UCSF
UC San Francisco Previously Published Works

Title

Evaluating brief digital interventions for university student mental health in mainland China: a randomized trial

Permalink

https://escholarship.org/uc/item/65v708hg

Journal

npj Mental Health Research

ISSN

2731-4251

Authors

Schueller, Stephen M
Liu, Ran
Lee, Michael D
et al.

Publication Date

2026-09-24

DOI

10.1038/s44184-026-00247-w

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

npj | mental health research

Article

Article in Press

https://doi.org/10.1038/s44184-026-00247-w

Evaluating brief digital interventions for university student mental health in mainland China: a randomized trial

Check for updates

Stephen M. Schueller[1,2,10,11] ✉, Ran Liu[3,10], Michael D. Lee[4], Emily Baum[5], Michelle Hue-Nam Chan[1], Chuansheng Chen[1], Y. Anthony Chen[1], Tannei Hong[6], Tianjiao Kong[3,7], Yufei Li[4], Elizabeth A. Martin[1], Peiyi Wang[1], Zexin Wen[6], Novia Wong[2], Jun Wu[8], Wailam Ng[6], Yumeng Yang[2,7], Zhaoxia Yu[9], Wenqi Zhang[1], Yanzhen Zhang[9] & Liang Luo[3,7,11] ✉

Mental health disorders are prevalent and burdensome, especially among university students. The need for mental health services, however, outstrips the supply, necessitating the exploration of new models of treatment. We investigated several brief and scalable digital mental health interventions among university students in mainland China. These interventions included two different single-session interventions and brief videos based on dialectical behavior therapy. We tested these interventions in a large trial (modified intention-to-treat analysis sample, n = 1872) that combined randomized and participant preference arms. Participants were randomized into either a randomized arm with five conditions (four intervention conditions plus waitlist control) or a participant preference arm where participants received choice for two conditions. Primary outcome measures were based on the hierarchical taxonomy of psychopathology. Using hierarchical Bayesian analysis, we found strong support that the paid DBT video condition had a small impact, but no strong evidence for the impact of other conditions. This study demonstrates that brief interventions can have reliable, but small, impacts. However, not all interventions were beneficial. Registry: Chinese Clinical Trial Registry, TRN: ChiCTR2400084558, Registration date: 20 May 2024.

Mental disorders are prevalent and burdensome health issues around the world. In China specifically, the prevalence of any mental health condition was estimated at nearly 10%, with the most common disorders being anxiety and depressive disorders at roughly 4% each[1]. Despite increases in mental health services and the mental health workforce in recent years, China still suffers a shortage in well-trained mental health professionals[2]. Although a shortage of mental health professionals is not unique to China, various factors complicate the ability to address these shortages. These include China's vast geographic size and its large, widely dispersed population, a history of mental health stigma, and negative attitudes towards the professions of psychology and psychiatry. Even those providing mental health services in China, often lack formal mental health training, which presents a challenge for providing evidence-based interventions[3]. Tackling the burden of mental disorders in China will require innovative solutions that overcome these factors.

One area of innovation is the development of digital mental health interventions (DMHIs). These interventions use internet websites and mobile applications to provide psychosocial interventions to promote affective, behavioral, and/or cognitive change. DMHIs have been developed for various mental health conditions, with considerable evidence for their effectiveness at addressing depression[4], anxiety[5], posttraumatic stress disorder[6], sleep difficulties[7], and emerging evidence in other conditions[8–10]. DMHIs can be provided at scale, with the interventions retaining their therapeutic potential even with large populations using them[11,12]. As such, health systems and public entities in many countries are using DMHIs as front-line treatments to address common mental health issues, including

[1]Department of Psychology, University of California, Irvine, Irvine, CA, USA. [2]Department of Informatics, University of California, Irvine, Irvine, CA, USA. [3]Institute of Developmental Psychology, Faculty of Psychology, Beijing Normal University, Beijing, China. [4]Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA. [5]Department of History, University of California, Irvine, Irvine, CA, USA. [6]Faculty of Psychology, Beijing Normal University, Beijing, China. [7]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China. [8]Department of Environmental & Occupational Health, University of California, Irvine, Irvine, CA, USA. [9]Department of Statistics, University of California, Irvine, Irvine, CA, USA. [10]These authors contributed equally: Stephen M. Schueller, Ran Liu. [11]These authors jointly supervised this work: Stephen M. Schueller, Liang Luo. ✉e-mail: s.schueller@uci.edu; luoliang@bnu.edu.cn

human-supported and self-guided programs[13–16]. A recent review identified 28 published randomized controlled trials conducted in China as of 2024[17]. Consistent with broader research on DMHIs, these trials consistently found positive impacts of DMHIs. Effect sizes varied, however, from small to large across individual trials. This review noted that most DMHIs were focused on adults, and further work is needed to be focused on reaching specific populations and tailoring those interventions for the target population.

Given the unique stressors posed by academic pressure and rapid social changes in China[18], university students in China may be an especially important target for mental health interventions, especially DMHIs. Most mental disorders first occur between adolescence and early adulthood[19]; therefore, by the time individuals reach university, most who will experience a mental disorder have already done so. Many universities also conduct screening or surveys that could help identify those most in need[20]. As such, universities provide an opportunity to intervene with students through targeted and accessible resources aimed at addressing diverse mental health challenges. Chinese university students often experience overlapping symptoms across various mental health conditions, such as anxiety, depression, and stress. These diverse needs call for an approach that captures a comprehensive range of psychopathological symptoms rather than relying on discrete diagnostic categories that might fail to fully reflect the complex psychological experiences of these students. Lastly, addressing mental disorders during the university years can help people on the important path to recovery, helping them succeed at forming relationships, finding jobs, and transitioning to the next stage of their lives.

Several reviews synthesizing results of DMHIs among university students have found strong evidence of effectiveness, particularly for depression and anxiety, with effect sizes consistently reported in the small to moderate range[21–23]. Increasingly, researchers are emphasizing the need to consider implementation-related aspects, such scalability and long-term use, early in the developmental and evaluation of interventions[24], especially when it comes to DMHI[25,26]. Brief, self-guided interventions, likes ones examined in the present study, have strong opportunities to reach university students at scale, and in the long run, if effective, could have a meaningful impact on public mental health. University students are often balancing various responsibilities, including coursework, extracurricular activities, work schedules, and family and social relationships that might make accessing traditional services more challenging. Several studies have identified that convenience is a major factor motivating university students' selection of mental health services and might prefer DMHIs, especially brief ones that can fit into the fabric of their busy lives[27,28].

Various brief DMHIs have demonstrated effectiveness, with considerable interest in recent years in single-session interventions (SSIs). An SSI is designed to provide a complete psychological intervention in one sitting, prioritizing each engagement with an individual rather than assuming that people will return for a subsequent session[29]. Several SSIs have been developed and validated for a variety of mental health targets, including anxiety, depression, externalizing problems, eating pathology, and substance use[30]. Many recently developed SSIs have been delivered in digital formats to increase their ability to scale and create a package easily accessible to diverse populations[31]. SSIs were selected for this study given their considerable evidence base. A recent umbrella review from Schleider and colleagues[30] summarized 24 systematic reviews that included 415 unique trials. The results of these reviews overwhelmingly (83.3%) found significant and positive outcomes for SSIs.

Other types of brief DMHIs have also been developed, this includes very brief interventions that use different types of media presentation. For example, one study explored brief dialectical behavior therapy (DBT) skills videos by sending a video each day to university students along with brief ecological momentary assessments four times a day[32]. The videos had high compliance rates, significant immediate reductions in negative affect and increases in positive affect, and, over time, helped support the tolerance of emotions, a key aspect of DBT. However, long-term impacts on negative affect and self-efficacy were not significant. Nevertheless, such an approach aligns with one way university students often receive content, in brief videos online, and could represent a highly scalable intervention. Another benefit of these DBT videos is that they focus on concepts that drive towards concrete behavioral skills. This type of concrete and focused content could be particularly useful for dissemination in a brief intervention. More research is needed to explore the use of these types of treatments to understand better which ones might has long-term impact at scale, and various ways to deliver them to increase their benefits.

Traditional diagnostic taxonomies have provided a common language for researchers, clinicians, and policymakers, advancing research and practice in mental health and its treatment and prevention. As noted earlier, the mental health of university students in China is diverse and complex, requiring assessments that can appropriately capture it. The Hierarchical Taxonomy of Psychopathology (HiTOP) was developed to address several limitations of other approaches, including heterogeneity within diagnoses, the use of arbitrary discrete boundaries categories, and the low reliabilities present in several diagnoses[33]. The Hierarchical Taxonomy of Psychopathology-Self Report measure (HiTOP-SR)[34] is a newly developed assessment tool that provides a dimensional measure to assess psychopathology across a broad range of experiences. This allows for a more complete snapshot of an individual's area or areas of concern. Thus, having the benefit of being able to capture potential change across a wide range of experiences rather than a single aspect captured by traditional measures.

This study investigated the use of several brief DMHIs for university students in mainland China. Our approach drew from diverse models of brief previously validated DMHIs, including SSIs and DBT videos with adaptations to our target population. We conducted a large trial combining both a randomized controlled trial (RCT) and participant preference design to allow comparisons between various types of interventions while maximizing the resources we had available to conduct this work. We examined which interventions improved mental health symptoms using the HiTOP measure, capturing a broad range of psychopathology.

Methods

Overview

In this study, we compared the effectiveness of several types of brief DMHIs, including brief DBT videos, two different SSIs, and a placebo control SSI. We also included a waitlist control condition. We conducted an RCT with an embedded participant preference arm. A target sample of 2000 participants was determined based on study logistics of funding availability, with a target to recruit 125 participants in each active condition. Participants were first randomized into the randomized arms with incentives for completion of study measures or a participant preference trial with no such incentives. Participants in the randomized arm were randomized at a 1:1:1:1:4 ratio to either one of four intervention conditions (brief DBT videos, two orderings of SSIs, and a placebo control) or a waitlist control condition. We used an unequal allocation, preferencing the waitlist comparison condition, to allow for a robust estimate of the effectiveness of interventions following our second hypothesis (described below). Thus, we had an equal number of participants who received an active condition compared to the waitlist control. The randomization allocation sequence was created using a computer-based program by a member of our research team to create a sequential assignment to study conditions. Following the generation of this sequence, the finalized allocation lists for each study condition were provided to the corresponding intervention researchers, who contacted participants assigned to group. The intervention researchers received only the finalized group assignments and were not involved in generating the allocation sequence or able to modify participant assignments. Participants were not blinded to condition, as they would know if they were assigned to a random assignment to conditions or participant preference or to waitlist control or active conditions. Participants randomized to the participant preference arm had their choice of the two orderings of the SSIs. All participants completed a baseline assessment battery, received their interventions over a six-week period, and then completed a follow-up assessment battery. We selected a six-week period in order to assess longer-term change rather than immediate reactions to the intervention conditions and because

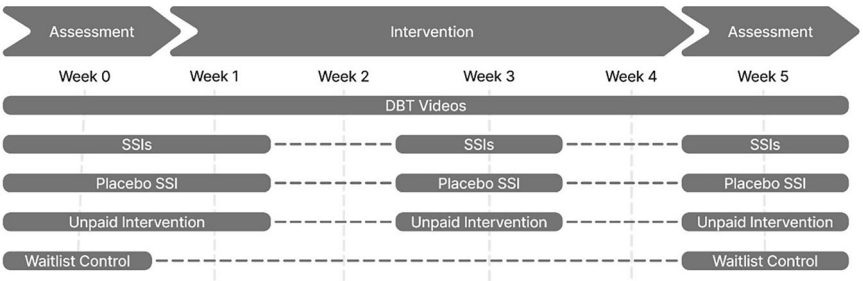


Fig. 1 | Study timeline illustrating timing of assessments and interventions. DBT dialectical behavior therapy, SSI single-session intervention. SSIs were provided in two different orders, referred to as Order 1 and 2 in the text. The placebo SSI was the Sharing Feelings Project. Assessment at Week 0 represents the baseline assessment. Assessment at Week 5 represents the endpoint assessment.

6 weeks is a common assessment period in trials DMHIs[35]. Compensation was provided for completion of the baseline assessment battery and all follow-up measures for those in the randomized arm. The participant preference arm was included, as we did not have sufficient funds to provide compensation to all individuals who completed the baseline assessment battery to complete follow-up assessments. Instead, we offered them the option of receiving the interventions and completing follow-up assessments if they chose. Participants were primarily contacted by email and text messages for follow-up assessments, including reminders to complete assessments and intervention activities within the designated time frame. For a small number of participants who did not respond promptly to emails or text messages, individualized follow-up was conducted through WeChat. The study was reviewed and approved by the Institutional Review Board at Beijing Normal University, all research was performed in accordance with the Declaration of Helsinki, and was prospectively registered on the Chinese Clinical Trial Registry: ChiCTR2400084558 on 20 May 2024. The clinical trial registry represents the anticipated sample size and randomized conditions. Our actual sample sizes for our analyzed sample are slightly smaller due to participants who failed attention checks. Our full analysis plan and hypotheses were preregistered (https://osf.io/qybtv/). There was no patient or public involvement in the design, conduct, or reporting of this study.

Interventions

Our DBT video condition used videos provided via Bilibili from the DBT-RU YouTube channel that were previously evaluated[32]. These videos include various skills and content based on DBT and are available in English, Spanish, and Mandarin. The previous evaluation and availability of Mandarin-language content were both contributing factors to our use of these videos in our study. We selected eight Mandarin videos that focused on emotion regulation and distress tolerance skills. Videos for each week were selected to go together, including mindfulness (HOW and WHAT skills), WISE mind and self-soothing, PLEASE and IMPROVE skills, and ACCEPTS and TIPP skills. Each participant received every video presented in the same order. When the videos were sent to participants, we added a brief follow-up question to each video as a call to action to practice the skills presented in the video. For example, a video on the mindfulness "what" skills was followed by a prompt, "Think of a time this week when you could practice the observe skill of mindfulness. When will you practice the observe skill, and what will you be doing?" Two videos and accompanying questions were sent to participants each week. The shortest video lasted 3 min, 21 s and the longest video lasted 5 min, 56 s (M = 4 min, 24 s; SD = 1 min, 9 s).

Our SSI conditions used SSIs developed by the Lab for Scalable Mental Health. Our two active SSIs were the ABC Project and Project CARE. We selected these two SSIs due to their considerable evidence base. The ABC Project is modeled on behavioral activation and teaches people to engage in value-based activities that lead to pleasure and meaning[36,37]. Project CARE supports self-compassion and self-guidance through providing solutions for overcoming self-doubt and activities to practice compassion towards others to promote self-kindness[37,38]. Behavioral activation is an effective element across several mental health interventions, including DMHIs[39–41]. We thought self-compassion might be particularly appropriate for the Chinese cultural context. Our placebo control SSI was the Sharing Feelings Project. The Sharing Feelings Project is modeled on supportive therapy and encourages people to share feelings with those close to them, but does not teach specific skills[42,43]. We linguistically and culturally adapted the interventions from existing materials to be relevant for Chinese university students. Adaptations included translating English to Chinese, adjusting examples that were initially intended for teens to be applicable to university students, and changing those that might not be as relevant in a Chinese context to be more culturally appropriate[44]. For instance, while sports-related examples are commonly used in the United States due to the widespread interest in athletics, the Chinese version incorporated examples related to academic competition, which represents a more salient source of pressure for Chinese students than sports performance. The adaptation process was conducted by members of our research team, including both our US- and China-based team members, who collectively had expertise in cross-cultural research, DMHI intervention design, and were bilingual speakers of Mandarin and English. Minor revisions were made to the translated materials to improve clarity, readability, and cultural appropriateness while preserving the original meaning and content. These revisions primarily involved adjustments to wording and phrasing to ensure that the materials were suitable for the Chinese cultural context. Many of the team members who participated in the adaptation process had attended university in China, and several were master's and doctoral students in psychology at Beijing Normal University with training and research experience in mental health. These interventions are text-based and self-guided through Qualtrics. Each SSI was designed to be approximately the same amount of content and length of engagement. All of our intervention materials are available on our OSF page (https://osf.io/qybtv/).

Due to logistical constraints of our funding, we only had enough compensation for half of our sample to complete additional follow-up assessments related to the interventions. As such, we designed our study to have ~1000 participants complete follow-up assessments, with half assigned to a waitlist and half assigned to active conditions. We offered other participants the option to complete interventions in a participant preference arm where assessments were offered without compensation. We excluded the DBT videos from the participant preference arm as we had concerns over whether participants would have sufficient information to make decisions between different types of interventions, and we did not propose any a priori comparisons within this arm. Given the vast evidence supporting SSIs[29], we decided to use those interventions over the DBT videos. The participant preference arm only included the two orderings of the SSIs. Therefore, potential comparison conditions were paid SSIs Order 1 (ABC then CARE), paid SSIs Order 2 (CARE then ABC), paid DBT videos, paid placebo SSI (Sharing Feelings), unpaid SSIs Order 1 (ABC then CARE), unpaid SSIs Order 2, or waitlist control. Note that the unpaid conditions still received compensation to complete the baseline assessment battery. A figure illustrating the timeline of the assessments and interventions is provided in Fig. 1.

Recruitment

We used various methods of recruitment via online posts (WeChat, QQ groups, and WeChat Moments) at four Chinese universities located in Beijing, Shaanxi Province, Jilin Province, and Shandong Province, China. Recruitment materials explained that eligibility required being an undergraduate student and that participants would receive notifications through email, text message, or WeChat. Participants were informed that participation would last six weeks, and they would receive RMB 140 to RMB 380 (roughly $20–50 USD) for compensation based on completion of study assessments. All participants needed to be 18 years or older to participate. All participants provided informed consent electronically prior to completing the baseline measure. Recruitment and study completion happened over two phases, with one set of participants ($n = 288$) completing their baseline and participation between May 27 to July 5, 2024 and the remainder of the participants completing their baseline and participation between July 22 to August 31, 2024.

Measures

Participants completed baseline and endpoint assessments consisting of demographics, background information, and the translated Chinese-version Hierarchical Taxonomy of Psychopathology-Self Report measure (HiTOP-SR). The HiTOP-SR[34] is a dimensional measure assessing a broad range of psychopathology organized into five spectra, including detachment, externalizing, internalizing, somatoform, and thought disorder symptoms ($N = 405$ items). Participants are asked to indicate on a 1 (not at all) to 4 (a lot) scale the extent to which each item applied to them over the past year. Because of the focus of the interventions used in the current project, only detachment, internalizing, and somatoform items were administered. We dropped internalizing items related to sexual problems due to concerns with appropriateness for the study population [e.g., ref. 44] and two items due to low factor loadings that had a disproportionate impact on internal consistency. One removed item was from the detachment-submissiveness subscale ("I allowed people to take advantage of me") and one from the internalizing-affective lability subscale ("I quickly went from being sad to being happy"). The resultant measure for this study was 213 items. This version of the HiTOP-SR showed good internal consistency as indicated by Cronbach's alphas between 0.67–0.95 for each spectrum (0.71–0.86 for detachment, 0.67–0.95 for internalizing, 0.74–0.86 for somatoform). Given the population and interventions used, we did not have any systematic assessment of harms or side effects.

Three attention check items were included in the baseline and follow-up assessments. These attention checks included one information question (identify the capital city of China) and two questions that asked participants to select specific response options ("Rarely" and "Never," respectively). All three attention check items were randomized in their presentation and therefore appeared at different positions throughout the survey. Participants were considered to fail the attention check if they answered any one of the three items incorrectly.

One reviewer raised concerns that the response period for the HiTOP measure, asking participants to indicate the extent to which each applied to them over the past year might not be sensitive to change during our intervention period. We had also collected the DSM-5 Cross-Cutting Symptom measure, which is a 23-item instrument that assesses mental health across 13 domains that are important across psychiatric diagnoses[45]. Participants are asked to indicate on a 0 (not at all) to 4 (severe, nearly every day) how much or how often they had been bothered by each problem during the past 2 weeks. The HiTOP measure was our preregistered primary outcome, so we focus on those results, but include a comparison with results using the DSM-5 Cross-Cutting Symptom measure, in the supplement. The DSM-5 Cross-Cutting Symptom analysis showed results consistent with the primary HiTOP analysis. Additional measures not used in these analyses are included in the OSF pre-registration methods.

Hypotheses

We predicted that: (1) the interventions will have a significant effect (the difference for paid SSIs, paid DBT videos, and unpaid SSIs, will be greater than 0). Addressed under preregistered hypothesis testing of directional alternatives (below), (2) interventions will be better than no treatment (the difference for paid SSIs, paid DBT videos, and unpaid SSIs will be greater than the difference for the waitlist control). Addressed under preregistered evaluation of preregistered orders (below), (3) interventions will be better than the placebo (the difference for paid SSIs, paid DBT videos, and unpaid SSIs will be greater than the difference for the placebo group receiving the Sharing Feelings Project). Addressed under preregistered evaluation of preregistered orders (below). All other comparisons between groups are exploratory.

Analysis plan

We used hierarchical Bayesian analysis to make inferences about the changes from baseline to endpoint responses for each of the conditions[46]. The hierarchical modeling approach allows us to measure the mean treatment effect of each condition, while allowing for variability in these effects at the level of each HiTOP spectra and their subscales, as well as for different individual participants. The Bayesian statistical approach allows us to quantify the uncertainty inherent in the inferences and to handle missing data in a model-based way.

The data take the form of responses $y^{pr}_{p,j} = 1, 2, 3, \ldots, 4$ and $y^{po}_{p,j} = 1, 2, 3, \ldots, 4$ for participant p answering item j in the baseline to endpoint, respectively. We model the change from the baseline to endpoint score, for every participant on every item, in terms of an effect δ_{p,s_j} of the treatment that is specific to the participant and the HiTOP scale to which the item belongs:

$$y^{po}_{p,j}\ Gaussian\left(y^{pr}_{p,j} + \delta_{p,s_j} \times \sigma, \frac{1}{\sigma^2}\right),$$

Equation 1. Modeling of change used to estimate treatment effects where σ is the standard deviation of the difference between the baseline and endpoint scores, which places δ_{p,s_j} on the scale of effect sizes.

Treatment effects are assumed to follow the hierarchical taxonomic structure of the HiTOP. At the topmost level is an overall effect of the treatment. At the second level, the effect of treatments on each spectrum are centered on the overall treatment effect, but with variability across spectra. At the third level, the effects of treatment on each scale are centered on the overall treatment for the relevant spectra effect, but with variability across subscales. Finally, each individual's treatment effect for an item comes from the distribution of treatment effects for the subscale to which the item belongs and the condition to which the participant was assigned. The formal statistical definition of the analysis model is provided in supplementary material, along with its JAGS code implementation[47].

We used the following statistical approaches to determine results. These approaches were preregistered on our OSF page.

Robustness checks

We considered three different prior assumptions about the overall mean effect of each treatment. The prior distribution for treatment t was defined as $\mu^{tr}_t\,\text{Gaussian}\left(0, \frac{1}{s^2}\right)$, where the constant s specifies the size of the effect that is expected. We used values of $s = 0.1$, $s = 0.2$, and $s = 0.4$, which specify small to moderate effects. These effect size estimates are based on past literature of the effect sizes found in brief digital interventions like the type we are using in this study[32,37,43].

As additional robustness checks, we considered both the nondirectional effect size priors specified in the JAGS code, for exploratory purposes, and the directional hypotheses specified above. We also tested the robustness of results to the priors on the standard deviations for the hierarchy of treatment effects. Given the number of participants in each treatment group, we expected these inferences to be robust.

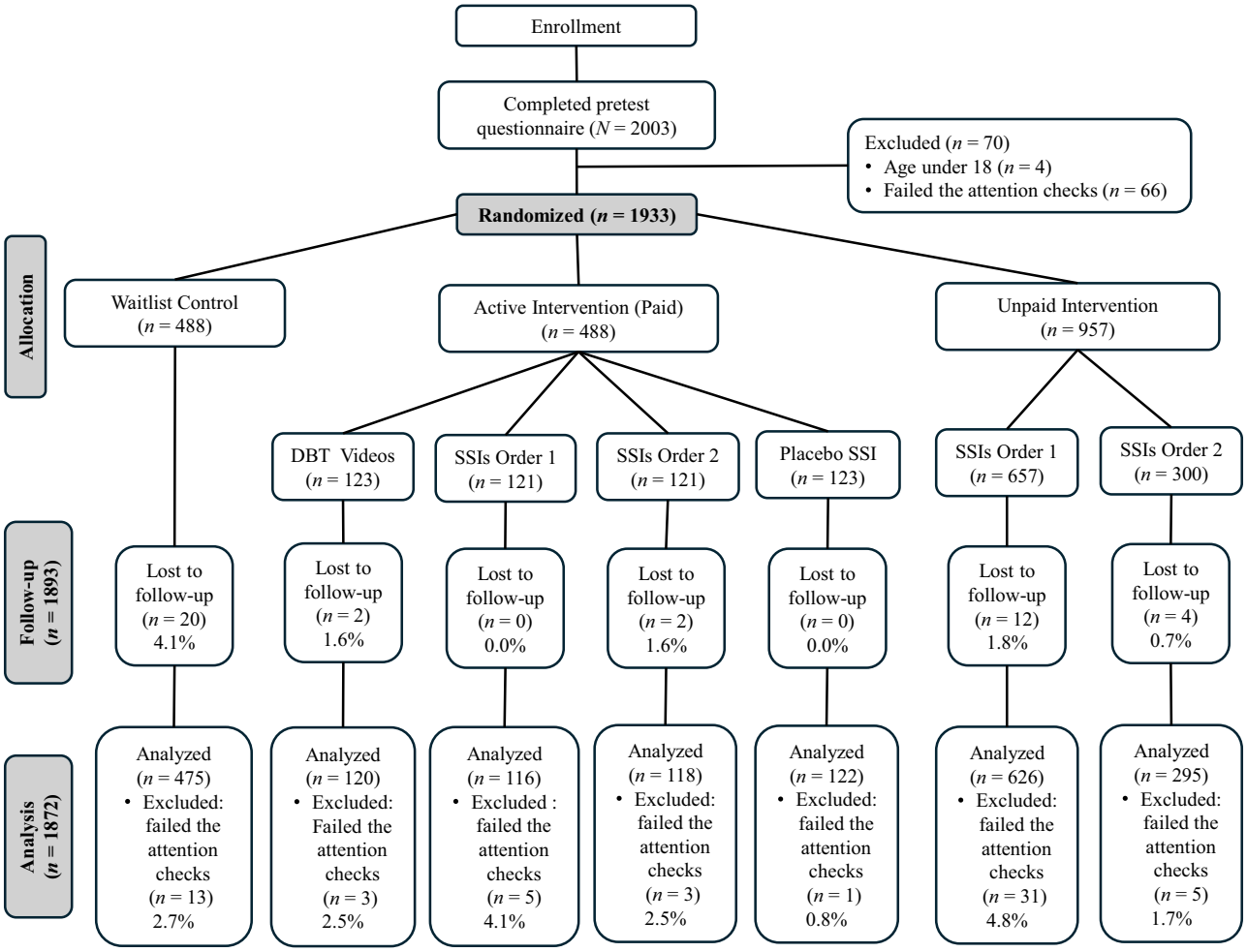


Fig. 2 | CONSORT diagram. DBT dialectical behavior therapy, SSI single-session intervention. SSI Order 1 was the ABC Project, followed by Project CARE, SSI Order 2 was Project CARE, followed by the ABC Project.

Comparison of preregistered hypotheses

We used the Savage–Dickey method[48] to approximate Bayes factors comparing null hypotheses against the (possibly directional) alternative hypotheses. The posteriors were approximated by binning posterior samples. The priors are defined by the hierarchical model for treatment, factor and scale effects. We did not use ROPE or similar procedures that use credible intervals to test hypotheses—for example, testing whether effect sizes of zero are contained in the intervals—because these procedures violate the principle of predictive irrelevance[49,50].

Evaluation of preregistered orders

We tested the ordering of treatments, factors, and scales by estimating the posterior probability of orders using posterior samples.

Evaluation of Effect Sizes

We visualized the marginal posterior distributions to examine effect sizes. Summaries of these marginals are based on the posterior means and 95% credible intervals. The credible intervals are based on the 2.5 and 97.5% percentiles, not highest density intervals, because highest density intervals are not reparameterization invariant.

Handling of missing data

We conducted analyses according to the modified intent-to-treat (mITT) principle. All randomized participants who provided valid baseline data were eligible for inclusion; however, participants who failed one or more attention checks at the endpoint assessment were excluded from the primary analysis ($n = 61$). Thus, unlike a conventional ITT analysis, the mITT analysis did not retain participants whose endpoint responses failed our prespecified data-quality criterion. We excluded these participants because we had reduced confidence in the validity of the endpoint data and, therefore, in conclusions based on their data. Bayesian methods automatically make model-based imputation for other missing data. Our preregistration specified conducting a completers analysis (see Fig. S1 for CONSORT diagram of completers analysis), which we also report in supplementary material and showed an identical pattern of findings.

Transparency and openness

The primary analysis was implemented using JAGS[47] to define a graphical model, which allows for fully Bayesian analysis based on computational sampling of the joint posterior distribution. The JAGS implementation model used is available on our OSF page, along with the de-identified analyzed data. MATLAB version 24.2[51] was used to run the JAGS 4.3.2 model as well as all accompanying analyses (plotting, summarizing inferences, etc.).

Results

Participants

Out of the 2003 individuals who completed our baseline questionnaire, 1933 met our inclusion criteria. Consistent with our mITT approach described above, the primary analysis sample included 1872 participants; 61 randomized participants failed were excluded because they failed our attention checks (see CONSORT Diagram, Fig. 2). Completion rates were

Table 1 | Baseline demographic and clinical characteristics

Baseline Characteristics Mean (SD) or *n* (%)	Overall (*n* = 1872)	Waitlist Control (*n* = 475)	Active Intervention (Paid) (*n* = 476)				Unpaid Intervention (*n* = 921)	
			DBT Videos (*n* = 120)	SSIs Order 1 (*n* = 116)	SSIs Order 2 (*n* = 118)	Placebo SSI (*n* = 122)	SSIs Order 1 (*n* = 626)	SSIs Order 2 (*n* = 295)
Age	20.48 (1.86)	20.40 (1.76)	20.80 (1.94)	20.90 (1.61)	20.42 (1.58)	20.19 (1.85)	20.60 (2.14)	20.18 (1.47)
Biological sex								
Female	1180 (63.0%)	298 (62.7%)	60 (50.0%)	71 (61.2%)	67 (56.8%)	86 (70.5%)	375 (59.9%)	223 (75.6%)
Male	692 (37.0%)	177 (37.3%)	60 (50.0%)	45 (38.8%)	51 (43.2%)	36 (29.5%)	251 (40.1%)	72 (24.4%)
School year								
Freshmen	454 (24.3%)	141 (29.7%)	23 (19.2%)	21 (18.1%)	31 (26.3%)	38 (31.1%)	130 (20.8%)	70 (23.7%)
Sophomore	566 (30.2%)	131 (27.6%)	39 (32.5%)	34 (29.3%)	37 (31.4%)	42 (34.4%)	193 (30.8%)	90 (30.5%)
Junior	471 (25.2%)	124 (26.1%)	28 (23.3%)	22 (19.0%)	21 (17.8%)	30 (24.6%)	164 (26.2%)	82 (27.8%)
Senior	220 (11.8%)	45 (9.5%)	10 (8.3%)	27 (23.3%)	19 (16.1%)	7 (5.7%)	81 (12.9%)	31 (10.5%)
Beyond senior year	161 (8.6%)	34 (7.2%)	20 (16.7%)	12 (10.3%)	10 (8.5%)	5 (4.1%)	58 (9.3%)	22 (7.5%)
Ethnicity								
Han	1750 (93.5%)	439 (92.4%)	110 (91.7%)	111 (95.7%)	113 (95.8%)	108 (88.5%)	593 (94.7%)	276 (93.6%)
Minority	122 (6.5%)	36 (7.6%)	10 (8.3%)	5 (4.3%)	5 (4.2%)	14 (11.5%)	33 (5.3%)	19 (6.4%)
Parental marital status								
Married	1652 (88.2%)	428 (90.1%)	107 (89.2%)	108 (93.1%)	105 (89.0%)	103 (84.4%)	556 (88.8%)	245 (83.1%)
Divorced	113 (6.0%)	20 (4.2%)	7 (5.8%)	5 (4.3%)	4 (3.4%)	12 (9.8%)	38 (6.1%)	27 (9.2%)
Re-married	43 (2.3%)	12 (2.5%)	1 (0.8%)	0 (0%)	4 (3.4%)	2 (1.6%)	14 (2.2%)	10 (3.4%)
One or both parents passed away	64 (3.4%)	15 (3.2%)	5 (4.2%)	3 (2.6%)	5 (4.2%)	5 (4.1%)	18 (2.9%)	13 (4.4%)
Parents' residence								
Urban	1155 (61.7%)	290 (61.1%)	70 (58.3%)	71 (61.2%)	73 (61.9%)	77 (63.1%)	378 (60.4%)	196 (66.4%)
Suburban	129 (6.9%)	26 (5.5%)	10 (8.3%)	12 (10.3%)	5 (4.2%)	11 (9.0%)	42 (6.7%)	23 (7.8%)
Rural	588 (31.4%)	159 (33.5%)	40 (33.3%)	33 (28.4%)	40 (33.9%)	34 (27.9%)	206 (32.9%)	76 (25.8%)
Sought help for mental health								
Yes	520 (27.8%)	142 (29.9%)	27 (22.5%)	23 (19.8%)	34 (28.8%)	31 (25.4%)	164 (26.2%)	99 (33.6%)
No	1352 (72.2%)	333 (70.1%)	93 (77.5%)	93 (80.2%)	84 (71.2%)	91 (74.6%)	462 (73.8%)	196 (66.4%)
Currently receiving help for mental health								
Yes	68 (3.6%)	23 (4.8%)	5 (4.2%)	5 (4.3%)	7 (5.9%)	4 (3.3%)	18 (2.9%)	6 (2.0%)
No	1804 (96.4%)	452 (95.2%)	115 (95.8%)	111 (95.7%)	111 (94.1%)	118 (96.7%)	608 (97.1%)	289 (98.0%)
HiTOP-SR								
Detachment	2.35 (0.58)	2.32 (0.59)	2.26 (0.55)	2.39 (0.56)	2.38 (0.54)	2.42 (0.54)	2.31 (0.61)	2.45 (0.54)
Internalizing-distress	1.91 (0.54)	1.91 (0.55)	1.82 (0.55)	1.91 (0.57)	1.95 (0.54)	1.91 (0.52)	1.89 (0.54)	1.97 (0.51)
Internalizing-eating pathology	2.02 (0.39)	2.01 (0.40)	1.99 (0.41)	2.04 (0.40)	2.04 (0.38)	2.04 (0.41)	2.02 (0.40)	2.04 (0.37)
Internalizing-fear	2.07 (0.51)	2.05 (0.53)	1.99 (0.51)	2.14 (0.55)	2.09 (0.48)	2.06 (0.49)	2.06 (0.52)	2.09 (0.49)
Internalizing-mania	1.95 (0.51)	1.95 (0.49)	2.03 (0.54)	1.90 (0.53)	1.97 (0.48)	1.92 (0.51)	1.94 (0.51)	1.95 (0.49)
Somatoform	1.95 (0.51)	1.95 (0.51)	1.87 (0.53)	2.01 (0.56)	1.95 (0.47)	1.95 (0.49)	1.94 (0.50)	1.98 (0.50)

consistently high, ranging from 96% to 100% across conditions. Participants had a mean age of 20.48 years (SD = 1.86). The sample was predominantly female (63%), and most identified as Han Chinese (93.5%), the dominant ethnic group in China. All participants were enrolled in undergraduate studies, with the largest proportions in their second (30.2%) and third (25.2%) years. Most reported that their biological parents were married (88.2%), and the majority lived in urban areas (61.7%). Regarding parental education, 60.5% of mothers and 58.4% of fathers had completed high school or less. In terms of mental health history, 27.8% of participants reported previously seeking help for psychological concerns, 3.6% were currently receiving help, and 6.6% had taken psychiatric medication. Table 1 provides baseline demographic and clinical characteristics for the overall mITT analysis sample and each condition including values for the HiTOP-SR at the domain and sub-factor, values at the scale- and subscale-levels are available elsewhere[52]. We provided baseline values for the DSM-5 Cross-Cutting Symptom Measure in the supplemental file (Table S1). No harms were reported by participants throughout the study.

Overview of results

Figure 3 shows the inferred posterior distributions of the effect sizes for the seven conditions, based on the prior of $s = 0.2$. The posterior distributions represent the relative probability that each possible effect size is the true effect size for the condition. The posterior distributions are all unimodal, symmetric, close to Gaussian, and are well summarized by their means and

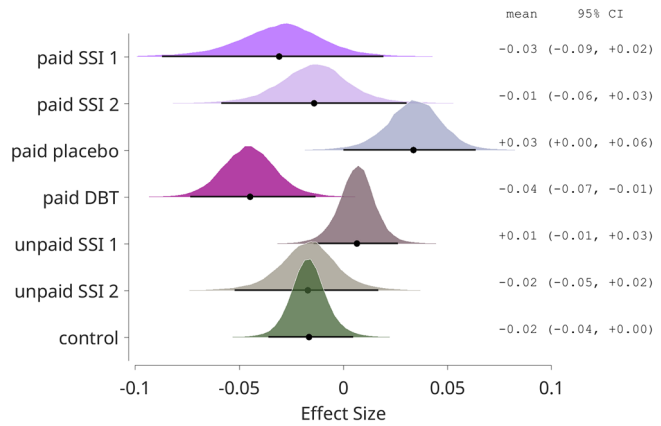


Fig. 3 | Posterior distributions of effect sizes for each of the seven conditions on the HiTOP measure. DBT dialectical behavior therapy, SSI single-session intervention. SSI1 SSI Order 1, i.e., the ABC Project, followed by Project CARE, SSI2 SSI Order 2, i.e., Project CARE followed by the ABC Project. CI credible intervals. Each condition corresponds to a different color distribution. The black dot represents the posterior mean, and the heavy black line at the bottom of each distribution represents the 95% credible intervals. The values for the posterior means and 95% credible intervals are displayed to the left of the distributions.

95% credible intervals. Figure S2 displays the same distributions for the DSM-5 cross-cutting symptom measure rather than the HiTOP measure. Table S4 provides a comparison of findings related to our hypotheses comparing the HiTOP and DSM-5 cross-cutting symptom measures.

The robustness checks with the alternative prior assumptions $s = 0.1$ and $s = 0.4$ for the effect size produced very similar posterior distributions, with almost identical means and credible intervals. This is a standard finding, because the posterior distributions are conditioned on the alternative hypothesis—that is, on there being an effect—and the information available in the data about the effect size overwhelms prior assumptions.

Prior assumptions about the effect size can, however, influence the results of hypothesis testing, because the evidence needed to support the existence of a small effect is different from the evidence needed to support the existence of a large effect. Accordingly, in testing the preregistered hypotheses, we considered the robustness of conclusions to prior assumptions. For all preregistered hypotheses reported below, we also provide the one- and two-sided Bayes Factors for both the mITT and completers analysis in the supplemental file (see Tables S2 and S3).

Hypothesis 1: The interventions will have a significant effect

The paid DBT videos condition showed the predicted effect, with a one-sided Bayes factor of $BF_{10} = 6.3$ in favor of the directional alternative hypothesis. This means that the data provide 6.3 times more evidence for the alternative hypothesis than the null hypothesis. In terms of the magnitude of the effect, the posterior mean is −0.04 and has a 95% credible interval of -0.07, -0.01. These magnitudes are smaller than even the smallest effect sizes around 0.1 we preregistered. The Bayes factors robustly favor the presence of an effect for our alternative prior assumptions about the size of treatment effects, with $BF_{10} = 12.2$ and $BF_{10} = 4.1$ for the smaller and larger effect size priors, respectively.

We found evidence suggesting that the paid SSI Order 2 (CARE, ABC), unpaid SSI Order 1 (ABC, CARE), and unpaid SSI Order 2 (CARE, ABC) conditions did not have an effect. The two-sided Bayes factors are $BF_{01} = 7.7$, $BF_{01} = 16.5$, and $BF_{01} = 5.9$ in favor of the null, respectively. Bayes factors robustly favor the null for the other prior assumptions about effect size. The evidence overall is that the paid SSI Order 1 (ABC, CARE) condition also did not have an effect, but it is less conclusive. The two-sided Bayes factor is $BF_{01} = 2.9$, which does not provide much evidence for either the presence or absence of an effect. This Bayes factor has a similarly non-decisive value, $BF_{01} = 1.3$, assuming small effect sizes, but increases to $BF_{01} = 5.0$ with a prior assuming large effect sizes. A reasonable conclusion is that there is no evidence for the presence of an effect, but also no compelling evidence that it is absent. If there was an effect, it is probably similar in magnitude to the DBT intervention, since the posterior mean is −0.03 with a 95% credible interval (−0.09, +0.02). Interestingly, there is more uncertainty about the possible size of an effect for the paid SSI Order 1 (ABC and CARE) condition than the paid DBT video condition.

It is unclear whether the paid placebo condition had an effect, but if there is an effect, it is almost certainly in the wrong direction (i.e., a small deterioration). The two-sided Bayes factor is $BF_{01} = 1.0$, the posterior mean is +0.03 with a 95% credible interval (0.00, +0.06).

There is weak evidence that the control condition did not have an effect. The two-sided Bayes factor is $BF_{01} = 3.6$. In terms of robustness, the Bayes factor is $BF_{01} = 1.8$, assuming smaller effects, which is non-decisive, but more definitive when assuming larger effects, with a $BF_{01} = 7.1$.

Hypothesis 2: The interventions will be better than no treatment

Given the results for Hypothesis 1, the relevant interventions are the paid DBT videos condition, for which there is evidence of an effect, and the paid SSI Order 1 (ABC, CARE) condition, for which there is no clear evidence of an effect. For the other conditions, the results for Hypothesis 1 found evidence for the absence of an effect, which makes comparison to our waitlist condition unnecessary. Thus, we compare the probabilities of effect orders just for the paid DBT videos and paid SSI Order 1 (ABC, CARE) conditions to the waitlist control. The implied Bayes factors provide evidence that there is no difference, with $BF_{01} = 3.5$ for the DBT condition and $BF_{01} = 10.4$ for the paid SSI Order 1 (ABC, CARE) condition.

Hypothesis 3. The interventions will be better than the placebo

As for Hypothesis 2, the relevant interventions are again the paid DBT videos and paid SSI Order 1 (ABC, CARE) conditions. The comparison is now to the paid placebo condition. The implied BF for the paid DBT videos condition was $BF_{10} = 6.6$, and for the paid SSI Order 1 (ABC, CARE) is $BF_{10} = 1.4$. Because these Bayes factors provide evidence of an effect or lack of evidence for no effect, we also calculated the probabilities of effect orders. These probabilities are 100 and 98% that the paid DBT videos and paid SSI Order 1 (ABC, CARE) conditions, respectively, had larger effects than the paid placebo condition. It is worth pointing out that these results depend not just on the possible presence of an effect for the interventions, but also from the possibility that the paid placebo had an effect in the wrong direction.

Discussion

This study evaluated various brief DMHIs among university students in China. It is also the first treatment study using the HiTOP framework as a primary outcome measure. Our results are bolstered by a large sample size and extremely high completion and high survey response rates, which are atypical among self-guided DMHIs [e.g., ref. 53]. Overall, we found significant evidence supporting the DBT videos condition, suggesting that even very brief interventions can lead to mental health benefits. However, the impact was small, although consistent with other studies of brief interventions for mental health[32,37,43]. For example, these studies reported effect sizes ranging from $d = 0.15$ to 0.31[43] and from 0.32 to 0.60 across a range of outcomes such as depressive symptoms, hopelessness, agency, and self-hate. Findings among our SSIs were more mixed, with no strong evidence for or against change in our paid conditions, but evidence suggesting no change in symptoms for one ordering in our unpaid condition. This seems inconsistent with previous evidence supporting the impact of SSIs. This study also contributes to a growing body of literature on DMHIs in non-Western countries and, to our knowledge, is the largest study of DMHIs in China and of digital SSIs outside of the United States.

Within brief interventions, it has been demonstrated that longer interventions may not necessarily be better [e.g., ref. 54], and our study added more support for this notion as the briefest DBT videos showed the most reliable effect. However, although the DBT video conditions were briefer than the SSIs in terms of each individual interaction, they were more

frequent, with participants receiving two videos each week. Other studies have demonstrated that individuals may use even more in-depth and feature-rich DMHIs in short bursts throughout their week and still receive considerable benefit[55]. The totality of the interaction with the DBT videos might reflect a more intense intervention than the SSIs that could have resulted in the observed change in that condition. The umbrella review of SSIs mentioned earlier[30] found four meta-analytic reviews comparing the effectiveness of SSIs and multi-session interventions, with two reporting no differences, one finding a significant difference favoring SSIs, and one finding a significant difference favoring multi-session interventions. Indeed, this seems like an unsolved issue requiring more research. Furthermore, developers and researchers of these interventions should think beyond each individual interaction to also consider broader patterns of use and engagement. In other types of DMHIs work has attempted to unpack patterns of users' engagement that correspond to differential individual traits and mental health outcomes[56].

Although we would caution that just because these types of interventions had an impact, does not necessarily mean they should be widely deployed. This is even the case when few other resources exist, as other studies have demonstrated that brief interventions may not be better than nothing. For example, one study found that among individuals with suicidal ideation who were offered care management or a brief online DBT skills training, the online DBT skills training actually increased nonfatal and fatal self-harm compared to usual care over an 18-month follow-up[57]. Individuals with self-harm recruited from outpatient settings represent a more severe population than we were working within our study, but it is a clear illustration of the potential negative impacts of the use of DMHIs. We would also note that we did not compare our interventions to more intense or longer-term interventions. Therefore, we cannot comment on their relative benefit compared to more standard care. As such, although we found support that the DBT videos could promote positive change in psychopathology, further work is needed before advocating for their widespread use.

Our study did not, however, find support for the benefits of the SSIs. In fact, both primary and supplemental analyses demonstrated the same pattern: strong confidence for a small change in the DBT video conditions and the lack of change in most other conditions. For the SSIs, we made adaptations both for cultural and age considerations. It is possible that the interventions we found to be ineffective because our adaptations were not sufficient for this population. Our research team had considerable expertise and lived experience in Chinese culture, yet we did not specifically validate the adapted interventions individually before conducting this study.

It is worth considering why the shortest intervention in terms of each interaction, the DBT videos, were the only intervention where we consistently found evidence of positive changes. These videos averaged about four and a half minutes in length, less than a quarter of the 20–30 minutes the SSIs were intended to take. We supplemented those videos, however, with follow-up questions, including calls to action. We also modified their dosage compared to past work. We sent participants two videos a week for 4 weeks, compared to daily videos for 14 days used in a previous study[32]. Along these lines, it is worth noting again, that although they were briefer than the SSIs in terms of each individual interaction, they were comparable in terms of overall interaction time and more frequent. Eight videos at a mean of 4.4 min per video yields ~35 min of total video time. These aspects may speak to the design elements of brief DMHIs. The best practices for designing SSIs have been proposed as using brain science to normalize concepts, empowering users to take an expert or helper role, using "saying is believing" exercises, featuring testimonials, and keeping SSIs targeted, brief, and focused on specific targets[58]. The SSIs in our study followed these best practices, which translates into more content, but the DBT videos were more focused and presented their intended skills and concepts directly. They did conform to the last design principle of being targeted, brief, and focused, and our call to action might have further emphasized the practical nature of these interventions. Future work should try to specify how to best operationalize concepts of targeted, brief, and focused within the context of DMHIs. For example, targeted might mean interventions that are selected for specific problem areas, brief might reflect a specific amount of time required, and focused could be the number of intervention elements present. Studies could compare DMHIs that vary in each of these factors to provide more concrete design recommendations for future interventions. Other conceptual models have proposed features of brief, targeted interventions, such as digital micro-interventions[59]. Further exploring the design space of brief DMHIs could advance the field and also support understanding of the specific mechanisms through which these interventions work.

Our analyses were based on the overall impact of the interventions across the HiTOP domains assessed in this study. Although this approach allowed us to efficiently evaluate intervention effects across a wide range of mental health symptoms, more nuanced analyses could explore whether specific interventions had stronger impacts for specific HiTOP domains or sub-domains. It is unlikely that interventions, especially brief interventions, would be uniformly beneficial for a wide range of psychopathologies. Conversely, existing evidence of the interventions used in this study have indicated support for cross-cutting aspects and common factors such as positive and negative affect, tolerance of emotions, hopelessness, and perceived control[32,37]. It is possible that various of the interventions examined in our study would have more clear evidence for change or benefit if we focused on specific spectra of the HiTOP, say internalizing or externalizing. Further studies might consider specific sub-domains of the HiTOP framework and evaluate interventions specifically tailored to those sub-domains. Given the number of HiTOP domains and sub-domains, we opted to focus on the overall benefit, yet we could explore questions like these in our data to guide future confirmatory work. That being said, our analytic approach modeled the mean treatment effect based on the distributions of each HiTOP spectra and subscale. This approach is different from merely summing these values into a composite score and is likely more conservative in terms of its estimation of change. Our hierarchical approach acknowledges that there is likely to be variability in how an overall treatment effect for a condition impacts each specific spectra, subscale, and participant. Given the conservative nature of this analytic strategy, finding any differences in overall change over time is noteworthy. Our approach, focusing on overall change at the treatment level, is a useful first step, especially given the novelty of the HiTOP framework in intervention research.

Critical questions remain, however, about how to best deliver such DMHIs to individuals. Some work has explored specific interventions among university students to increase intentions to use mental health services and found significant results[60]. These interventions combined personalized feedback, psychoeducation, and information about services. Other work has called out challenges and opportunities in recruiting university students into DMHIs[61] and highlighted the need to think not only about the effectiveness of such interventions, but also their marketing and dissemination. Brief interventions might have various advantages over longer-term programs, including the attractiveness and the lack of concern over attrition impacting the receipt of active ingredients. However, the long-term impact of such brief interventions needs to be evaluated, and our study did not look at outcomes after six weeks. Considerations of deployment should also address whether the amount of change found in this study justifies their use. Despite robust evidence suggesting a change in the DBT videos condition, the observed effect size was still quite small. Some calls have focused on the potential of brief interventions at the population-level [i.e., ref. 62]. However, it is also worth noting that small individual-level changes, even when applied across a population, still may or may not be impactful in producing meaningful change.

One unanticipated strength of our study was the extremely high retention rate and the completion of study measures across our different conditions (ranging from 96 to 100%). Although these rates are much higher than those typically reported for trials of DMHI, a recent review noted that engagement with DMHIs demonstrates a great deal of variability both in research studies and real-world engagement[63]. A few other studies conducted in mainland China have demonstrated similarly high retention rates in trials of DMHIs (i.e., 89%)[64] and longitudinal cohort studies (i.e., 87.7 to

96.5%)[65–67]. However, one study of a DMHI for depression for young adults in China found dropout rates of 56.1% at post-assessment and 72.7% at a three-month follow-up[68]. So, high retention does not seem to be guaranteed even among this population. A meta-analysis of treatment dropout in internet cognitive-behavioral therapy found an average estimate of 16.3% with an I^2 of 62% that indicates substantial heterogeneity among studies[69]. If we had anticipated such a high retention rate, we would have designed our study differently. As noted earlier, the inclusion of the participant preference arm was due to a lack of sufficient funds to provide compensation to all individuals. Given that compensation did not seem to substantially improve retention, and we did not have specific hypotheses about the role of compensation or preference in shaping the outcomes, we could have simplified our design. Future work should acknowledge that high attrition is not characteristic of all DMHI studies and attempt to better understand the characteristics of studies that achieve higher or lower retention rates. In the meta-analysis referenced above, the only significant subgroup difference was that attrition was significantly higher in internet cognitive-behavioral therapy compared to face-to-face cognitive-behavioral therapy when support was provided synchronously. However, the authors cautioned that this comparison was based on a few studies, indicating the need for future research.

Our findings should be considered in light of several limitations. Although we had a large sample, the sample was collected from four universities. These universities may not be representative of other universities in China, or students enrolled in these universities may not represent individuals in the same age range in other settings. The implications of our findings more broadly across China are unclear, and it would be interesting to consider their deployment and impact outside of university settings. Our sample was also a broad sample without inclusion or exclusion criteria based on symptom level. In light of this, our small effects where change was observed and null findings for several interventions might have been due to floor effects. We might find larger effects in a selected sample. Our sample might also contribute to the low internal reliabilities on some of our subscales. Those subscales consisted of several items that had low endorsement, which indicates low trait variance in our sample. Such low trait variance often leads to lower reliabilities[70]. Future research could use selected samples with individuals likely to endorse those symptoms to see if higher reliabilities occur. The use of the HiTOP framework as a primary outcome measure is novel, but also makes it hard to compare our findings to more traditional outcome measures used in this type of intervention research, including affect, depression, anxiety, hopefulness, and agency. The consistency of the primary analyses with the supplemental analyses using the DSM-5 Cross-Cutting Symptom Measure, however, supports that our findings are robust when considering other models of psychopathology as well. We cannot comment on the durability of any change, as our follow-up period was only a few weeks, and it is possible we might observe that effects dissipate or that sleeper effects would emerge over a longer follow-up. It is also worth noting that the intervention length differed and that the last DBT videos were delivered after the last SSI and closer in time to the follow-up assessment. This condition might appear more effective simply because participants engaged with the material more recently. We did not analyze engagement and how it related to outcome, as the scope of this analysis was to understand which interventions improved mental health symptoms. It is possible that differences in impacts reflect differences in engagement, and further analysis could help understand this relationship. Despite these limitations, our findings expand previous research in the areas of DMHIs, brief interventions, and the use of the HiTOP framework, and these limitations offer promising directions for future research.

Our analyses suggest that even brief DMHIs can be impactful for university student mental health in China. These findings carry potential clinical implications, especially given the well-documented shortage of mental health professionals and the growing burden of mental health conditions among young adults in this country. We also note that China is not unique in these challenges, as mental health demands among young people and university students outstrips supply elsewhere as well[71,72]. In this study, brief skills-based videos had a significant impact and outperformed other validated brief intervention models. Given substantial treatment gaps, SSIs and brief DMHIs hold strong potential as scalable, low-barrier solutions that could be systematically embedded within university mental health infrastructures, integrated into existing campus apps or digital platforms, and extended to broader healthcare settings such as primary care or community health centers. These tools are well-suited for incorporation into a stepped-care model, in which brief, self-guided interventions serve as a first-line resource before students are stepped up to more intensive or formal treatment as needed. Guided DMHIs have demonstrated utility in stepped-care formats compared to psychotherapy[73], and further investigating more low-intensity options holds potential. Even if these videos are a stepping stone into subsequent care, they could prepare students to gain basic mental health literacy and expose them to important principles of mental health treatment (e.g., change is possible, coping skills can alleviate distress) before they begin more formal treatment. Tiered-care approaches, where students are assessed and triaged to appropriate levels of care, are another model where brief DMHIs could hold value, especially when other tiers might involve human-supported DMHIs or other traditional treatments [i.e., ref. 74]. These approaches could meaningfully expand the reach of mental health support beyond what overstretched clinical services alone can provide. Future research should continue to validate these interventions across diverse settings, examine individual-level moderators to better understand what works for whom[75], and explore optimal strategies for integrating DMHIs and SSIs into both university systems and broader public health frameworks in China and similar contexts.

This study is a pioneering effort in the DMHIs space in several areas: the evaluation in China, a large sample size, high completion rate and retention, and direct comparison of various types of interventions. Our study found the strongest support for brief DBT videos in promoting positive mental health change. Our SSI conditions did not seem to result in change, with the exception of the paid SSI Order 1 (ABC, CARE). Participants were highly adherent to our procedures with completion rates ranging from 96–100%, although some participants did fail our manipulation checks, suggesting attention might have been slightly lower. Still, our mITT analysis sample was 97% of our randomized sample, which is extremely high for an RCT. This provides strong support for the feasibility of deploying these types of interventions among university students in China.

Data availability

De-identified analyzed dataset is available on our OSF page https://osf.io/qybtv.

Received: 15 August 2025; Accepted: 4 September 2026;
Published online: 24 September 2026

References

1. Huang, Y. et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry* **6**, 211–224 (2019).
2. Xu, Z. et al. The state of mental health care in China. *Asian J. Psychiatry* **69**, 102975 (2022).
3. Yue, J. L. et al. Workforce situation of the Chinese mental health care system: results from a cross-sectional study. *BMC Psychiatry* **22**, 562 (2022).
4. Moshe, I. et al. Digital interventions for the treatment of depression: a meta-analytic review. *Psychol. Bull.* **147**, 749–786 (2021).
5. Pauley, D., Cuijpers, P., Papola, D., Miguel, C. & Karyotaki, E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. *Psychol. Med.* **53**, 567–579 (2023).
6. Steubl, L., Sachser, C., Baumeister, H. & Domhardt, M. Mechanisms of change in Internet-and mobile-based interventions for PTSD: a systematic review and meta-analysis. *Eur. J. Psychotraumatol.* **12**, 1879551 (2021).

7. Hasan, F. et al. Comparative efficacy of digital cognitive behavioral therapy for insomnia: a systematic review and network meta-analysis. *Sleep. Med. Rev.* **61**, 101567 (2022).
8. Morales-Pillado, C. et al. Efficacy of technology-based interventions in psychosis: a systematic review and network meta-analysis. *Psychol. Med.* **53**, 6304–6315 (2023).
9. Anmella, G. et al. Smartphone-based interventions in bipolar disorder: systematic review and meta-analyses of efficacy. A position paper from the International Society for Bipolar Disorders (ISBD) Big Data Task Force. *Bipolar Disord.* **24**, 580–614 (2022).
10. Torok, M. et al. Suicide prevention using self-guided digital interventions: a systematic review and meta-analysis of randomised controlled trials. *Lancet Digit. Health* **2**, e25–e36 (2020).
11. Muñoz, R. F. Using evidence-based internet interventions to reduce health disparities worldwide. *J. Med. Internet Res.* **12**, e1463 (2010).
12. Schueller, S. M., Muñoz, R. F. & Mohr, D. C. Realizing the potential of behavioral intervention technologies. *Curr. Dir. Psychol. Sci.* **22**, 478–483 (2013).
13. Titov, N. et al. From research to practice: ten lessons in delivering digital mental health services. *J. Clin. Med.* **8**, 1239 (2019).
14. Nordberg, S. S. et al. Evaluating the Implementation and clinical effectiveness of an innovative digital first care model for behavioral health using the RE-AIM framework: quantitative evaluation. *J. Med. Internet Res.* **26**, e54528 (2024).
15. Richards, D. et al. Digital IAPT: the effectiveness & cost-effectiveness of internet-delivered interventions for depression and anxiety disorders in the improving access to psychological therapies programme: study protocol for a randomised control trial. *BMC Psychiatry* **18**, 1–13 (2018).
16. Borghouts, J. et al. Characteristics associated with the use of the mindfulness meditation app headspace in a large public health deployment: cross-sectional survey study. *JMIR Form. Res.* **9**, e73457 (2025).
17. Sit, H. F. et al. Digital mental health: a potential opportunity to improve health equity in China. *Lancet Public Health* **9**, e1136–e1141 (2024).
18. Zhang, C. et al. Associations between academic stress and depressive symptoms mediated by anxiety symptoms and hopelessness among Chinese college students. *Psychol. Res. Behav. Manag.* **15**, 547–556 (2022).
19. De Girolamo, G., Dagani, J., Purcell, R., Cocchi, A. & McGorry, P. D. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. *Epidemiol. Psychiatr. Sci.* **21**, 47–57 (2012).
20. Kodish, T. et al. Enhancing racial/ethnic equity in college student mental health through innovative screening and treatment. *Admin. Policy Ment. Health* **49**, 267–282 (2022).
21. Harrer, M. et al. Internet interventions for mental health in university students: a systematic review and meta-analysis. *Int. J. Methods Psychiatr. Res.* **28**, e1759 (2019).
22. Lattie, E. G. et al. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. *J. Med. Internet Res.* **21**, e12869 (2019).
23. Taylor, M. E. et al. The reach, effectiveness, adoption, implementation, and maintenance of digital mental health interventions for college students: a systematic review. *Curr. Psychiatry Rep.* **26**, 683–693 (2024).
24. Beidas, R. S., Saldana, L. & Shelton, R. C. Testing psychosocial interventions in the contexts they are meant to be delivered. *J. Consult. Clin. Psychol.* **91**, 189–191 (2023).
25. Mohr, D. C., Weingardt, K. R., Reddy, M. & Schueller, S. M. Three problems with current digital mental health research… and three things we can do about them. *Psychiatr. Serv.* **68**, 427–429 (2017).
26. Mohr, D. C., Lyon, A. R., Lattie, E. G., Reddy, M. & Schueller, S. M. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. *J. Med. Internet Res.* **19**, e7725 (2017).
27. Bautista, J. & Schueller, S. M. Understanding the adoption and use of digital mental health apps among college students: secondary analysis of a national survey. *JMIR Ment. Health* **10**, e43942 (2023).
28. Hamza, S. Z., Aguilar Silvan, Y. & Ng, L. C. Use of online tools for mental health among racially and ethnically diverse college students: mixed methods study. *JMIR Hum. Factors* **12**, e60628 (2025).
29. Schleider, J. L., Dobias, M. L., Sung, J. Y. & Mullarkey, M. C. Future directions in single-session youth mental health interventions. *J. Clin. Child Adolesc. Psychol.* **49**, 264–278 (2020).
30. Schleider, J. L. et al. Single-session interventions for mental health problems and service engagement: Umbrella review of systematic reviews and meta-analyses. *Annu. Rev. Clin. Psychol.* **21**, 279–303 (2025).
31. Jans, L. K., Sotomayor, I., & Schleider, J. L. in *Brief CBT and Science-Based Tailoring for Children, Adolescents, and Young Adults* (eds Storch, E. A. & Davis III, T. E.) Ch 13 (Springer Nature, 2024).
32. Rizvi, S. L. et al. Randomized clinical trial of a brief, scalable intervention for mental health sequelae in college students during the COVID-19 pandemic. *Behav. Res. Ther.* **149**, 104015 (2022).
33. Kotov, R. et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* **126**, 454–477 (2017).
34. Hierarchical Taxonomy of Psychopathology Society. The HiTOP self-report measures. https://www.hitop-system.org/hitop-self-report-measures (2024).
35. Potts, C. et al. Digital mental health interventions for young people aged 16-25 years: scoping review. *J. Med. Internet Res.* **27**, e72892 (2025).
36. Schleider, J. L., Mullarkey, M. C., Mumper, E., & Sung, J. Y. The ABC Project: action brings change. *OSF.* https://doi.org/10.17605/OSF.IO/QJ94C (2019).
37. Schleider, J. L., Dobias, M. L., Sung, J. Y., Mumper, E. & Mullarkey, M. C. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. *JMIR Ment. Health* **7**, e20513 (2020).
38. Dobias, M., Mullarkey, M. C. & Schleider, J. L. Project CARE. https://doi.org/10.17605/OSF.IO/YDQXJ (2019).
39. Alber, C. S., Krämer, L. V., Rosar, S. M. & Mueller-Weinitschke, C. Internet-based behavioral activation for depression: systematic review and meta-analysis. *J. Med. Internet Res.* **25**, e41643 (2023).
40. Jia, E., Macon, J., Doering, M. & Abraham, J. Effectiveness of digital behavioral activation interventions for depression and anxiety: systematic review and meta-analysis. *J. Med. Internet Res.* **27**, e68054 (2025).
41. Wasil, A. R., Venturo-Conerly, K. E., Shingleton, R. M. & Weisz, J. R. A review of popular smartphone apps for depression and anxiety: assessing the inclusion of evidence-based content. *Behav. Res. Ther.* **123**, 103498 (2019).
42. Schleider, J. & Weisz, J. R. Sharing feelings open. *OSF* https://doi.org/10.17605/OSF.IO/SWV5B (2019).
43. Schleider, J. et al. A randomized trial of online single-session interventions for adolescent depression during COVID-19. *Nat. Hum. Behav.* **6**, Article 2 (2022).
44. Song, Y. The sexuality education and attitudes of college students in China. *Health Educ.* **115**, 93–104 (2015).
45. Clarke, D. E. & Kuhl, E. A. DSM-5 cross-cutting symptom measures: a step towards the future of psychiatric care?. *World Psychiatry* **13**, 314 (2014).
46. Gelman, A., Hill, J. & Vehtari, A. *Regression and Other Stories* (Cambridge Univ. Press, 2021).
47. Plummer, M. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proc. 3rd International Workshop on*

Distributed Statistical Computing (eds Hornik, K., Leisch, F. & Zeileis, A) (2003).
48. Wetzels, R., Grasman, R. P. P. P. & Wagenmakers, E. An encompassing prior generalization of the Savage-Dickey density ratio test. *Comput. Stat. Data Anal.* **54**, 2094–2102 (2010).
49. Etz, A., Chávez de la Peña, A. F., Baroja, L., Medriano, K. & Vandekerckhove, J. The HDI+ ROPE decision rule is logically incoherent but we can fix it. *Psychol. Methods*. https://doi.org/10.1037/met0000660 (2024).
50. Wagenmakers, E.-J., Lee, M. D., Rouder, J. N. & Morey, R. D. In *The Theory of Statistics in Psychology* (Springer International Publishing, 2020).
51. The MathWorks Inc. MATLAB version 24.2 (R2024b). The MathWorks Inc. https://www.mathworks.com (2024).
52. Zhang, W. et al (2026). Validation of the Mandarin Chinese version of the detachment, internalizing, and somatoform spectra of the hierarchical taxonomy of psychopathology self-report HiTOP-SR. Preprint at https://osf.io/preprints/osf/t6vxy_v1 (2026).
53. Linardon, J. & Fuller-Tyszkiewicz, M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. *J. Consult. Clin. Psychol.* **88**, 1–13 (2020).
54. Kaveladze, B., Ghosh, A., Funkhouser, C., Schueller, S. & Schleider, J. Longer single-session interventions may not be better: evidence from two randomized controlled trials with online workers facing mental health struggles. *Clin. Psychol. Sci*. https://doi.org/10.1177/21677026251358836 (2025).
55. Mohr, D. C. et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. *J. Med. Internet Res.* **19**, e10 (2017).
56. Hoffman, V. et al. User engagement clusters of an 8-week digital mental health intervention guided by a relational agent (woebot): exploratory study. *J. Med. Internet Res.* **25**, e47198 (2023).
57. Simon, G. E. et al. Effect of offering care management or online dialectical behavior therapy skills training vs usual care on self-harm among adult outpatients with suicidal ideation: a randomized clinical trial. *JAMA* **327**, 630–638 (2022).
58. Manvelian, A. et al. A practical guide to adapting evidence-based treatments into digital single-session interventions for youth. *OSF*. https://doi.org/10.31234/osf.io/vjgxr_v1 (2025).
59. Baumel, A., Fleming, T. & Schueller, S. M. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. *J. Med. Internet Res.* **22**, e20631 (2020).
60. Ebert, D. D. et al. Increasing intentions to use mental health services among university students. Results of a pilot randomized controlled trial within the World Health Organization's World Mental Health International College Student Initiative. *Int. J. Methods Psychiatr. Res.* **28**, e1754 (2019).
61. Kodish, T., Schueller, S. M. & Lau, A. S. Barriers and strategies to improve digital mental health intervention uptake among college students of color: a modified Delphi study. *J. Behav. Cogn. Ther.* **33**, 10–23 (2023).
62. Dodge, K. A. et al. Population mental health science: guiding principles and initial agenda. *Am. Psychol.* **79**, 805–823 (2024).
63. Boucher, E. M. & Raiker, J. S. Engagement and retention in digital mental health interventions: a narrative review. *BMC Digit. Health* **2**, 52 (2024).
64. Wang, Y., Li, X., Zhang, Q., Yeung, D. & Wu, Y. Effect of a cognitive behavioral therapy–based AI chatbot on depression and loneliness in Chinese university students: randomized controlled trial with financial stress moderation. *JMIR mHealth uHealth* **13**, e63806 (2025).
65. Dou, K., Wang, M. L., Li, Y. Y., Yuan, X. Q. & Wang, L. X. The longitudinal association between peer victimization and problematic social media use among Chinese college students: testing a moderated mediation model. *Int. J. Ment. Health Addict.* **23**, 3540–3557 (2024).
66. Guo, L., Jiang, L. & Huang, H. A longitudinal study of four-year changes in physical fitness among university students before and after COVID-19: 2019–2022. *PLoS ONE* **20**, e0334088 (2025).
67. Lu, P. et al. Cohort profile: China undergraduate cohort for environmental health study. *BMC Public Health* **24**, 828 (2024).
68. Li, G. et al. A WHO digital intervention to address depression among young Chinese adults: a type 1 effectiveness-implementation randomized controlled trial. *Transl. Psychiatry* **14**, 102 (2024).
69. Linardon, J., Messer, M., Reid, R., Bolger, T. & Andersson, G. Absolute and relative rates of treatment non-initiation, dropout, and attrition in internet-based and face-to-face cognitive-behavioral therapy: a meta-analysis of randomized controlled trials. *Cogn. Behav. Ther.* **55**, 614–627 (2025).
70. McCrae, R. R., Kurtz, J. E., Yamagata, S. & Terracciano, A. Internal consistency, retest reliability, and their implications for personality scale validity. *Personal. Soc. Psychol. Rev.* **15**, 28–50 (2011).
71. McGorry, P., Gunasiri, H., Mei, C., Rice, S. & Gao, C. X. The youth mental health crisis: analysis and solutions. *Front. Psychiatry* **15**, 1517533 (2025).
72. McGorry, P. D. et al. The Lancet psychiatry commission on youth mental health. *Lancet Psychiatry* **11**, 731–774 (2024).
73. Mohr, D. C. et al. A randomized noninferiority trial evaluating remotely-delivered stepped care for depression using internet cognitive behavioral therapy (CBT) and telephone CBT. *Behav. Res. Ther.* **123**, 103485 (2019).
74. Wolitzky-Taylor, K., Wen, A., Freimer, N. & Craske, M. G. Anxiety and depression in emerging adults: the STAND program as a model of scalable screening and intervention. *Neuropsychopharmacology* **51**, 244–258 (2026).
75. Benjet, C. et al. A precision treatment model for internet-delivered cognitive behavioral therapy for anxiety and depression among university students: a secondary analysis of a randomized clinical trial. *JAMA Psychiatry* **80**, 768–777 (2023).

Acknowledgements

We acknowledge the efforts of the research assistants at Beijing Normal University who supported data collection and preparation. This study was supported by funding from the Cyrus Tang Foundation. S.M.S. has received consulting payments from Boehringer Ingelheim and Otsuka Pharmaceuticals and is a member of the Headspace Scientific Advisory Board, for which he receives compensation. The other authors do not have a competing interest.

Author contributions

S.M.S. had the primary responsibility for drafting the initial manuscript and incorporating feedback, contributed to obtaining funding for this project, and prepared Fig. 2. R.L. had the primary responsibility for study management and data collection and also provided feedback on the manuscript and contributed to obtaining funding for this project. M.D.L. had the primary responsibility for data analysis and initial drafting of sections related to data analysis, also provided feedback on the manuscript and to obtaining funding for this project, and prepared Fig. 3. E.B. contributed to obtaining funding, translation, and providing feedback on the manuscript. M.H.-N.C. contributed to translation, adaptation, and provided feedback on the manuscript. C.C. had the primary responsibility for obtaining and managing funding, as well as overseeing translation and cultural adaptation of all materials and provided feedback on the manuscript. Y.A.C. contributed to translation, adaptation, and provided feedback on the manuscript. T.H. contributed to study management, adaption, data collection, and provided feedback on the manuscript. T.K. contributed to study management, adaption, data collection, and provided feedback on the manuscript. Y.L. contributed to study planning, translation and adaptation of the study materials, and provided feedback on the manuscript. E.A.M. had the primary

responsibility for oversight on the use of the HiTOP framework and also contributed to obtaining funding and provided feedback on the manuscript. P.W. contributed to data management, reporting, translation, adaptation, provided feedback on the manuscript, and prepared Fig. 2. Z.W. contributed to study management, adaption, data collection, and provided feedback on the manuscript. N.W. contributed to translation, adaptation, provided feedback on the manuscript, and prepared Fig. 1. J.W. contributed to obtaining funding and provided feedback on the manuscript. W.N. contributed to study management, adaption, data collection, and provided feedback on the manuscript. Y.Y. contributed to study management, adaption, data collection, and provided feedback on the manuscript. Z.Y. contributed to obtaining funding, translation and adaptation of study materials, and provided feedback on the manuscript. W.Z. contributed to data management, reporting, translation, adaptation, provided feedback on the manuscript, and prepared Fig. 2. Y.Z. contributed to study planning, translation and adaptation of the study materials, and provided feedback on the manuscript. L.L. contributed to obtaining funding, study management, data collection, and provided feedback on the manuscript.

Competing interests

S.M.S. has received consulting payments from Boehringer Ingelheim and Otsuka Pharmaceuticals and is a member of the Headspace Scientific Advisory Board, for which he receives compensation.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44184-026-00247-w.

Correspondence and requests for materials should be addressed to Stephen M. Schueller or Liang Luo.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.