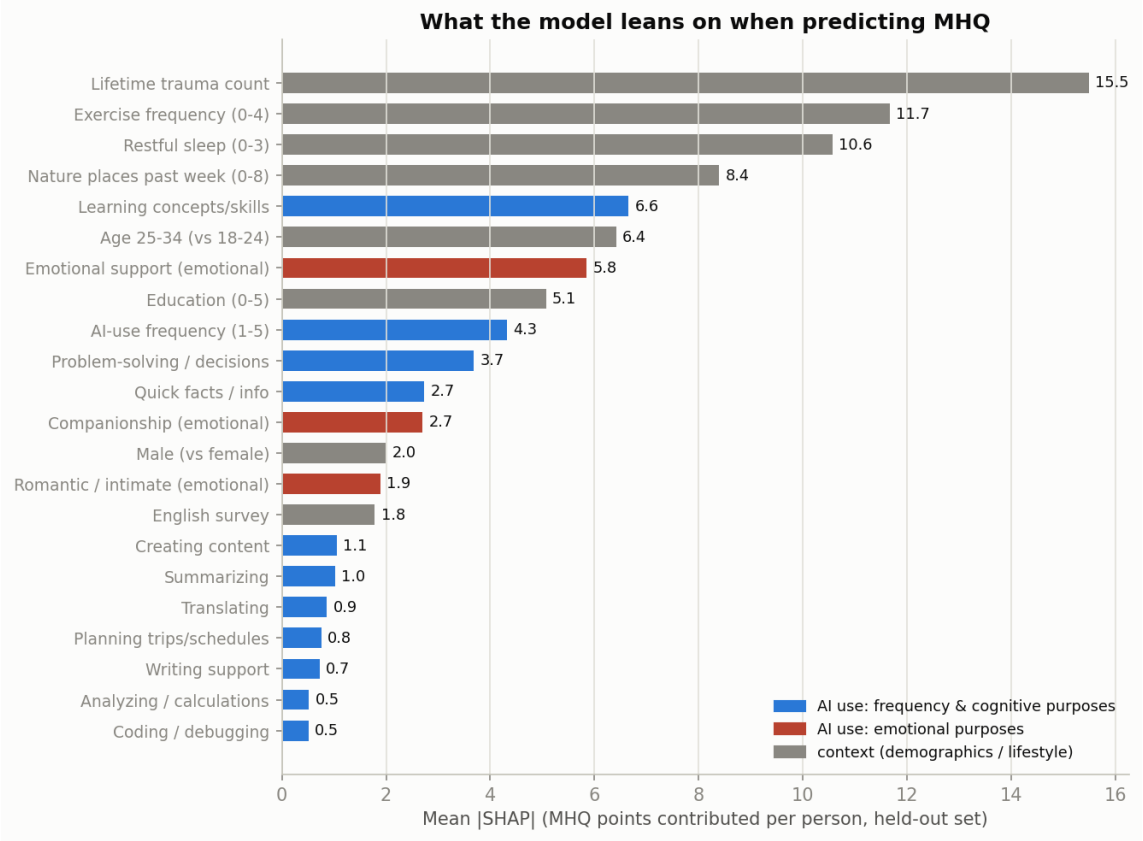


Figure S3. Mean absolute SHAP contribution per feature (held-out set). Grey: context; blue: AI-use frequency and cognitive purposes; red: emotional purposes. Features are ranked by this mean absolute contribution.

H. Direction and spread of contributions

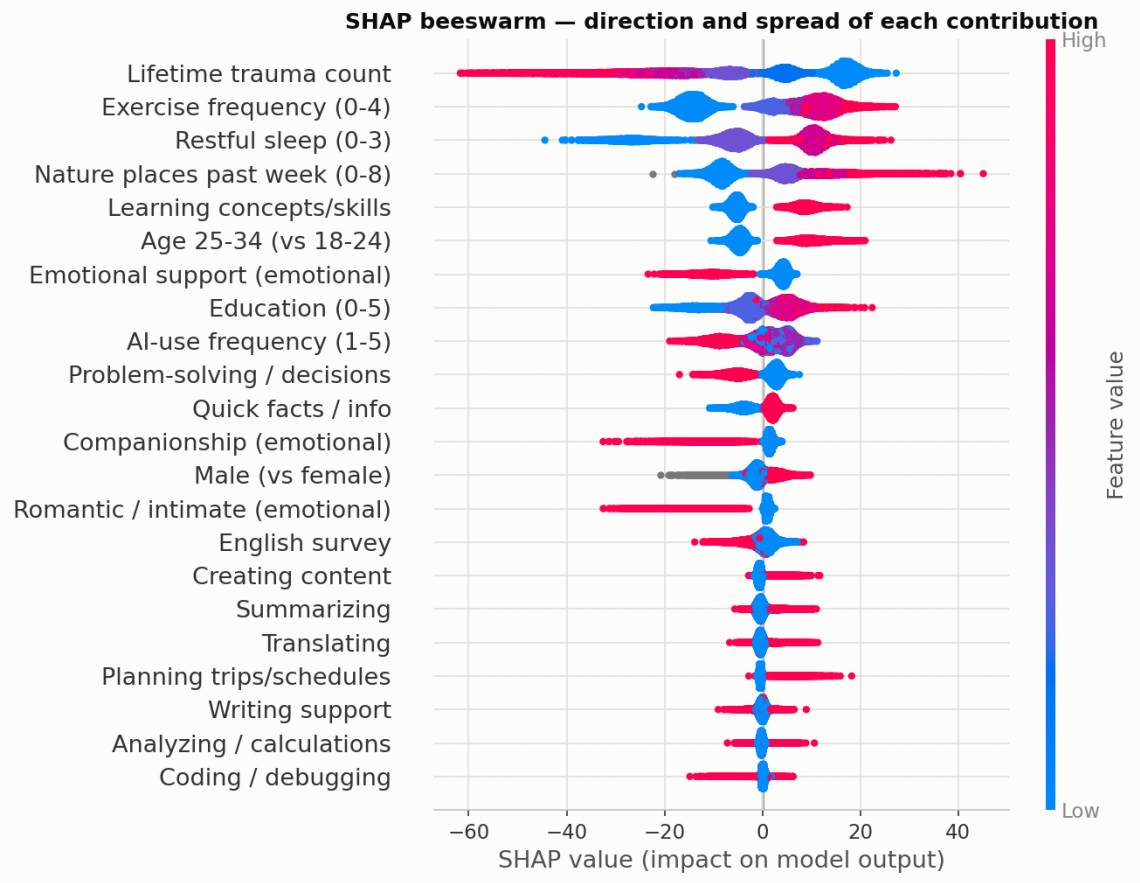


Figure S4. Beeswarm of per-respondent SHAP values; each dot is one respondent, colored by their feature value (red high, blue low). High trauma pulls predictions down and good sleep up; reporting an emotional use pulls predictions down with tails reaching −20 to −30 points. Frequency's scrambled color pattern (mid values right, top value far left) is the signature of its non-monotone, inverted-U effect.

Table S16. Average SHAP contribution of each purpose of use among respondents who report it (held-out set). The ordering matches the descriptive Figure 6A; the emotional-use deficits of 28-30 raw points attenuate to 11-14 attributed points once the model accounts for trauma, sleep, exercise, time in nature and the other context, consistent with part of the raw gap running through who those users are.

Purpose of use	Share reporting it	Mean SHAP when reported
Romantic / intimate (emotional)	6.700	−14.18
Emotional support (emotional)	27.200	−10.88
Companionship (emotional)	11.700	−10.66
Problem-solving / decisions	36.000	−5.93
Coding / debugging	8.400	−0.95
Writing support	23.200	0.60
Translating	19.000	1.89
Analyzing / calculations	13.000	2.22
Quick facts / info	65.200	2.50
Summarizing	22.900	3.02
Creating content	14.800	4.15
Planning trips/schedules	6.400	10.63

Learning concepts/skills	37.100	10.84

I. The dose response by type of use

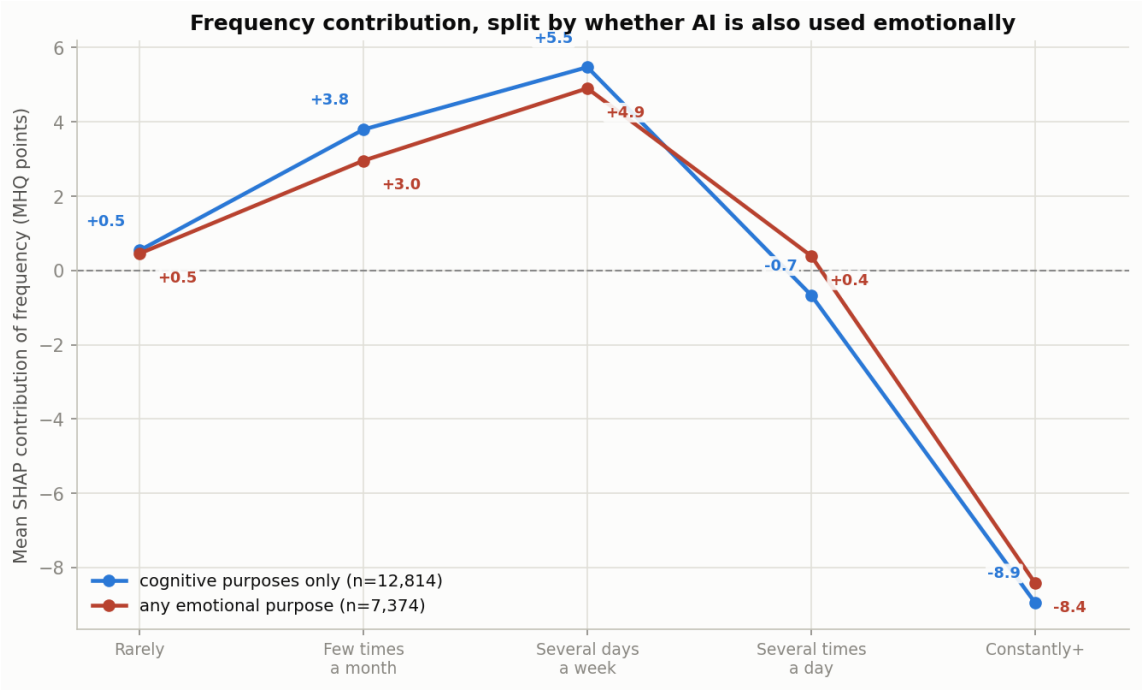


Figure S5. Mean SHAP contribution of use frequency, computed separately for respondents with and without any emotional use. Both groups trace the same rise-then-fall; emotional use sets the level through its own contributions while frequency produces the same dose pattern in both, the multivariate counterpart of Figure 5B.

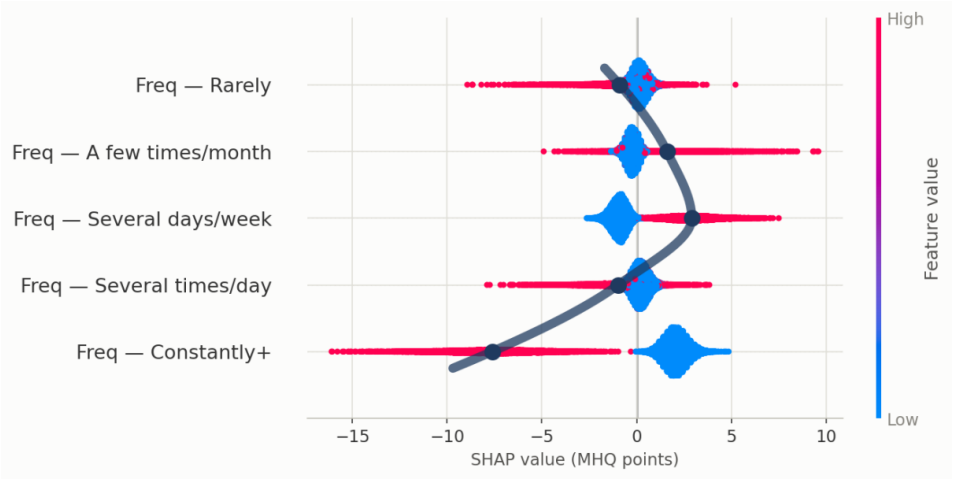


Figure S6. Beeswarm of per-respondent SHAP contributions with AI-use frequency represented as five binary features, one per use-frequency level. The curve through the five-level means reproduces the inverted U of Figure 8.