Mindfulness
https://doi.org/10.1007/s12671-022-01952-1

ORIGINAL PAPER

Can the Academic and Experiential Study of Flourishing Improve Flourishing in College Students? A Multi-university Study

Matthew J. Hirshberg[1] · Blake A. Colaianne[2] · Mark T. Greenberg[2] · Karen Kurotsuchi Inkelas[3,4] · Richard J. Davidson[1] · David Germano[4] · John D. Dunne[1] · Robert W. Roeser[2]

Accepted: 14 July 2022

Abstract

Objectives Significant concerns have been raised about the "mental health crisis" on college campuses, with attention turning to what colleges can do beyond counseling services to address students' mental health and well-being. We examined whether primarily first-year (89.1%) undergraduate students ($n = 651$) who enrolled in the Art and Science of Human Flourishing (ASHF), a novel academic and experiential for-credit elective course on human flourishing, would demonstrate improved mental health and strengthen skills, perspectives, and behaviors associated with flourishing relative to students who did not enroll in this course.

Methods In a two-wave, multi-site, propensity-score matched controlled trial (ASHF $n = 217$, control $n = 434$; $N = 651$), we used hierarchal linear models and false discovery rate corrected doubly robust estimates to evaluate the impact of the ASHF on attention and social-emotional skill development, flourishing perspectives, mental health, health, and risk behavior outcomes.

Results ASHF participants reported significantly improved mental health (i.e., reduced depression) and flourishing, improvements on multiple attention and social-emotional skills (e.g., attention function, self-compassion), and increases in prosocial attitudes (empathic concern, shared humanity; Cohen's ds = 0.18–0.46) compared to controls. There was no evidence for ASHF course impacts on health or risk behaviors, raising the possibility that these outcomes take more time to change.

Conclusions This research provides initial evidence that the ASHF course may be a promising curricular approach to reduce and potentially prevent poor mental health while promoting flourishing in college students. Continued research is needed to confirm these conclusions.

Keywords Flourishing · College student mental health · Meditation · Depression · Mindfulness

The transition to college is a major life event that is becoming normative for young adults in the USA. In 2017, 48% of American 18- and 19-year-olds enrolled in higher education (US Department of Education, 2019). During the transition into college, students must adapt to a demanding intellectual environment, develop new social networks, and begin to independently make lifestyle choices (e.g., studying, sleep, alcohol consumption). These transitions are especially salient for students who move away from home to attend college, where there is less adult support and supervision, heightened peer pressures, and greater personal autonomy (Roisman et al., 2004). Research shows that navigating these new environmental demands can be stressful and can affect young adults' mental health (Gall et al., 2000).

Mental health can be defined as the absence of distress (e.g., depression) and the presence of well-being (e.g., life satisfaction; Kazdin, 1993). Mental health disorders become more common during the college years, and college student distress has been increasing while well-being has been

✉ Matthew J. Hirshberg
hirshberg@wisc.edu

1 Center for Healthy Minds, University of WI–Madison, 625 W. Washington Ave, Madison, WI 53703, USA

2 The Pennsylvania State University, Human Development and Family Studies, HHD Building, University Park, State College, PA 16801, USA

3 University of Virginia, School of Education and Human Development, 405 Emmet St S, Charlottesville, VA 22904, USA

4 Contemplative Sciences Center, University of Virginia, 102 Cresap Rd, Charlottesville, VA 22903, USA

Published online: 23 July 2022

declining for some time. For example, two large nationally representative cohort surveys of American college students found substantial increases in mental health disorders (e.g., depression + 34%, suicidal ideation + 74%) between 2007 and 2018 (Duffy et al., 2019). A World Health Organization–sponsored global investigation of college student mental health reached similar conclusions (Auerbach et al., 2018). In 2018, around 40% of American college students experienced one or more significant mental health challenges (e.g., depression; Duffy et al., 2019; Lipson et al., 2018). At the same time, fewer students report a sense of flourishing, defined as a life imbued with meaning, purpose, strong and rewarding social connections, optimism, and engagement (Seligman, 2018). The rate of college students reporting low flourishing (i.e., one standard deviation below the mean) increased by 119%, to nearly a quarter of all respondents from 2012 to 2018 (Duffy et al., 2019). Finally, we know that the emergence or worsening of mental health problems during the transition to college can impact subsequent college achievement, college completion, and labor market outcomes in adulthood (Fletcher, 2010; Mojtabai et al., 2015).

In sum, significant concerns have been raised about the "mental health crisis" on many college campuses today, and focus has turned to what colleges can do beyond counseling services to address students' mental health (Center for Collegiate Mental Health, 2020). While the burden of these issues is often placed on student affairs, university counseling centers are struggling to adequately address the growing demands on their services (Orben et al., 2020; Xiao, et al., 2017). As a consequence, many colleges and universities are implementing new strategies to supplement traditional efforts to stem the tide of deteriorating college student mental health.

Preventive approaches that bring multiple campus units together (e.g., student services, residential life, academic courses) to intervene and bolster mental health may improve student retention, psychological functioning, and graduation rates (Renn & Reason, 2012). For instance, first-year courses that teach students skills for managing transition-related and other common life stressors are a promising and scalable prevention approach (Lang, 2007). In addition, there is some evidence that meditation programs generally (Shapiro et al., 2011) and mindfulness-based interventions (MBIs) in particular can help to strengthen college students' cognitive (e.g., attentional) and social-emotional (e.g., compassion for self and others) skills, and overall well-being during this transitional period (Dvořáková et al., 2019).

MBIs comprise a range of interventions that aim, through didactive instruction and practice, to strengthen skills of awareness (e.g., the ability to pay attention with an attitude of openness or acceptance). As noted, experimental evidence from research with college students indicates that the approach has promise. For example, Dvořáková et al. (2017) randomly assigned 109 first-year college students to either a MBI or wait-list condition. Assignment to the MBI predicted significantly increased life satisfaction and significantly reduced anxiety and depressive symptoms. A recent meta-analysis of randomized controlled trials (RCTs) of MBIs in college students reported that this class of intervention results in reduced distress (e.g., anxiety and depressive symptoms), while increasing mindfulness, self-compassion, and well-being. These results are largely consistent with the more mature literature on the benefits of MBIs in healthy adult and clinical populations (e.g., Goldberg et al., 2022). Extant evidence also suggests, importantly, that participation in MBIs is not associated with increased rates of adverse outcomes relative to population base rates (Hirshberg et al., 2020, 2021).

Loving-kindness and compassion practice interventions (i.e., connection practices) have also been shown to cultivate important skills for mental health — specifically the strengthening of prosocial dispositions and attitudes toward one's self and others (Dahl et al., 2015, 2020). While less researched than MBIs, a meta-analysis on connection interventions in the general population found significant positive effects on depressive symptoms, mindfulness, and self-compassion (Galante et al., 2014). Another meta-analysis of 21 adult-focused connection RCTs found that program participation predicted increases in compassion, self-compassion, mindfulness, well-being, and reductions in anxiety, depression, and distress (Kirby, 2017). Research with college students has shown that exposure to brief loving-kindness meditation increased resilience to the negative impacts of an acute stressor (Hirshberg et al., 2018) and longer duration connection training supports reductions in anxiety and depressive symptoms as well as increases in well-being (Smeets et al., 2014; Totzeck et al., 2020).

A limitation of most research with college students is that the programs are either (a) implemented outside of typical academics (e.g., extracurricular activities) or (b) structured so that they add onto or take away time from academic time (Dvořáková et al., 2017; Huberty et al., 2019). Courses offered outside of or in addition to typical academic demands may present participation barriers that disproportionately affect some students. For example, students who must work to support their education might have less available time to enroll in extracurricular activities. These courses may also be more challenging to implement as universal prevention strategies, for example, because they may be perceived as detracting from academic time. Thus, a curricular approach may be more equitable and could in theory be offered to all in-coming college students as an important part of an overall strategy of putting well-being at the center of a university's mission (Renn & Reason, 2012).

Scholars of higher education have argued for a return to holistic approaches to education in which developing student

knowledge about various topics (i.e., declarative or content knowledge) is balanced with the structured cultivation of important cognitive, social, and emotional skills (i.e., knowing how to do something; Palmer et al., 2010). Research has demonstrated the centrality of skills to well-being (Dahl et al., 2020), positive social relationships, health behaviors, labor market outcomes (Moffitt et al., 2011), and academic achievement across the lifespan (Chiteji, 2010). General education courses may be a prime opportunity to introduce curricula that integrates declarative knowledge with procedural skill development, particularly around improving well-being, healthy behaviors, and enriching social relationships (Keeling, 2013; Palmer et al., 2010). Such an educational approach would not only address students' mental health and flourishing, but is also consistent with the intent of undergraduate general education, which Walker and Soltis (2009) summarize as threefold: (1) pass knowledge into perpetuity; (2) cultivate a productive citizenry; and (3) prepare students for life by helping them explore their potential.

We are aware of a few innovative attempts to provide undergraduates with academic courses that also emphasize the development of self-knowledge and flourishing skills. As one example, Yale University's "The Science of Well-Being" course provides undergraduates with an overview of the current science on well-being and happiness as well as opportunities to practice some flourishing skills (Santos, 2018). Although this course provides experiential opportunities to apply course content, it focuses more on cognitive strategies (e.g., attributional reframing; Santos, 2018) than on mindfulness or compassion meditation for developing flourishing-related skills.

We propose that declarative learning and procedural learning are both equally important to mental health and flourishing, and strengthening flourishing skills through practices of awareness, connection, and other contemplative strategies may be especially conducive to durable enhancements of mental health and flourishing (Dahl et al., 2015). Therefore, courses on flourishing that address both a declarative understanding of flourishing concepts and strengthen skills through forms of practice that help one to actually flourish, according to our theory of change, are those that are most likely to produce substantive and lasting benefits on mental health outcomes and more distal risk and health behavior outcomes (Fig. 1).

Beginning in 2016, a multidisciplinary team of scholars (e.g., education, neuroscience, developmental and prevention science, religious studies) from three American research universities set out to develop, based on available evidence, a scalable, for-credit course that could become a general education requirement, similar to courses such as composition or calculus that are required for first-year students on many campuses. Out of this effort, the *Art and Science of Human Flourishing* (ASHF) course was developed.

The ASHF is an academically rigorous credit-bearing course, cross-listed in multiple units, that fulfills undergraduate general education requirements. It melds intellectual rigor (e.g., declarative learning) on what constitutes a "life of flourishing" with semester-long experiential (i.e., procedural) learning in awareness, connection, and other meditation techniques that support flourishing directly (Dahl et al., 2020). Several features differentiate the ASHF from similar prior efforts. For example, while concepts and meditation practices associated with MBIs and connection interventions are integral, so too are diverse disciplinary (e.g., religious studies, neuroscience, social psychology) and historical perspectives on and practices of flourishing (e.g., Stoicism, Taoism). As a result, the course provides students with multiple potential "ingredients" of flourishing (e.g., knowledge, skills, and perspectives) from multiple perspectives, allowing them to investigate which are most resonant to their lives now. This multifaceted approach to flourishing fits within university credit requirements such that, should it become an undergraduate

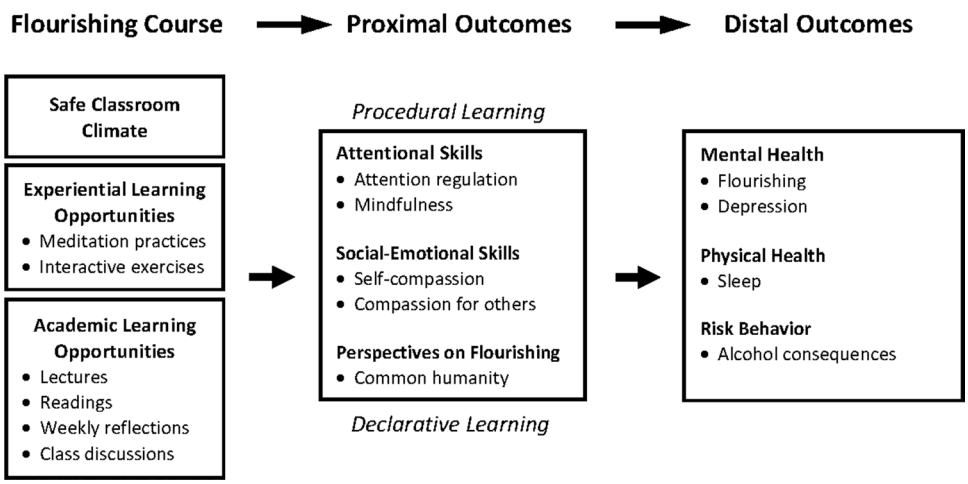

Fig. 1 Theory of change: Art and Science of Human Flourishing course on student outcomes. *Note*: Experiential learning opportunities are predicted to lead to attention and social-emotional skills learning. Academic learning opportunities are predicted to lead to effects on knowledge of and perspectives on flourishing. Both experiential and academic learning outcomes are predicted to impact distal outcomes through students' application of these skills and perspectives

prerequisite like other first-year seminars, it could easily be implemented as a universal prevention strategy.

To conceptualize what it means to "flourish," we borrow from multiple theories of flourishing (Dahl et al., 2020; Ryff & Keyes, 1995; Seligman, 2018). We define flourishing as leading a meaningful, fulfilling, and engaged life that benefits self and others. As illustrated in our model of change (Fig. 1), we propose that flourishing emerges from skills and perspectives related to awareness (e.g., attention, mindfulness), connection with others (e.g., kindness to self/others, social connection), personal insight (i.e., clarity regarding one's values, worldviews, and identities), and the ability to integrate and embody these skills and perspectives in the pursuit of a meaningful life. Thus, the ASHF course is structured along five main dimensions: (1) Foundations of Flourishing, (2) Awareness, (3) Connection, (4) Insight, and (5) Integration (Fig. 2) that are based upon a framework for understanding the plasticity of well-being (Dahl et al., 2020). Each main dimension of the course is further subdivided into three themes, each discussed for 1 week. This summates to a curriculum that fits a typical 15-week semester undergraduate course.

The *Foundations Dimension* of the course encompasses explorations of diverse definitions of flourishing; the science of personal transformation through effort, practice, and support; and the need for resilience on the path of flourishing. The *Awareness Dimension* explores the importance of emotion, focus, and mindfulness to flourishing. The *Connection Dimension* examines the qualities of interdependence, compassion, and belonging — and how these relational qualities are fundamental to not only our own flourishing, but our ethical responsibility to others and their flourishing. The *Insight Dimension* investigates the qualities of identity, values, and aesthetics/imagination in service of crafting a vision and plan for flourishing. Finally, the *Integration Dimension* completes the course by pulling together all prior dimensions while exploring the role of courage, community, and intention on the path to and in the embodiment of flourishing (Fig. 2).

Academic learning (i.e., declarative knowledge) is supported through assigned scholarly readings, mini-class lectures, weekly written student reflections, whole- and small-group discussion and exercises, and other academically focused assessments. Experiential learning is supported through once per week "meditation labs" in which practices intended to development attentional and social-emotional skills are taught and practiced (e.g., mindfulness, compassion, gratitude practices). Practice-based experiential learning is also integrated into the twice-weekly seminars through additional brief opportunities to strengthen skills relevant to current course content. Students are given resources and guidance on how to develop a home meditation practice outside of class and are asked to keep a journal of weekly practice.

The current study consists of a two-wave, multi-site, quasi-experimental, propensity-score matched (PSM) controlled study of the ASHF course in primarily first-year college students (ASHF $n = 217$, Control $n = 434$; $N = 651$; 89.1% first-year students). The waves were conducted during the fall semester of 2018 and 2019, respectively. Based on our theory of change (Fig. 1), we predicted that, relative to PSM controls, ASHF students would demonstrate larger improvements in attention functioning, social-emotional skills, perspectives of flourishing, mental health, health, and risk behaviors. To examine these hypotheses, we assessed at pre- and post-test attention function, attention regulation, and mindfulness (awareness skills); distress tolerance, self-compassion, perspective-taking, empathic concern, roommate compassion, and compassion for others (social-emotional skills); shared humanity and fairness and equity thinking (perspectives on flourishing); depressive symptoms and flourishing (mental health); sleep quality (health); and alcohol consequences (risk behaviors).

Method

Participants

Participants were primarily first-year undergraduates attending one of the three universities involved in this

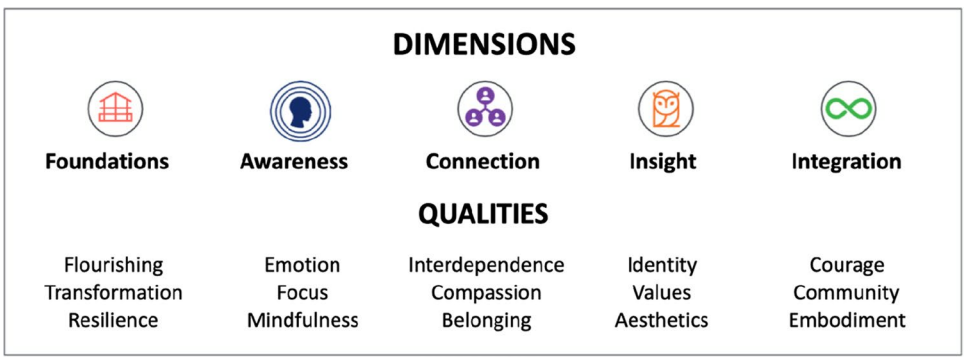

Fig. 2 Art and Science of Human Flourishing curriculum: dimensions and quality of flourishing. *Note*: Dimensions represent the five overarching domains of flourishing in the course. Qualities represent the specific knowledge, perspectives, and skills of flourishing taught during the course. Each quality reflects 1 week of course content (15 weeks total)

research in the fall of 2018 and 2019. Students who elected to enroll in the ASHF course for general education credits were recruited to participate in research on the effects of the ASHF course through in-class announcements and emails. Separately at each university, control students were recruited through emails describing a research opportunity on student well-being. The initial sample (i.e., before PSM) was $n = 918$ (217 ASHF, 701 control).

Following PSM (methods described below), the analysis sample was $N = 651$ ($n = 217$ ASHF, $n = 454$ PSM control). ASHF participants were 74.19% female ($n = 161$), 65.44% White ($n = 142$), 15.67% Asian/Pacific Islander ($n = 34$), 5.53% Hispanic ($n = 12$), 4.61% African American or Black ($n = 10$), and 0.46% American Indian/Native American ($n = 1$), with 8.29% preferring not to report their race ($n = 18$). PSM control participants were 75.35% female ($n = 322$), 67.51% White ($n = 293$), 15.67% Asian/Pacific Islander ($n = 68$), 3.00% Hispanic ($n = 13$), 3.92% African American or Black ($n = 17$), and 0.23% American Indian/Native American ($n = 1$), with 9.68% preferring not to report their race ($n = 42$). There were 89 ASHF and 160 PSM controls from university a, 69 ASHF and 159 PSM controls from university b, and 59 ASHF and 115 PSM controls from university c. The majority of students ($n = 580$; 89.1%) were first-years, 2.8% ($n = 18$) were second-years, 5.5% ($n = 36$) were third-years, and 2.6% ($n = 17$) were ≥ fourth-years. Just over 5% ($n = 33$) of the sample were international students.

The target sample size in each study wave (i.e., fall of 2018 and 2019) was based on the required number of participants to conduct two (control) to one (ASHF) PSM. The target control sample size was based on the expectation that 50% of eligible ASHF enrolled students would participate in the research and that three control participants for each ASHF enrolled student would be required for the planned PSM procedure. However, we did not limit the number of ASHF enrolled students who could participate in the research.

Procedure

In both waves, students selecting into the ASHF course were recruited through an in-class visit on the first day of class and emails. At each university, control students were recruited through emails describing an opportunity to participate in research on student well-being. Eligible students (i.e., > 18 years old) were required to read and sign an electronic consent document before beginning the pre-test survey. The pre-test survey was collected during the first 2 weeks of the semester. In the case of the ASHF study arm, approximately 40% of students had experienced one class of ASHF content prior to pretest (the first class involved an overview of the syllabus and introductions). About 14 weeks after the pre-test window, at the end of the same semester, an email with the post-test survey link was sent to participants. Both the pre- and post-test surveys were completed by participants online via Qualtrics® at a time and place of their choosing. Participants with > 50% of pre-test items completed were included in data processing and analysis regardless of post-test completion. In wave 1, participants were compensated $20 for each completed survey. In wave 2, participants were entered into a lottery to win one of 12 $200 awards (each university held a separate lottery). All procedures and study materials were approved by each university's respective ethics board. A Certificate of Confidentiality was obtained to protect participant privacy.

Before this research was conducted, the ASHF was piloted on the campuses of the three collaborating universities in Fall 2017 ($n = 150$). Formative and course evaluation data informed substantive course revisions reflected in the curriculum taught in wave I (fall 2018). Further but comparatively minor refinements were made to the course in wave II (fall 2019). The overall structure of the course, topic domains, and types of pedagogical practices utilized were equivalent between universities and across waves. To date, the only analyses on any part of the present data examined the distribution of pedagogical practices employed by each university and whether any observed differences were associated with ASHF student outcomes in wave I. Although one instructor utilized lecture to a greater degree than the other two (21%, 34.6%, and 55%, respectively) and the amount of time dedicated to class activities (35.2%, 46.3%, and 19.5%, respectively) differed as well, no significant differences in student outcomes were observed (Inkelas et al., 2021). These results suggest that student outcomes are robust to at least the level of variability observed in wave I. This article is the first to address the question of ASHF effectiveness and to use wave I and wave II data.

Although manualized curricula provide experimental benefits and help ensure fidelity of implementation, they may limit the generalizability of observed effects in real-world contexts. The goal of this consortium was to construct and evaluate a course that could be widely disseminated at colleges and universities. The nature of this goal implies that rigid control during implementation must give way to pragmatic constraints, and that a degree of implementation variability is inevitable and potentially desirable given natural variation in contexts (Bryk, 2016). For instance, some instructors may prefer didactic lecture or discussion while others may prefer discussion over in-class experiential learning. Instead of pursuing "fidelity of implementation" to a scripted course, we instead pursued "implementation with integrity" (Bryk, 2016).

Instructors had autonomy to mold the course around their preferred instructional styles, within bounds. Instructors

were in close communication before, during, and after implementation to ensure consistency in course scope and sequence and core content. They all used the same readings, course videos, and the 15-week structure described above (Fig. 2). In addition, the instructional teams involved at each university possessed expertise in areas deemed essential to high fidelity implementation of an academic and experiential (i.e., meditation-based) curriculum. In particular, all instructional teams included experts in psychology or humanities who had a longstanding personal meditation practice. This research provides real-world estimates of the average effects of the ASHF across contexts and waves when implemented by instructors who the consortium believed possessed the necessary skills and knowledge to instruct the course with fidelity. We acknowledge, however, optimizing the course and determining which elements are most effective in improving student flourishing may require tighter experimental control.

Measures

Measures were selected to correspond to weekly content on flourishing qualities and key mental health outcomes (Fig. 2). Measure names along with pre- and post-test reliability statistics, number of items, and a sample item are provided in Table 1. To reduce participant burden, we selected a subset of items from some measures based on item face validity and anchored all five option Likert-like scales at 1 (Not at all true for me) to 5 (Very true for me). Otherwise, we used original response anchors.

For depression, in addition to average depressive scores on the Patient Health Questionnaire-8 (PHQ; Kroenke et al., 2009), we used standardized cut-points to construct two dichotomous (No/Yes) symptom category scores: ≤ 8.87 = not clinical and > 8.87 = clinical depressive symptoms; and ≤ 13.32 = not severe depression and scores > 13.32 = severe depression (Ettman et al., 2020). For analyses of alcohol consequences, all available data was included in analyses. Students who do not drink would necessarily report no alcohol consequences. Although including in analyses students who do not drink may lead to a floor effect in which pre-test levels of alcohol consequences are so low that detecting change is difficult, we include all data because we did not prespecify analyses on a subset of participants (i.e., those that drink at baseline).

Data Analyses

We conducted PSM using the nearest-neighbor approach to allow for causal inferences without random assignment given the challenges of randomly assigning participants to for-credit university classes (Dehejia & Wahba, 2002). Data from the waves were combined. We then followed a two-step approach for PSM (Pishgar et al., 2020). Because the MatchIt package in R requires complete data and there was a small amount of missingness on T1 variables (0.33% [attention function] to 8.17% [compassionate roommate goals]), we first multiply imputed 50 complete T1 datasets using the MICE package (Pishgar et al., 2020). All variables to be included in PSM were included in the multiple imputation procedure. After examining imputed data for plausibility, we used the "matchthem" function in MatchIt to select the two control participants for every one ASHF participant with the nearest propensity score (i.e., 2:1 nearest neighbor approach). Multiply imputed data were only used for the PSM procedure. The matched participants' IDs along with matching statistics were then concatenated with the raw data, resulting in the final dataset used in outcome analyses.

As described below, during outcome analyses, missing data were addressed with maximum likelihood estimation in primary outcome models and pattern-mixture modeling in sensitivity analyses. In addition to baseline scores on all outcomes, participant gender, race, university, wave, country of childhood (USA/not USA), and search for meaning in life were included in both the multiple imputation and PSM procedures. The latter two covariates were included because before matching, the groups differed significantly on them and including them improved the results of the PSM procedure (i.e., resulted in increased balance). After constructing the PSM sample, we evaluated balance statistics (i.e., standardized mean differences), qqplots, and a love plot, and estimated independent sample *t*-tests on all pre-test scores to determine the effectiveness of the PSM procedure.

We used hierarchical linear models (HLMs) in which participants are nested within university, group (PSM controls/ASHF), and wave (I or II) for a total of 12 level-2 clusters (i.e., universities a, b, and c, ASHF and PSM controls, in waves I and II). Although 12 is a small number of level-2 clusters, research shows that HLMs produce unbiased estimates with as few as 10 (Huang, 2018a; McNeish & Stapleton, 2016) and that failing to properly model nesting effects can increase type I error (Huang, 2018b). Because of the small number of level-2 units, we used restricted maximum likelihood estimation with a Satterthwaite approximation (McNeish & Stapleton, 2016) to conduct intention-to-treat analyses on all ASHF and PSM control participants. We achieved doubly robust estimates by including the propensity-score as a weight in HLM models, providing unbiased estimates of average treatment effects when at least one of the models (i.e., PSM or HLM) is correctly specified (Funk et al., 2011).

In HLM, we regressed the post-test outcome score on the pre-test outcome score, while including as covariates study wave, participant gender, race, and group (ASHF/control), with the group coefficient (PSM control/ASHF) of primary interest. For all models, we report the unconditional

Table 1 Measure, construct, number of items, internal reliability, and example item

Measure	Construct (n items)	α T1/T2	Example item
Proximal skills: attention skills			
Attentional Function Index (Cimprich et al., 2011)	Attention function (10)	.85/.86	"Following through on your plans"
Multidimensional Assessment of Interoceptive Awareness (MAIA;Mehling et al., 2012)	Attention regulation (3)	.57/.67	"I can refocus my attention from thinking to sensing my body."
Five Facet Mindfulness Questionnaire (Baer et al., 2008)	Mindfulness (24)	.80/.83	"It seems I am 'running on automatic' without much awareness of what I'm doing."
Social-emotional skills			
MAIA	Distress tolerance (2)	.74/.78	"When I feel physical pain, I become upset."
Self-Compassion Short-Form (Raes et al., 2011)	Self-compassion (12)	.85/.85	"When something upsets me, I try to keep my emotions in balance."
Interpersonal Reactivity Index (Davis, 1980)	1. Perspective-taking (3) 2. Empathic concern (4)	.81/.82 .81/.83	"Before criticizing somebody, I try to imagine how I would feel if 'I were in their place.'" "I feel very sorry for people when they are having problems."
Compassionate Goals (Crocker & Canevello, 2008)	Roommate compassion (9)	.90/.90	"Make a positive difference in your roommate's life."
Compassion for Others (Gilbert et al., 2017)	Empathy (8)	.81/.82	"I reflect on and make sense of other people's distress."
Perspectives on flourishing			
Spiritual Transcendence Scale (Piedmont, 1999)	Shared humanity (4)	.88/.89	"I believe that all of life is interconnected."
Social Awareness Index (Lerner et al., 2008)	Fairness and equity (7)	.92/.93	"It is important to me to make the world a better place to live in."
Short-term developmental outcomes: mental health — distress			
Patient Health Questionnaire-8 (Kroenke et al., 2009)	Depressive symptoms (8)	88/.88	"Feeling down, depressed, or hopeless"
Mental health — flourishing			
Pemberton Happiness Index (Hervás & Vázquez, 2013)	Flourishing (11)	.93/.92	"I am satisfied with myself."
Physical health			
Pittsburgh Sleep Quality Index (Buysse et al., 1989)	Sleep quality (5)	.52/.55	"During the past month, rate your overall sleep quality?"
Risk behaviors			
Alcohol Consequences Questionnaire (Kahler et al., 2005)	Alcohol consequences (11)	.88/.89	"I have passed out from drinking"

Two of three distress tolerance items were used. Five of 19 Pittsburgh Sleep Quality Inventory items were used

intraclass correlation coefficient (ICC) which can be understood of the proportion of variance in student outcomes accounted for by their cluster. We use false discovery rate correction (FDR; Benjamini & Hochberg, 1995) on the group contrasts of all outcomes to ensure that the ratio of false positives to true positives does not exceed the two-tailed $p < 0.05$ threshold for statistical significance. As a magnitude of an effect's size and its variability, we estimated a model-based equivalent of Cohen's d and its 95% confidence interval (CI) for continuous outcomes. As there is no consensus on calculating standardized effect sizes from HLMs, we calculated Cohen's d as the unstandardized beta of the group contrast divided by the standard deviation of the sample pre-test score on the outcome (Feingold, 2009). We describe $d = 0.20$ as a small effect, $d = 0.50$ as a moderate effect, and $d = 0.80$ as a large effect (Cohen, 1977). For dichotomous depression severity outcomes, we report the percent change in prevalence within each group and the odds ratio (OR) as estimates of an effect's magnitude. R statistical software was used for all analyses (R Core Team, 2021). A full list of packages is provided in Supplemental materials. Due to complications arising from multiple ethics boards, the data involved in this study cannot be made publicly available.

Missing data can bias estimates of treatment effects when data are missing not at random (MNAR). The

full-information maximum likelihood estimation used in primary models is robust to data missing at random (MAR; Enders, 2001). We examine the robustness of these estimated treatment effects to multiple MNAR assumptions by regressing outcome T2 scores on T1 scores, extracting the residuals, and then simulating results under the following assumptions: (1) missing data are equal to the predicted T2 sample average; (2) missing data are 0.20 standard deviations worse than the predicted sample T2 average; (3) missing data are 0.50 standard deviations worse than the predicted sample T2 average; and (4) missing data are 0.80 standard deviations worse than the predicted sample T2 average. After imputing missing values based on these assumptions, we compare between group change using non-parametric Wilcoxon rank-sum tests to evaluate the robustness of effects to each assumption (Goldberg et al., 2021).

We conducted one exploratory analysis. Although evidence suggests that moderate exposure to meditation does not lead to adverse events (i.e., increase the prevalence of symptoms/disorders), several researchers have noted the lack of examination of the potential for harm in meditation research (Hirshberg et al., 2020; Lindahl et al., 2017). Examining the potential for adverse effects is particularly important in a classroom setting and to inform future policies that might consider implementing the ASHF or similar curricula more broadly. We operationalize an adverse event as a categorical worsening in depressive symptoms from pre- to post-test (i.e., from non-clinical to clinical, clinical to severe, or non-clinical to severe symptoms) (see "Measures"; Jacobson & Truax, 1991). We then estimate and compare the prevalence of adverse events between the ASHF and PSM groups.

Results

We matched 217 ASHF participants with 434 control participants using a 2:1 nearest neighbor approach ($N = 651$). After matching, standardized mean differences for all covariates were below 0.10, indicating adequate balance (Fig. 3). There were no differences between the groups in race or gender distributions ($ps > 0.05$). Similarly, the distributions of undergraduate year (ASHF: 90.78% first-year, 3.23% second-year, 3.69% third-year, 2.3% ≥ fourth-year; PSM: 88.25% first-year, 2.53% second-year, 6.45% third-year, 2.76% ≥ fourth-year) and international student status (ASHF 6.45%; PSM 4.38%) were not different following matching. We observed no statistically significant differences (i.e., $p < 0.05$) at baseline on any variable using independent sample *t*-tests on continuous outcomes and chi-square tests on categorical outcomes following matching.

T2 missingness was significantly higher in the PSM control group than in the ASHF group $z = 2.03$, $p = 0.042$. ASHF T2 missingness ranged from 28.60% (mindfulness) to 30.00% (shared humanity, flourishing). PSM control group T2 missingness ranged from 36.60% (mindfulness) to 41.20% (compassionate roommate goals). Predicted between group standardized mean differences (i.e., Cohen's *d*s) for all outcomes are presented in Fig. 4.

Attention Skills

Compared to PSM controls, participation in the ASHF predicted significant post-test improvements on all three attention skills: attention function $b = 0.54$, $se = 0.15$,

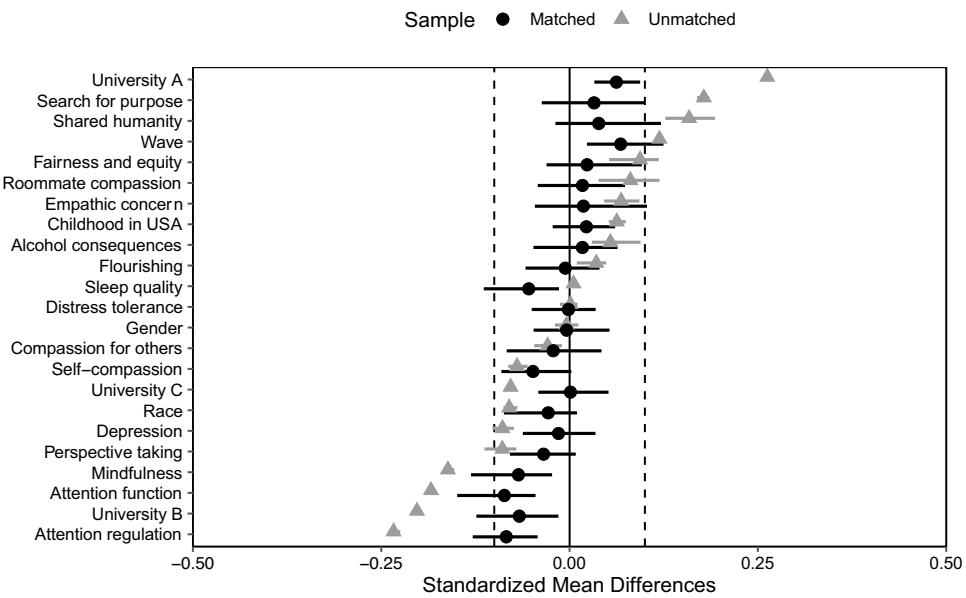

Fig. 3 Love plot of balance following propensity-score matching. *Note:* All variables are T1 scores. Standardized mean difference is between the ASHF and control group. Gray triangles: difference between the ASHF and control groups prior to matching. Black circles: difference between ASHF and the PSM control group (i.e., after matching) based on the 50 complete multiply imputed datasets. Black error bars represent the spread of imputed values across the 50 imputed datasets. Standardized mean differences between − 0.10 and 0.10 indicate adequate balance

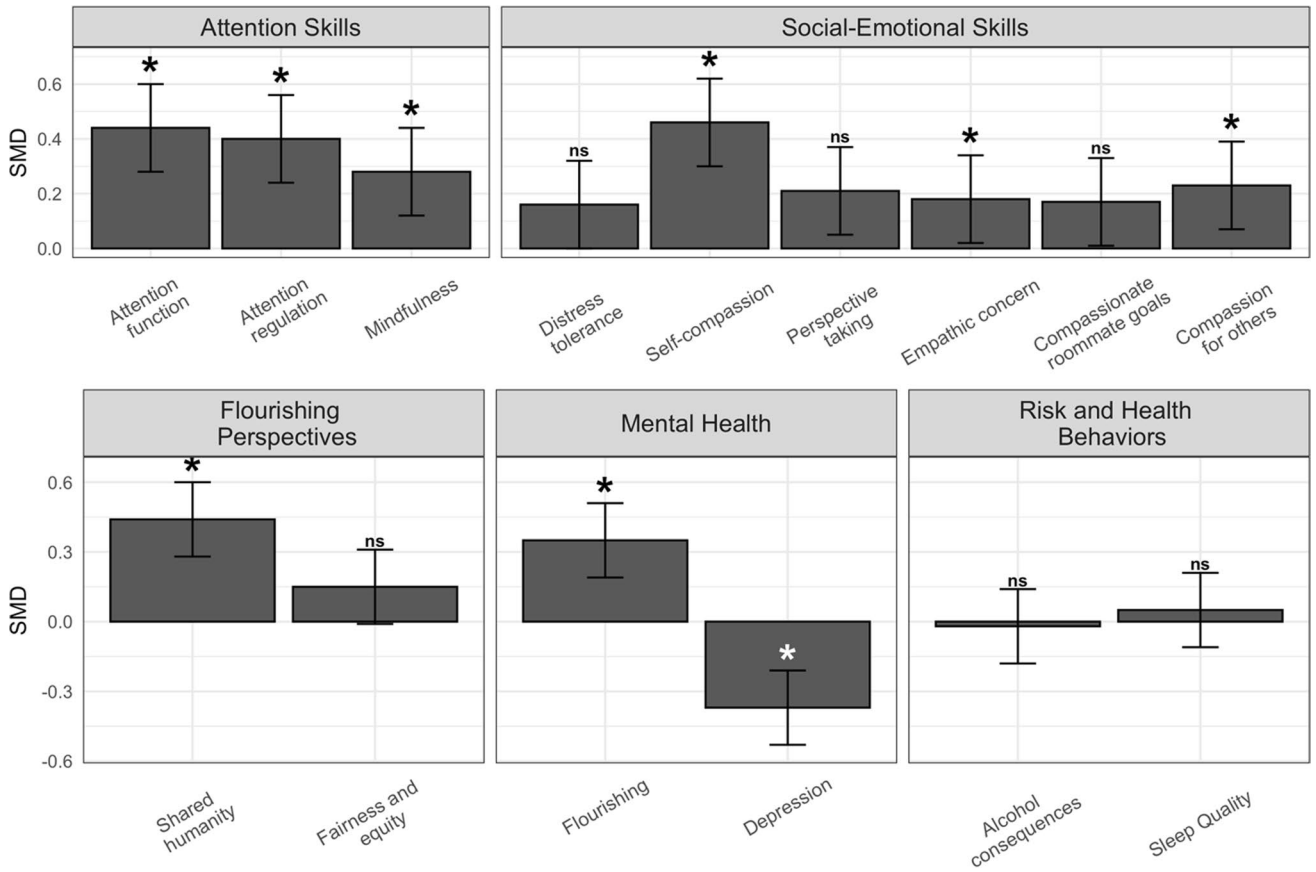

Fig. 4 Model-based post-test standardized mean group differences on all outcomes. *Note*: SMD, standardized mean difference between ASHF and PSM controls at post-test (T2). * $p < .05$ false discovery rate corrected (FDR). ns, not statistically significant ($p > .05$ FDR). Error bars: 95% confidence interval

$t(7.67) = 3.60$, $p_{FDR} = 0.018$, $d = 0.44$ CI[0.28, 0.61], unconditional ICC = 0.16; attention regulation $b = 037$, $se = 0.09$, $t(9.60) = 3.54$, $p_{FDR} = 0.016$, $d = 0.40$ CI[0.24, 0.57], ICC = 0.19; and mindfulness $b = 0.13$, $se = 0.05$, $t(9.46) = 2.52$, $p_{FDR} = 0.049$, $d = 0.28$ CI[0.12, 0.44], ICC = 0.07.

Social-Emotional Skills

Compared to PSM controls, participation in the ASHF predicted significant post-test improvements in self-compassion $b = 0.31$, $se = 0.08$, $t(9.46) = 3.94$, $p_{FDR} = 0.016$, $d = 0.46$ CI[0.29, 0.63], ICC = 0.07; empathic concern $b = 0.14$, $se = 0.06$, $t(404) = 2.16$, $p_{FDR} = 0.049$, $d = 0.18$ CI[0.02, 0.35], ICC = 0.03; and compassion for others $b = 0.15$, $se = 0.05$, $t(405) = 2.95$, $p_{FDR} = 0.016$, $d = 0.23$ CI[0.07, 0.39], ICC < 0.01. On distress tolerance, perspective taking, and compassionate roommate goals, the ASHF group made non-significant improvements relative to PSM controls ($b = 0.14$, $se = 0.09$, $t(9.65) = 1.49$, $p_{FDR} = 0.203$, $d = 0.16$ CI[− 0.00, 0.33], ICC = 0.06; $b = 0.17$, $se = 0.09$, $t(11.10) = 1.83$, $p_{FDR} = 0.131$, $d = 0.21$ CI[0.05, 0.37], ICC = 0.10; and $b = 0.17$, $se = 0.11$, $t(10.27) = 1.53$, $p_{FDR} = 0.203$, $d = 0.17$ CI[0.00, 0.33], ICC = 0.04, respectively).

Perspectives of Flourishing

ASHF participation predicted significantly larger gains in shared humanity $b = 0.39$, $se = 0.11$, $t(9.51) = 3.55$, $p_{FDR} = 0.011$, $d = 0.44$ CI[0.27, 0.61], ICC = 0.16 and non-significant improvements in fairness and equity $b = 0.11$, $se = 0.08$, $t(9.75) = 1.40$, $p_{FDR} = 0.412$, $d = 0.15$ CI[− 0.01, 0.31], ICC < 0.01.

Mental Health — Flourishing and Distress

Participation in the ASHF predicted significantly larger reductions in depressive symptoms $b = -1.97$, $se = 0.65$, $z = -3.04$, $p_{FDR} = 0.029$, $d = -0.37$ CI[− 0.54, − 0.21], ICC = 0.11 and improvements in flourishing $b = 0.55$,

$se = 0.15$, $z = 3.64$, $p_{FDR} = 0.002$, $d = 0.35$ CI[0.19, 0.51], ICC = 0.05 at T2. Enrollment in the ASHF also predicted a significantly lower likelihood of clinical and severe depressive symptoms at T2 ($b = -0.86$, $se = 0.27$, $z = -3.24$, $p_{FDR} = 0.011$, $OR = 0.42$ CI[0.25, 0.71], ICC = 0.02 and $b = -1.51$, $se = 0.45$, $z = -3.38$, $p_{FDR} < 0.001$, $OR = 0.22$ CI[0.09, 0.53], ICC = 0.03, respectively). The prevalence of clinical depression decreased nearly in half from 33% at pre-test to 17% at T2 in the ASHF group (PSM controls 29% at T1 and 23% at T2). The prevalence of severe depression decreased by two-thirds from 12% at pre-test to 4% at T2 in the ASHF group and did not change in PSM controls (11% at T1 to 10% at T2).

Physical Health and Risk Behavior

There were no significant group differences in sleep quality (ICC = 0.11) or alcohol consequences (ICC = 0.02, $ps > 0.10$).

Sensitivity Analyses

All significant improvements observed in the ASHF group under the MAR assumption were robust to the MNAR assumption that missing data was equivalent to the predicted samplewide T2 average (all $ps < 0.02$), with the exception of empathic concern ($p = 0.057$). Compassionate roommate goals, which in primary analyses did not show a significant effect favoring the ASHF, significantly favored the ASHF group under this assumption. All other significant improvements favoring the ASHF group observed in primary analyses, including empathic concern, were robust to all additional MNAR assumptions (all $ps < 0.05$). In addition, under the MNAR assumption that missing data were 0.20 standard deviations worse than predicted at T2, the ASHF group made significant improvements relative to PSM controls in perspective-taking and compassionate roommate goals ($ps < 0.05$). Under the MNAR assumption that missing data were 0.50 and 0.80 standard deviations worse than predicted at T2, the ASHF group demonstrated significant improvements relative to PSM controls in perspective-taking, compassionate roommate goals, and distress tolerance ($ps < 0.05$).

Adverse Change and ASHF Implementation Variability

Using observed data, 12 out of 151 (7.95%) participants in the ASHF group reported a categorical worsening in depressive symptoms from pre- to post-test compared to 61 out of 283 (21.56%) of PSM controls. In intention-to-treat analyses, adverse change was significantly more likely in PSM controls than in the ASHF group $b = -1.19$, $se = 0.35$, $z = -3.41$, $p < 0.001$.

Discussion

In a multi-wave, multi-site propensity-score matched controlled trial ($N = 651$) of a novel college course on human flourishing, we observed significant improvements on multiple attention and social-emotional skills, flourishing perspectives, and mental health and well-being outcomes. Specifically, students who enrolled in the ASHF course, relative to PSM control students, reported significant improvements in their attention function, attention regulation, mindfulness, self-compassion, empathic concern, compassion for others, sense of shared humanity, flourishing, and multiple metrics of depressive symptoms, including the prevalence of severe depression, at the end of the semester in which they took the course. While not reaching statistical significance, we observed small magnitude ASHF improvements in perspective-taking, compassionate roommate goals, fairness and equity, and distress tolerance as well. Taken together, these findings suggest that the ASHF class is a promising approach to supporting the mental health, attention and social-emotional skills, and flourishing perspectives of college students.

Colleges are struggling to address the growing mental health concerns among students. Based on a skill strengthening model of mental health and flourishing (Fig. 1; Dahl et al., 2020), we investigated whether a 15-week semester-long for-credit course that integrated rigorous academic learning about flourishing with experiential practice of flourishing skills would promote adaptation to college life. These data indicate that taking the ASHF, relative to a PSM control group, enhances a number of skills of relevance to college academic (e.g., attention function, attention regulation) and social (e.g., empathic concern, compassion, and shared humanity with others) life. In addition, enrollment in the ASHF predicted significantly greater flourishing and significantly reduced depressive symptoms at the end of the semester.

Considering the high prevalence of clinical and severe depression on college campuses, concerns about the COVID-19 pandemic exacerbating these problems, and the debilitating impact that depressive symptoms can have on students, ASHF effects on the prevalence of clinical and severe depression are noteworthy. Relative to the sample at baseline, enrollment in the ASHF reduced the prevalence of clinical depression by over 51% and severe depression by two-thirds. These benefits were accomplished in an academic setting in which students were earning credit toward undergraduate requirements. Whether ASHF effects on

depressive symptoms during the pandemic are consistent with the observed results will require additional research to determine, but it is clear that the on-going effects of the pandemic only serve to highlight the importance of programs that support college student mental health and flourishing. Also of importance, we find no evidence for higher rates of adverse change following the ASHF and preliminary evidence that the ASHF may be protective against increased depressive symptoms during the semester duration of this study. This research occurred during what was the first semester of college for the majority of participants, a time typically associated with deteriorations in mental health (Wyatt et al., 2017), not with the pronounced, multidimensional improvements in mental health and flourishing reported by ASHF students.

This study contributes to extant research on promoting college student flourishing in three major ways. First, the ASHF is unique in the degree to which it marries academic rigor with experiential learning. Second, the ASHF presentation of flourishing is based on a well-articulated theory of flourishing that is multidimensional, and study outcomes were selected to reflect the breadth of these domains. Third, the consortium that developed the ASHF was motivated to not only address the mental health crisis on college campuses but also to begin to recenter a holistic notion of student development as a defining feature of the undergraduate experience. This research suggests that courses such as the ASHF that emphasize academic and skills learning may help address the mental health crisis on college campuses by reducing symptoms (e.g., depression) while increasing flourishing. Furthermore, these data indicate that this approach may also support the development of attention skills critical to academic achievement and well-being (Moffitt et al., 2011) and perspectives essential to enriching social relationships and engagement with the social challenges of our time. Although our research design did not allow us to evaluate whether the observed improvements in psychological functioning translated into improvements in academic achievement or persistence in college, there is evidence that poorer mental health is associated with worse academic outcomes (Fletcher, 2010; Mojtabai et al., 2015). An important extension of this research is to examine whether, by improving the psychological functioning of college students, their longer-term academic and labor market outcomes also improve. Should future research find this to be the case, implementing a course like the ASHF during the first semester of college, as was done here, holds appeal as a prevention strategy as it may support adaptation to college life, improving outcomes during and after college.

We additionally note that many of the social-emotional skills the ASHF appears to improve are crucial to healthy human development and to functioning liberal democratic societies (e.g., shared humanity, perspective-taking, empathy). Recent events in America and other liberal democratic societies demonstrate the threat that heightened polarization and the inability to identify with others has on the functioning of government and the cohesiveness of society (Heltzel & Laurin, 2020). If a key purpose of public education (Mann, 1848) and undergraduate general education is to prepare a productive citizenry (Walker et al., 2009), education must purposefully strengthen skills and attitudes that allow individuals and liberal democratic societies to flourish.

Although these data support many aspects of our theory of change, including effects on the distal outcomes of flourishing and depression, we did not observe ASHF impacts on health and risk behavior outcomes. There are several potential explanations for these null effects. One possibility is that the ASHF course does not affect these outcomes, perhaps because these behaviors are strongly affected by social contexts occurring outside of academic classes. An alternative possibility is that effects on these outcomes occur over a longer-time course than the roughly 15-week duration of this study. For instance, strengthening of proximal skills may have immediate effects on depressive symptoms and flourishing, but their impact on social and health behaviors may only become apparent months or even years later. It is also possible that sample and measurement characteristics prevented detection of effects on these outcomes. This possibility is especially salient to our measure of risk behavior — alcohol consequences. Alcohol consequences were near floor at pre-test, leaving little room for improvement. Similarly, our sleep quality measure demonstrated poor reliability and high scores at baseline, suggesting that ceiling effects may have been at work. Future studies might include additional measures of key health and risk behavior, including objective measures of sleep behaviors (e.g., actigraphy) as well as a longitudinal follow-up to understand (a) whether sleeper effects on distal outcomes are present and (b) whether the observed effects on skills, mental health, and flourishing are durable over time.

Limitations and Future Research

The main limitation in this study is the strong set of assumptions that underly causal inference through PSM. We address this assumption in multiple ways. First, we include a large and multidimensional set of measured baseline variables into the PSM procedure. Second, we achieve doubly robust estimates by weighting outcome models after matching with the propensity score, a statistical approach that has been shown to provide unbiased treatment effect estimates if one or both of the PSM and outcome models are properly specified (Funk et al., 2011). Not directly an issue of causal inference, we model the data according to its clustered nature (i.e., waves, universities, groups), accounting for any variance in student outcomes explained by these higher-level clusters.

Despite the logistical challenges of randomly assigning college students to for-credit courses, designs such as large-scale cluster randomized controlled trials, ideally with long-term follow-ups assessing risk and health behaviors, afford strong causal claims about course impacts on proximal and distal outcomes. We predict that such a study would detect long-term ASHF effects on key risk and health behaviors.

A second limitation is the relatively high attrition rates which present a threat to internal validity. Our sensitivity analyses modeled effects under multiple MNAR assumptions, including the extreme assumption that missing data were 0.80 standard deviations worse than the observed data, and in nearly every case the inferences drawn from primary analyses remained valid. This fact provides confidence that our principal conclusions were not the result of attrition bias, but when considering this research and possible longitudinal future work, understanding and working to mitigate the causes of attrition are important areas for consideration.

A third limitation is that our sample is less diverse in terms of its race and gender composition than students nationwide (though not within the universities involved), limiting the generalizability of these results. As a result of the relative homogeneity of our sample, we collapsed multiple race and gender identifications into overarching categories. It is important to know whether ASHF effects differ based on specific gender or race identities. Further research with large and sufficiently diverse samples to examine impacts across multiple gender and racial identities, and their intersections, is needed. Although most ASHF students completed pre-test before exposure to core course content, this subset of ASHF students' pre-test data could have been affected by the exposure to class content. It is possible that average baseline values of ASHF students were improved as a result, potentially reducing the observed magnitude of ASHF group change and thereby the likelihood of detecting significant group differences.

Finally, the lack of ASHF standardization across universities and waves is both a limitation and a strength. These results indicate that on average, when implemented in real-world contexts, the ASHF produced multidimensional benefits in students. Prior research (Inkelas et al., 2021) indicated that some variability in ASHF course implementation does not affect outcomes, but it remains possible that our methods have not yet captured important differences in student outcomes based on ASHF implementation. The consortium's approach to ASHF implementation assumed that instructors needed considerable training in the academic and experiential content involved in the course to teach it with fidelity. We therefore predict that although some variability in course implementation is acceptable, too much elasticity has the potential to reduce the impact of the course. Future research may wish to interrogate this claim. Relatedly, optimizing the course by identifying the content or instructional practices that are most strongly associated with positive outcomes may require greater standardization in future implementations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12671-022-01952-1.

Author Contribution MJH: designed and executed the study, led the data analyses, and wrote the paper. BAC: designed and executed the study, assisted with the data analyses, and collaborated in the final writing and editing of the study. MTG: designed the study, assisted with the data analyses, and collaborated in the final writing and editing of the study. KKI: designed and executed the study, and collaborated in the final writing and editing of the study. RJD: designed the study, provided research infrastructure, and collaborated in the final writing and editing of the study. DG: provided research infrastructure and collaborated in the final writing and editing of the study. JDD: collaborated in the final writing and editing of the study. RWR: designed and executed the study, assisted with the data analyses, and collaborated in the final writing and editing of the study.

Funding This study was supported by generous individual donations to the Center for Healthy Minds, a 2019 National Academy of Education/Spencer Foundation Postdoctoral fellowship (first author), the Bennett Pierce Chair in Care and Compassion (last author), and the Contemplative Sciences Center at the University of Virginia.

Data Availability Due to complications arising from multiple ethics boards, the data involved in this study cannot be made publicly available.

Declarations

Ethics Approval The institutional review boards of the University of Wisconsin–Madison, the Pennsylvania State University, and the University of Virginia each approved the research described in this article. In addition, a certificate of confidentiality was obtained to protect participant privacy.

Informed Consent All participants provided written informed consent before study activities commenced.

Conflict of Interest Richard J. Davidson is the founder, president, and serves on the board of directors for the non-profit organization, Healthy Minds Innovations, Inc. No donors, either anonymous or identified, have participated in the design, conduct, or reporting of research results in this manuscript.

References

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., & Hasking, P. (2018). WHO world mental health surveys international college student project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology, 127*(7), 623.

Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. *Assessment, 15*(3), 329–342. https://doi.org/10.1177/1073191107313003

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological), 57*(1), 289–300.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Bryk, A. (2016). Fidelity of implementation: Is it the right concept. *Carnegie Commons Blog*, *17*. Available from: https://www.carnegiefoundation.org/blog/fidelity-of-implementation-is-it-the-right-concept/

Center for Collegiate Mental Health. (2020). Annual report. 40. Available from https://ccmh.psu.edu/annual-reports

Chiteji, N. (2010). Time-preference, non-cognitive skills and well-being across the life course: Do non-cognitive skills encourage healthy behavior? *The American Economic Review, 100*(2), 200–204.

Cimprich, B., Visovatti, M., & Ronis, D. L. (2011). The Attentional Function Index—A self-report cognitive measure. *Psycho-Oncology, 20*(2), 194–202.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Routledge.

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: The role of compassionate and self-image goals. *Journal of Personality and Social Psychology, 95*(3), 555.

Dahl, C. J., Lutz, A., & Davidson, R. J. (2015). Reconstructing and deconstructing the self: Cognitive mechanisms in meditation practice. *Trends in Cognitive Sciences, 19*(9), 515–523. https://doi.org/10.1016/j.tics.2015.07.001

Dahl, C. J., Wilson-Mendenhall, C. D., & Davidson, R. J. (2020). The plasticity of well-being: A training-based framework for the cultivation of human flourishing. *Proceedings of the National Academy of Sciences, 117*(51), 32197–32206. https://doi.org/10.1073/pnas.2014859117

Davis, M. H. (1980). *Interpersonal reactivity index*. Edwin Mellen Press.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*(1), 151–161.

Duffy, M. E., Twenge, J. M., & Joiner, T. E. (2019). Trends in mood and anxiety symptoms and suicide-related outcomes among US undergraduates, 2007–2018: Evidence from two national surveys. *Journal of Adolescent Health, 65*(5), 590–598. https://doi.org/10.1016/j.jadohealth.2019.04.033

Dvořáková, K., Greenberg, M. T., & Roeser, R. W. (2019). On the role of mindfulness and compassion skills in students' coping, well-being, and development across the transition to college: A conceptual analysis. *Stress and Health, 35*(2), 146–156. https://doi.org/10.1002/smi.2850

Dvořáková, K., Kishida, M., Li, J., Elavsky, S., Broderick, P. C., Agrusti, M. R., & Greenberg, M. T. (2017). Promoting healthy transition to college through mindfulness training with first-year college students: Pilot randomized controlled trial. *Journal of American College Health, 65*(4), 259–267. https://doi.org/10.1080/07448481.2017.1278605

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement, 61*(5), 713–740.

Ettman, C. K., Abdalla, S. M., Cohen, G. H., Sampson, L., Vivier, P. M., & Galea, S. (2020). Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Network Open, 3*(9), e2019686. https://doi.org/10.1001/jamanetworkopen.2020.19686

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*(1), 43–53. https://doi.org/10.1037/a0014699

Fletcher, J. M. (2010). Adolescent depression and educational attainment: Results using sibling fixed effects. *Health Economics, 19*(7), 855–871.

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology, 173*(7), 761–767. https://doi.org/10.1093/aje/kwq439

Galante, J., Galante, I., Bekkers, M.-J., & Gallacher, J. (2014). Effect of kindness-based meditation on health and well-being: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology, 82*(6), 1101–1114. https://doi.org/10.1037/a0037249

Gall, T. L., Evans, D. R., & Bellerose, S. (2000). Transition to first-year university: Patterns of change in adjustment across life domains and time. *Journal of Social and Clinical Psychology, 19*(4), 544–567. https://doi.org/10.1521/jscp.2000.19.4.544

Gilbert, P., Catarino, F., Duarte, C., Matos, M., Kolts, R., Stubbs, J., Ceresatto, L., Duarte, J., Pinto-Gouveia, J., & Basran, J. (2017). The development of compassionate engagement and action scales for self and others. *Journal of Compassionate Health Care, 4*(1), 4. https://doi.org/10.1186/s40639-017-0033-3

Goldberg, S. B., Bolt, D. M., & Davidson, R. J. (2021). Data missing not at random in mobile health research: Assessment of the problem and a case for sensitivity analyses. *Journal of Medical Internet Research, 23*(6), e26749. https://doi.org/10.2196/26749

Goldberg, S. B., Riordan, K. M., Sun, S., & Davidson, R. J. (2022). The empirical status of mindfulness-based interventions: A systematic review of 44 meta-analyses of randomized controlled trials. *Perspectives on Psychological Science, 17*(1), 108–130.

Heltzel, G., & Laurin, K. (2020). Polarization in America: Two possible futures. *Current Opinion in Behavioral Sciences, 34*, 179–184. https://doi.org/10.1016/j.cobeha.2020.03.008

Hervás, G., & Vázquez, C. (2013). Construction and validation of a measure of integrative well-being in seven languages: The Pemberton Happiness Index. *Health and Quality of Life Outcomes, 11*(1), 66.

Hirshberg, M. J., Goldberg, S. B., Rosenkranz, M., & Davidson, R. J. (2020). Prevalence of harm in mindfulness-based stress reduction. *Psychological Medicine, 52*(6), 1080–1088. https://doi.org/10.1017/S0033291720002834

Hirshberg, M. J., Goldberg, S. B., Rosenkranz, M., & Davidson, R. J. (2021). Response to Van Dam and Galante's correspondence 'Underestimating harm in mindfulness-based stress reduction.' *Psychological Medicine*, 1–3.

Hirshberg, M. J., Goldberg, S. B., Schaefer, S. M., Flook, L., Findley, D., & Davidson, R. J. (2018). Divergent effects of brief contemplative practices in response to an acute stressor: A randomized controlled trial of brief breath awareness, loving-kindness, gratitude or an attention control practice. *PLoS ONE, 13*(12), e0207765.

Huang, F. L. (2018a). Using cluster bootstrapping to analyze nested data with a few clusters. *Educational and Psychological Measurement, 78*(2), 297–318.

Huang, F. L. (2018b). Multilevel modeling and ordinary least squares regression: How comparable are they? *The Journal of Experimental Education, 86*(2), 265–281. https://doi.org/10.1080/00220973.2016.1277339

Huberty, J., Green, J., Glissmann, C., Larkey, L., Puzia, M., & Lee, C. (2019). Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: Randomized controlled trial. *JMIR MHealth and UHealth, 7*(6), e14273.

Inkelas, K. K., Colaianne, B. A., Hirshberg, M. J., Greenberg, M. T., Davidson, R. J., Dunne, J. D., Germano, D., & Roeser, R. W. (2021).

Does variability across three universities in the implementation of a college course on human flourishing affect student outcomes? Journal of American College Health, 1–14. https://doi.org/10.1080/07448481.2021.1920956

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12.

Kahler, C. W., Strong, D. R., & Read, J. P. (2005). Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: The Brief Young Adult Alcohol Consequences Questionnaire. *Alcoholism Clinical and Experimental Research, 29*(7), 1180–1189.

Kazdin, A. E. (1993). Adolescent mental health: Prevention and treatment programs. *American Psychologist, 48*(2), 127.

Keeling, R. (2013). *Learning reconsidered: A campus-wide focus on the student experience*: NASPA Publications

Kirby, J. N. (2017). Compassion interventions: The programmes, the evidence, and implications for research and practice. *Psychology and Psychotherapy: Theory, Research and Practice, 90*(3), 432–455. https://doi.org/10.1111/papt.12104

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders, 114*(1–3), 163–173.

Lang, D. (2007). The impact of a first-year experience course on the academic performance, persistence, and graduation rates of first-semester college students at a public research university. *Journal of the First-Year Experience & Students in Transition, 19*(1), 9–25.

Lerner, R. M., Roeser, R. W., & Phelps, E. (2008). *Positive youth development & spirituality: From theory to research*. Templeton Foundation Press West Conshohocken.

Lindahl, J. R., Fisher, N. E., Cooper, D. J., Rosen, R. K., & Britton, W. B. (2017). The varieties of contemplative experience: A mixed-methods study of meditation-related challenges in Western Buddhists. *PLoS ONE, 12*(5), e0176239. https://doi.org/10.1371/journal.pone.0176239

Lipson, S. K., Lattie, E. G., & Eisenberg, D. (2018). Increased rates of mental health service utilization by U.S. college students: 10-year population-level trends (2007–2017). *Psychiatric Services, 70*(1), 60–63. https://doi.org/10.1176/appi.ps.201800332

Mann, H. (1848). Twelfth annual report to the Massachusetts Board of Education. *The Republic and the School Horace Mann and the Education of Free Men New York Teachers College Press Retrieved November, 12*, 2006.

McNeish, D., & Stapleton, L. M. (2016). Modeling clustered data with very few clusters. *Multivariate Behavioral Research, 51*(4), 495–518. https://doi.org/10.1080/00273171.2016.1167008

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartmess, E., & Stewart, A. (2012). The Multidimensional Assessment of Interoceptive Awareness (MAIA). *PLoS ONE, 7*(11), e48230. https://doi.org/10.1371/journal.pone.0048230

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*(7), 2693–2698.

Mojtabai, R., Stuart, E. A., Hwang, I., Susukida, R., Eaton, W. W., Sampson, N., & Kessler, R. C. (2015). Long-term effects of mental disorders on employment in the National Comorbidity Survey ten-year follow-up. *Social Psychiatry and Psychiatric Epidemiology, 50*(11), 1657–1668.

Orben, A., Tomova, L., & Blakemore, S.-J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health, 4*(8), 634–640. https://doi.org/10.1016/S2352-4642(20)30186-3

Palmer, P. J., Zajonc, A., & Scribner, M. (2010). *The heart of higher education: A call to renewal*. Wiley.

Piedmont, R. L. (1999). Does spirituality represent the sixth factor of personality? Spiritual transcendence and the five-factor model. *Journal of Personality, 67*(6), 985–1013. https://doi.org/10.1111/1467-6494.00080

Pishgar, F., Greifer, N., Leyrat, C., & Stuart, E. (2020). MatchThem: Matching and weighting after multiple imputation. *arXiv preprint* arXiv:2009.11772.

R Core Team (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012*. ISBN 3–900051–07–0.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the self-compassion scale. *Clinical Psychology & Psychotherapy, 18*(3), 250–255.

Renn, K. A., & Reason, R. D. (2012). *College students in the United States: Characteristics, experiences, and outcomes*. Wiley.

Roisman, G. I., Masten, A. S., Coatsworth, J. D., & Tellegen, A. (2004). Salient and emerging developmental tasks in the transition to adulthood. *Child Development, 75*(1), 123–133. https://doi.org/10.1111/j.1467-8624.2004.00658.x

Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*(4), 719–727.

Santos, Laurie. (2018). Yale University Psychology 157: *Psychology and the good life*. Available from: https://www.psychologyandthegoodlife.com/

Seligman, M. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology, 13*(4), 333–335.

Shapiro, S. L., Brown, K. W., & Astin, J. (2011). Toward the integration of meditation into higher education: A review of research evidence. *Teachers College Record: The Voice of Scholarship in Education, 113*(3), 493–528. https://doi.org/10.1177/016146811111300306

Smeets, E., Neff, K., Alberts, H., & Peters, M. (2014). Meeting suffering with kindness: Effects of a brief self-compassion intervention for female college students: Self-compassion intervention for students. *Journal of Clinical Psychology, 70*(9), 794–807. https://doi.org/10.1002/jclp.22076

Totzeck, C., Teismann, T., Hofmann, S. G., von Brachel, R., Pflug, V., Wannemüller, A., & Margraf, J. (2020). Loving-kindness meditation promotes mental health in university students. *Mindfulness, 11*(7), 1623–1631. https://doi.org/10.1007/s12671-020-01375-w

US Department of Education. National Center for Educational Statistics. (2019). *Historical statistics of the United States, colonial times to 1970; Current Population Reports. Current Population Survey (CPS), October, 1970 through 2017*. U.S. Department of Commerce.

Walker, D. F., Soltis, J. F., & Schoonmaker, F. (2009). *Curriculum and aims* (5th edition). Teachers College Press.

Wyatt, T. J., Oswalt, S. B., & Ochoa, Y. (2017). Mental health and academic performance of first-year college students. *International Journal of Higher Education, 6*(3), 178–187.

Xiao, H., Carney, D. M., Youn, S. J., Janis, R. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2017). Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychological Services, 14*(4), 407–415.