

A Six-year Longitudinal Dataset of College Student Well-being, Daily Experience, and Academic Performance

HAN ZHANG, University of Chicago, USA

YASAMAN SEFIDGAR, Texas A&M University , USA

XUHAI XU, Columbia University, USA

YIYI REN, University of Washington, USA

KEVIN KUEHN, University of California San Diego, USA

WOOSUK SEO, Yale University, USA

PAULA NURIUS, University of Washington, USA

MARGARET E. MORRIS, University of Washington, USA

EVE A. RISKIN, Stevens Institute of Technology, USA

ANIND K. DEY, University of Washington, USA

JENNIFER MANKOFF, University of Washington, USA

The University of Washington Student Experience study is a six-year longitudinal dataset of undergraduate students collected from 2018 to 2023. The dataset comprises 765 unique participants (1,234 person-years) collected through four complementary modalities: annual demographic surveys; pre and post-quarter surveys assessing psychological well-being, social functioning, academic experience, and health behaviors using 32 validated measures and team-developed questionnaires; ecological momentary assessments capturing momentary affect, stress, coping, and daily behavior; and academic performance records obtained from the university registrar. Data span six annual cohorts with rolling enrollment, enabling both within-person longitudinal and cross-sectional analyses. We release the dataset in three access tiers: open-access for non-sensitive variables; restricted (under a data use agreement) for sensitive variables; and by-request (managed) access for potentially identifying variables—demographics, academic performance records, and the cross-year participant linkage. A complete codebook and data cleaning scripts are provided; the cross-year participant identifier mapping is available by request.

Additional Key Words and Phrases: College student; mental health; well-being; academic performance; longitudinal dataset

Draft citation:

Han Zhang, Yasaman Sefidgar, Xuhai Xu, Yiyi Ren, Kevin Kuehn, Woosuk Seo, Paula Nurius, Margaret E. Morris, Eve A. Riskin, Anind K. Dey, and Jennifer Mankoff. 2026. A Six-year Longitudinal Dataset of College Student Well-being, Daily Experience, and Academic Performance. Manuscript in preparation.

Authors' Contact Information: Han Zhang, University of Chicago, USA, micohan@cs.uchicago.edu; Yasaman Sefidgar, Texas A&M University , USA, ; Xuhai Xu, Columbia University, USA, xx2489@columbia.edu; Yiyi Ren, University of Washington, USA, yiyiren92@gmail.com; Kevin Kuehn, University of California San Diego, USA, kskuehn@health.ucsd.edu; Woosuk Seo, Yale University, USA, woosuk.seo@yale.edu; Paula Nurius, University of Washington, USA, nurius@uw.edu; Margaret E. Morris, University of Washington, USA, margie.mem@gmail.com; Eve A. Riskin, Stevens Institute of Technology, USA, eriskin@stevens.edu; Anind K. Dey, University of Washington, USA, anind@uw.edu; Jennifer Mankoff, University of Washington, USA, jmankoff@cs.washington.edu.

1 Background & Summary

The college years represent a critical period of emerging adulthood marked by elevated risk for onset of mental health conditions, increased exposure to academic stressors, and significant shifts in health behaviors [2, 3, 8, 17, 25, 32]. Epidemiological evidence indicates that approximately one-third of university students meet criteria for at least one common mental health condition [4], yet the longitudinal pathways through which disability, health, and academic success interact with contextual stressors over the course of undergraduate study remain poorly understood. This gap is compounded by students' diverse and intersecting social identities (e.g., race, gender, socioeconomic background), which shape exposure to stressors and access to resources in ways that aggregate-level analyses obscure [25, 29]. Understanding these dynamics requires data that follow the same individuals over time, capture both retrospective and real-time assessments of psychological and behavioral functioning, and are sufficiently rich in demographic detail to support intersectional analyses.

Several existing datasets have examined related questions in college populations. The Healthy Minds Study [25] is an annual web-based survey spanning 373 U.S. campuses with over 350,000 respondents from 2013 to 2021, documenting worsening mental health trends and persistent racial and ethnic disparities in help-seeking. The StudentLife dataset [39] and the College Experience Study [28] capture longitudinal passive mobile sensing and ecological momentary assessments (EMAs) from cohorts of Dartmouth College students over one term and five years respectively, enabling behavioral modeling of student mental health. However, each of these datasets captures only a subset of the dimensions needed to characterize student well-being comprehensively. A key gap therefore remains: no existing publicly available dataset combines multi-year within-person longitudinal tracking, intensive EMAs, a comprehensive set of validated psychological measures, academic performance records, and data collected from students with diverse backgrounds.

The University of Washington Student Experience (UWEXP) study was designed to address this gap. Conducted from 2018 to 2023 at the University of Washington (UW)—which operates on a quarter system dividing the academic year into fall, winter, and spring terms of roughly 10 weeks each—UWEXP is a six-year longitudinal cohort study of undergraduate students spanning the pre-pandemic, COVID-19 pandemic, and post-pandemic periods, capturing psychological well-being, daily experiences and behaviors, and academic performance. A companion dataset from the same research group—GLOBEM [48]—has previously released passively sensed behavioral data collected from smartphones and wearable devices for an overlapping subset of participants (2018–2021). The present dataset releases the complementary self-report layer, comprising six years of psychological assessments (established measures supplemented by team-developed measures), EMAs, and academic performance records limited to GPA—the only registrar variables released. In total, 765 unique students participated across the full study period, representing 1,234 person-years of data, with 67 followed for four years.

Data were collected through four complementary modalities each annual cycle. A demographic survey captured students' sociodemographic characteristics (e.g., race, gender, disability status, and immigration history), academic context, and life circumstances. Pre- and post-quarter surveys assessed a broad range of psychological, behavioral, and contextual domains—including mental health, social functioning, academic experience, substance use, adversity and discrimination, and life context—using 32 validated measures and team-developed questionnaires. EMAs captured momentary affect, stress, coping, social interactions, and daily behavior in near real-time, combining twice-weekly surveys across the quarter with an intensive four-times-daily burst over two weeks. Academic performance records were obtained directly from the university registrar each year. Together, these modalities yield a multimodal self-report

dataset released in three access tiers¹: an open-access tier for non-sensitive variables; a restricted tier for sensitive variables, available to qualified researchers under a data use agreement (DUA); and a by-request (managed) tier holding potentially identifying variables—demographics, GPA, and the cross-year participant linkage—available under a DUA and current Collaborative Institutional Training Initiative (CITI) certification.

These data have supported a range of published investigations, including studies of student distress, marginality, and resilience [29]; the impact of COVID-19 and remote learning on general population of college students [27], and students with disabilities and mental health concerns [49]; passively-sensed behavioral correlates of discrimination [33]; personalized depression detection [45, 46]; longitudinal depressive-behavior modeling generalization [47]; and early prediction of academic performance [50]. The combination of within-person longitudinal tracking, intensive EMAs, validated psychological measures, and registrar-verified academic records makes UWEXP suited to future investigations across psychology, public health, education, social science, and computer science. For example, researchers can model how momentary stress and affect (EMAs) related to end-of-quarter well-being (pre/post surveys) and academic outcomes (GPA) within the same individuals; examine how psychological trajectories diverge across demographic subgroups; or compare pre-pandemic, pandemic, and post-pandemic cohorts to study the effects of a major societal disruption on emerging adults. The dataset can also serve methodological work in missing data handling, EMA compliance modeling, and cross-dataset generalization when linked to GLOBEM.

2 Methods

2.1 Study Design

We employed a longitudinal cohort design with rolling enrollment. The 2018 data collection served as a pilot year, during which we developed and refined measures and protocols. Primary longitudinal data collection began in spring 2019 and continued each spring quarter through 2023, with each annual cycle corresponding to a 10-week academic quarter. In June 2020, an additional cohort was recruited concurrent with the spring 2020 data collection to supplement COVID-19-related assessments.

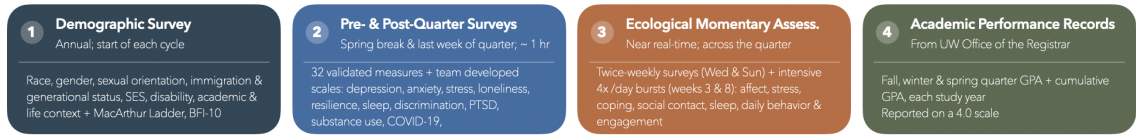
Data were collected through four complementary modalities each cycle: a demographic survey, pre- and post-quarter surveys, and EMAs, all administered online via Qualtrics, and academic performance records obtained directly from the UW Office of the University Registrar. We adapted study measures across cycles to reflect emerging contexts, most notably the addition of pandemic-specific measures beginning in 2020. As the 2018 cohort served as a pilot year, certain measure versions and item sets in 2018 differ from subsequent years. The complete codebook (`UWEXP_codebook.xlsx`), openly available in the deposit, is the authoritative reference for all measures: it provides the exact item wording presented to participants, response options, subscale and composite scoring procedures, the access tier of each variable, and year-by-year availability and version differences for every measure across all data types. Readers seeking the precise content of any measure described below should consult the codebook directly.

2.2 Ethics Statement

This study was reviewed and approved by UW’s Institutional Review Board (Study#: 00003244). All participants provided written informed consent prior to enrollment and were free to withdraw at any time without penalty. Given the sensitive nature of study measures, participants were provided with information about campus resources including mental health services, victim advocacy and support services, and financial crisis assistance at each data collection cycle.

¹To be deposited in ICPSR upon acceptance.

FOUR DATA MODALITIES



STUDY TIMELINE, ZOOMED INTO ONE 10-WEEK QUARTER

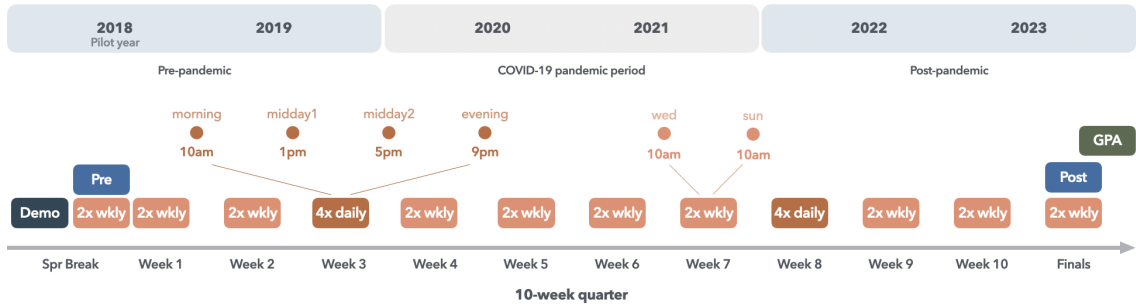


Fig. 1. **Overview of the UWEXP data collection design.** The dataset comprises four complementary modalities (top): (1) an annual demographic survey at the start of each cycle; (2) pre- and post-quarter surveys (about 1 hour each) of 32 validated and team-developed measures; (3) ecological momentary assessments (EMAs) capturing affect, stress, coping, sleep, and daily behavior in near real-time; and (4) annual academic performance records (quarterly and cumulative GPA) from the University of Washington registrar. The timeline (middle) spans six annual cohorts (2018–2023), with 2018 as a pilot year, across pre-pandemic, COVID-19, and post-pandemic periods. The lower panel expands a single 10-week spring quarter: demographic and pre-quarter surveys at the start and the post-quarter survey at Finals; twice-weekly EMAs (“2x wkly,” Wednesdays and Sundays at 10am) across the quarter; and four-times-daily EMA bursts (“4x daily,” at 10am, 1pm, 5pm, and 9pm) during weeks 3 and 8. Note that the exact EMA window varied by cohort—most also collected during Spring Break (all years except 2019) and about half during Finals week; see the codebook and Figure 5. GPA spans the full academic year and is not tied to a within-quarter timepoint.

Data collected from participants who withdrew were excluded from the dataset. All data were de-identified prior to deposit, with original participant IDs replaced by randomized 16-character alphanumeric identifiers. To protect participant privacy, the dataset is deposited in three access tiers, as detailed in the relevant subsections below. All research staff with direct access to participant data completed human subjects research training through the CITI program prior to data collection. Data collection staff were additionally briefed on research ethics protocols, including informed consent procedures and data privacy requirements.

2.3 De-identification and Disclosure Risk

To limit re-identification risk, the open tier excludes all direct and indirect identifiers. A k-anonymity analysis on the deposited open-tier files confirmed every file exceeds $k \geq 11$ (minimum $k = 153$). Demographic combinations are the primary identifiability risk: gender + race + income alone place 30–44% of each cohort in equivalence classes smaller than 5, so all demographics, GPA, and the cross-year linkage were moved to the by-request (managed) tier, where any release is held to a $k \geq 5$ suppression floor. Free-text responses (the only response type that can carry a direct identifier) are removed from every tier.

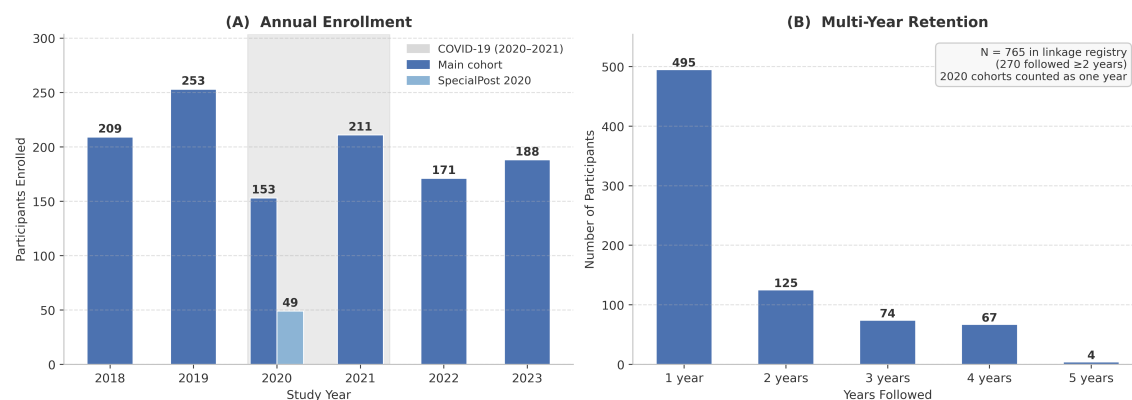


Fig. 2. **Participant Enrollment and Retention.** (A) Annual enrollment by cohort year (2018–2023). Dark blue bars represent the primary longitudinal cohort; light blue bar represents the SpecialPost 2020 cohort recruited concurrently with the 2020 cohort. The COVID-19 pandemic period (2020–2021) is indicated in light gray. (B) Multi-year retention among the 765 unique participants in the cross-year PID mapping followed for one through five years. A total of 270 participants were followed for two or more years.

2.4 Participants

To be eligible, participants were required to be enrolled as a full-time undergraduate student at the UW and to be 18 years of age or older. Recruitment was open to all eligible students; we intentionally recruited students with diverse demographic backgrounds to ensure representation across all student groups. Each year, participants were recruited through campus flyers, university mailing lists, and student social media groups. The 2018 cohort entered the study as incoming first-year students; subsequent cohorts included both newly enrolled participants and returning students from prior years. All participants provided written informed consent prior to enrollment. Compensation was provided in the form of cash or check, with participants able to earn up to \$245 per year depending on their completion rate. Study staff monitored participation rates throughout each data collection period and proactively contacted participants who missed surveys to confirm continued engagement. Participants received reminders prior to the final quarterly survey and met with study staff in person to receive payment.

In June 2020, concurrent with the spring 2020 post-quarter survey, we recruited an additional cohort of 49 participants (the SpecialPost 2020 cohort) through the same eligibility criteria and recruitment channels as the primary cohort, specifically to supplement COVID-19-related assessments. Recruitment of this supplemental cohort was initiated in response to the onset of the COVID-19 pandemic, with the goal of capturing COVID-19-related experiences. The target sample size of approximately 50 participants was determined based on the capacity of study team to onboard and support new participants mid-cycle while maintaining data quality standards for the existing cohort; it was not intended to power standalone subgroup analyses but rather to supplement the primary cohort's COVID-19 data and enable cohort comparisons. These participants were carried forward into subsequent annual data collection cycles alongside returning participants from prior cohorts.

Overall, enrollment and retention across the six study years are shown in Figure 2. Annual enrollment ranged from 171 to 253 participants across all cohort years, representing 1,234 person-years of observation across 765 unique participants. Of these, 270 participants were followed for two or more years, with 125 followed for two years, 74 for three years, 67 for four years, and 4 for five years.

Table 1. Sample demographic characteristics by cohort year. Values are N (%) of enrolled participants in each cohort year. SP2020 = SpecialPost 2020 cohort (N = 49), recruited concurrently with the 2020 main cohort. LGBTQ+ includes all participants who did not identify as heterosexual. URM = underrepresented racial or ethnic minority (Black, Latinx, and Native American students). SES = socioeconomic status. ‘—’ indicates data not collected in that year.

Demographics	2018	2019	2020 Cohort		2021	2022	2023	Overall
			2020	SP2020				
N enrolled	209	253	153	49	211	171	188	1,234
Woman	133 (63.6%)	126 (49.8%)	86 (56.2%)	33 (67.3%)	128 (60.7%)	104 (60.8%)	120 (63.8%)	730 (59.2%)
Asian	97 (46.4%)	123 (48.6%)	84 (54.9%)	35 (71.4%)	112 (53.1%)	91 (53.2%)	99 (52.7%)	641 (51.9%)
White	64 (30.6%)	79 (31.2%)	43 (28.1%)	11 (22.4%)	56 (26.5%)	38 (22.2%)	41 (21.8%)	332 (26.9%)
Race-based URM	35 (16.7%)	32 (12.6%)	17 (11.1%)	3 (6.1%)	28 (13.3%)	29 (17.0%)	36 (19.1%)	180 (14.6%)
Biracial / multiracial	28 (13.4%)	31 (12.3%)	11 (7.2%)	3 (6.1%)	19 (9.0%)	21 (12.3%)	19 (10.1%)	132 (10.7%)
LGBTQ+	23 (11.0%)	25 (9.9%)	25 (16.3%)	8 (16.3%)	48 (22.7%)	35 (20.5%)	57 (30.3%)	221 (17.9%)
First-generation student	62 (29.7%)	76 (30.0%)	51 (33.3%)	18 (36.7%)	72 (34.1%)	49 (28.7%)	75 (39.9%)	403 (32.7%)
International student	16 (7.7%)	33 (13.0%)	18 (11.8%)	5 (10.2%)	20 (9.5%)	11 (6.4%)	15 (8.0%)	118 (9.6%)
Low-income household ^a	—	31 (15.2%)	25 (19.8%)	4 (11.8%)	30 (17.6%)	16 (11.5%)	21 (14.1%)	127 (12.4%)
Reported a disability	—	23 (9.1%)	21 (13.7%)	9 (18.4%)	16 (7.6%)	15 (8.8%)	16 (8.5%)	100 (9.8%)
SES, mean (SD) ^b	—	—	5.74 (1.82)	5.85 (1.75)	5.77 (1.85)	5.93 (1.84)	5.75 (1.85)	5.80 (1.84)

^a Annual household income below \$25,000.

^b SES available from 2020 onward only; Overall SES computed over 2020–2023 and SP2020 (N = 772).

2.5 Demo Surveys

Demographic information was collected annually via a dedicated survey administered at the start of each data collection cycle (2018–2023). Variables include: sociodemographic characteristics (gender, sexual orientation, age, race and ethnicity, immigration history, and generational status, all spanning 2018–2023; sex assigned at birth, 2019–2023); socioeconomic background (household income, parental education, family size, financial aid receipt, and first-generation college student status); academic context (college, major, and enrollment status); and life context (housing type, living arrangements, relationship status, employment status and hours, and home languages). We assessed disability status via detailed multi-select items covering sensory, mobility, learning, mental health, and other impairments; registration with UW Disability Resources for Students (DRS) was also recorded (2019–2023). We additionally administered the MacArthur Community Ladder [1] for subjective socioeconomic status (SES, 2018, 2020–2023) and the Big Five Inventory-10 (BFI-10 [30]) for personality (2018–2023). Item availability varied across cohort years and is documented in the accompanying codebook.

Demographics of enrolled participants are summarized in Table 1. Across all cohort years, 59.2% of participants identified as women. The sample was racially and ethnically diverse, with 51.9% identifying as Asian, 26.9% as White, 14.6% as underrepresented racial or ethnic minorities (including Black and Latinx students), and 10.7% as biracial or multiracial; no participants identified as Native American. Approximately 17.9% identified as LGBTQ+, with this proportion increasing notably across cohort years (9.9% in 2019 to 30.3%). Additionally, 32.7% were first-generation college students, 9.6% were international students, 9.8% reported a disability, 12.4% came from low-income household (annual household income below \$25,000), and the average SES was 5.80 out of 10 (where higher values indicate higher perceived social standing). The SpecialPost 2020 cohort (N = 49) was notably more Asian (71.4%), had a higher disability rate (18.4%), and lower underrepresented minority representation (6.1%) compared to the primary cohort. All demographic variables are held in the by-request (managed) tier—available to qualified researchers under a DUA and current CITI certification—and are not included in the open or restricted deposit.

2.6 Pre and Post Surveys

We administered pre- and post-quarter surveys during the first and last weeks of each spring quarter, respectively, and surveys took approximately one hour to complete. In 2018, surveys were additionally administered during the pilot winter quarter. Participants received an email notification with a personalized survey link at the start of each administration window and were given approximately one week to complete the survey. Study coordinators sent reminder notifications to non-respondents midway through the window.

We assessed a broad range of psychological, behavioral, and contextual domains using both established validated measures and measures developed by our research team. Validated measures covered mood and depression, anxiety and stress, social support and belonging, emotion regulation and rumination, resilience and coping, mindfulness, sleep quality, wellbeing and physical health, academic self-efficacy, substance use, adversity and discrimination, posttraumatic stress, and personality. Table 2 summarizes the validated and team-developed measures administered in the pre- and post-quarter surveys, including their abbreviations, item counts, score ranges, years of availability, and access tier. To limit disclosure of specific sensitive responses, every validated scale that has a composite is released as its composite/subscale score only; the exact item wording is retained in the codebook for transparency, but item-level responses are not deposited. Full item text and the specific scoring procedures for each measure, including the exact items used for measures that changed version across years, are provided in the codebook.

Certain measures varied in version or length across cohort years for several reasons: adaptation to emerging contextual factors (e.g., COVID-19 pandemic measures, accessibility barrier items), reduction of participant burden, or correction of missing items identified in earlier administrations. Affected measures include the CESD (9-, 10-, and 20-item versions), PSS (10- and 14-item), UCLA Loneliness Scale (10- and 20-item), MAAS (7- and 15-item), and 2-Way SSS (21- and 22-item); composite scores were calculated consistently across versions to ensure cross-year comparability. Not all measures were administered in every study year. Full details including item text, variable names, subscales, scoring procedures, and year-by-year availability are provided in the accompanying codebook.

Table 2. **Summary of Pre- and Post-Quarter Survey Measures.** DUA = Data Use Agreement. “–” represents that composite scores are not calculated. For every measure that has a composite, only the composite/subscale score is released; the individual item responses are documented in the codebook but are not deposited in any tier.

Measure	Abbrev.	Items	Score Range	Years Available	Tier
Mood & Depression					
Beck Depression Inventory-II [5]	BDI2	21	0–63	2018–2023	Public
Center for Epidemiologic Studies Depression Scale [16]	CESD	9, 10, or 20	0–27 / 0–30 / 0–60	2018–2023	Public
Anxiety & Stress					
State-Trait Anxiety Inventory (State) [22]	STAI-S	20	20–80	2018	Public
State-Trait Anxiety Inventory (Trait) [6]	STAI-T	20	20–80	2019–2023	Public
Perceived Stress Scale [13]	PSS	10 or 14	0–40 / 0–56	2018–2023	Public
Social Support & Belonging					
UCLA Loneliness Scale [31]	UCLA	10 or 20	10–40 / 10–80	2018–2023	Public
2-Way Social Support Scale [34]	2WaySSS	21 or 22	0–105	2018–2023	Public
Interpersonal Support Evaluation List [14]	ISEL12	12	12–48	2018, 2022–2023	Public
Social Fit Scale [38]	SocialFit	17	17–119	2018–2023	Public
Emotion Regulation & Rumination					

Continued on next page
Draft manuscript

Table 2 continued

Measure	Abbrev.	Items	Score Range	Years Available	Tier
Emotion Regulation Questionnaire [21]	ERQ	10	1–7	2018–2023	Public
Rumination-Reflection Questionnaire [36]	RRQ	24	1–5	2019–2023	Public
Response Styles Questionnaire [37]	RSQ	10	10–40	2018	Public
Resilience & Coping					
Brief Resilience Scale [35]	BRS	6	1–5	2018–2023	Public
Brief COPE [10]	BriefCope	28 or 29	0–3	2019–2023	Public
Mindfulness & Sleep					
Mindful Attention Awareness Scale [7]	MAAS	7 or 15	1–6	2018–2023	Public
Pittsburgh Sleep Quality Index [9]	PSQI	9 or 10	0–21	2018–2023 ^a	Public
Wellbeing & Physical Health					
Flourishing Scale–Psychological Well-Being Scale [18]	FSPWB	8	8–56	2019–2023	Public
Gratitude Questionnaire [26]	GQ	6	6–42	2019–2023	Public
Short-Form Health Survey [40]	SF12	12	—	2018–2023	Public
Cohen-Hoberman Physical Symptoms [12]	CHIPS	33	0–132	2018–2023	Public
Food Security					
Food Security Survey Module [19]	foodSecurity	8	—	2019–2023 (post only)	DUA
Academic Functioning					
Academic Self-Efficacy & Efficacy for Self-Regulated Learning [11, 51]	AE	19	1–7	2019–2023	Public
Adversity & Discrimination					
Everyday Discrimination Scale [44]	EDS	28	0–45	2019–2023	Public ^c
Chronic Education Discrimination & Harassment [43]	CEDH	39	0–48	2019–2023	Public ^c
Major Experiences of Discrimination	MED	4	0–12	2019–2023	DUA
Adverse Childhood Experiences [20]	ACE	17	0–17	2019 (pre only)	DUA
PTSD Checklist for DSM-5 [42]	PTSD	20	0–80	2019–2023 ^b	Public ^c
Substance Use					
Brief Young Adult Alcohol Consequences Questionnaire [23]	B-YAACQ	24	0–24	2018–2023	DUA
Current Substance Use	CSU	varies	—	2018–2023	DUA
Life Context					
Major Life Events	MLE	varies	—	2018–2023	DUA
Social Media Use Questionnaire	SMuse	varies	—	2018–2023	Public
COVID-19 Questionnaire	COVID19	varies	—	2020–2023	DUA
Personality					
Big Five Inventory-10 [30]	BFI-10	10	1–5	2018–2023	Public

^a Total PSQI scores available for post-quarter surveys only in 2019–2023 due to missing items in other administrations.

^b PTSD administered at pre in 2019, and pre and post in 2020–2023.

^c The composite score is open-access; individual items are documented in the codebook but are not released in any tier.

2.7 Ecological Momentary Assessment

We administered EMA surveys in two formats each quarter. Twice-weekly surveys were delivered on Wednesdays and Sundays at 10am during weeks 1, 2, 4, 5, 6, 7, 9, and 10 of the quarter. Intensive daily surveys were delivered four times per day—at 10am (morning), 1pm (midday1), 5pm (midday2), and 9pm (evening)—during weeks 3 and 8 of the quarter. Participants were notified of each survey via push notification, and study coordinators contacted participants who missed EMA surveys to ensure continued engagement.

Through the twice-weekly surveys, we administered depression and anxiety screening (PHQ-4 [24]), perceived stress (PSS-4 [15]), everyday unfair treatment (occurrence, type, perpetrator, context, and coping response), and substance use (alcohol, cannabis, and other drugs). The intensive daily surveys administered momentary affect (PANAS-SF [41]), stress experience (severity, type, and primary coping strategy), social interactions and felt connectedness, emotion regulation, sleep quality and duration, emotional resources, physical activity and fatigue, daily activity context, and school engagement and academic workload.

We introduced additional modules in later cohorts to reflect emerging contextual factors. Beginning in 2020, community and online interaction items were added, alongside measures of COVID-19 academic impact, assessing perceived effects on degree progress, GPA, financial aid, and graduation. In 2021, accessibility barrier items were introduced to capture disability- and chronic illness-related barriers in academic, health, and social contexts, along with COVID-19 exposure and behavioral response items. As with the pre- and post-quarter surveys, certain EMA items in the 2018 pilot year differ from subsequent years and should be interpreted with caution; notably, negative affect items from 2018 are not directly comparable to those from 2019–2023 due to scale differences. Table 3 lists the EMA measures by survey type, year availability, and access tier. As with the pre/post measures, the codebook contains the full item text, response options, and year-by-year administration details for each EMA module.

Table 3. Summary of Ecological Momentary Assessment (EMA) Measures. DUA = Data Use Agreement. TW = Twice-weekly surveys (Wednesdays and Sundays); D = Daily surveys (4x/day during burst weeks). More details can be found in the complementary codebook.

Measure	Survey Type	Years Available	Tier
Mood & Affect			
Current Affect (PANAS-SF) ^a	TW & D	2018–2023	Public
Anxiety & Stress			
Depression & Anxiety Screening (PHQ-4)	TW & D	2019–2023	Public
Perceived Stress (PSS-4)	TW & D	2019–2023	Public
Stress & Coping	TW & D	2018–2023	Public
Social Support & Belonging			
Social Interactions	TW & D	2018–2023	Public
Community & Online Interaction	TW & D	2020	Public
Belongingness	D	2018	Public
Emotion Regulation & Resources			
Momentary Emotion Regulation ^c	D	2018–2023	Public
Emotional Resources ^c	D	2018–2023	Public
Mindfulness			
State Mindfulness (MAAS-S)	D	2019–2023	Public
Time Management	TW & D	2021	Public
Sleep			
Daily Sleep Diary	D	2019–2023	Public
Adversity, Trauma & Discrimination			
Everyday Unfair Treatment ^d	TW & D	2018–2023	DUA

Continued on next page

Draft manuscript

Table 3 continued

Measure	Survey Type	Years Available	Tier
Accessibility Barriers	TW & D	2021–2023	DUA
Substance Use			
Alcohol Use	TW & D	2018–2023	DUA
Drug Use	TW & D	2018–2023	DUA
Marijuana Use	TW & D	2020–2023	DUA
Life Context / Contextual Factors			
Activity	TW & D	2018, 2020–2023	Public
COVID-19 Exposure	TW & D	2020–2021	Public
Physical Health			
Physical Activity & Fatigue	D	2018	Public
Academic			
School Engagement	TW & D	2020	Public
Student Concerns & Workload	D	2018	Public
Overall Day Quality	D	2018	Public

^a Negative affect items in 2018 are not directly comparable to 2019–2023. Positive affect is not available in 2018.

^b Emotion regulation scale in 2018 differs slightly (missing response option 2).

^c resources_use_level in 2018 is not comparable to 2019–2023.

^d Discrimination reason items (by identity) available under DUA.

2.8 Academic Performance Records

We obtained academic performance data directly from the UW Office of the University Registrar for all participants who consented to registrar-data linkage. GPA coverage therefore varied across cohort years (73.4%–95.2%; Figure 3C). Records include fall quarter GPA (AutGPA), winter quarter GPA (WinGPA), spring quarter GPA (SprGPA), and cumulative GPA (CUMGPA) for each study year from 2018 to 2023. GPA values are reported on a 4.0 scale. GPA data are held in the by-request (managed) tier, available to qualified researchers under a DUA and current CITI certification; it is not part of the open or restricted deposit. Researchers should note that fall and winter quarter GPA records reflect performance in quarters prior to the spring data collection window and thus serve as retrospective academic performance indicators within each annual cycle.

2.9 Data Processing and Quality Control

We exported raw survey data from Qualtrics in comma-separated values (CSV) format and underwent a structured cleaning and quality assurance pipeline prior to deposit. The data cleaning scripts can be found in this [repo](#).

Response Exclusions. We excluded incomplete submissions—defined as responses where the participant did not reach the final page—along with empty responses and duplicate submissions. We further manually investigated duplicates and retained the most complete or most recent valid responses.

Out-of-Range Values. We flagged and removed responses falling outside the valid range for each item.

Score Recoding and Harmonization. We recoded reverse-scored items prior to scale computation. Composite scores were derived according to each measure’s published or study-specific scoring guide and standardized across cohort years to ensure cross-year comparability. Where measure versions differed across years—including the CESD (9-, 10-, and 20-item versions), PSS (10- and 14-item), UCLA Loneliness Scale (10- and 20-item), MAAS (7- and 15-item), and 2WaySSS (21- and 22-item)—we calculated composite scores using consistent item sets across versions. Version differences and scoring procedures are documented in detail in the accompanying codebook.

Participant De-Identification. We replaced original participant IDs with randomized 16-character alphanumeric identifiers prior to deposit. Identifiers are randomized independently per cohort year, so deposited files cannot be linked across years. A secured cross-year mapping file linking identifiers across study years is maintained by our research team in the by-request (managed) tier, available to qualified researchers under a DUA and current CITI certification. Disclosure-risk analysis for the deposit is described in Section 2.3.

3 Data Records

All data are available at [this link](#). Upon acceptance, all files will be transferred to a single ICPSR dataset and this link replaced with the accession number. Open-access files are available for download; restricted files require completion of a DUA, requested from the research team during review and through the ICPSR access request system after transfer; managed-tier files (demographics, GPA, and the cross-year participant linkage) are available by request to the research team under a DUA and current CITI certification.

3.1 File Organization

We organized data files by access tier, data type, and cohort year as a directory tree. At the top level, the two openly deposited tiers are separated into `public/` (open access) and `dua/` (restricted, released under a DUA); a separate `managed_access/` tree holds the by-request tier, including demographics, GPA, and the cross-year participant linkage, which is not part of the open deposit. Within each tier, files are grouped by data type into `composite/`, containing the composite and subscale scores (`prepost.csv` for the pre/post/demographic scales and `ema.csv` for the EMA summary scores), and `item-level/`, containing the raw item responses of measures that have no composite score (e.g., SF-12, social-media use, food security, COVID-19) together with the raw EMA items (`ema.csv` plus one file per EMA administration); each is further divided into one subfolder per cohort year (2018–2023). Every validated scale that has a composite is released as its composite/subscale score only—its item-level responses are documented in the codebook but are not deposited in any tier. A survey’s variables are partitioned across tiers and data types by their access tier and by whether they are a score or a raw item, so the same survey can appear in more than one file (for example, the demographic survey’s BFI-10 personality scores are in `public/composite/`, while its identifying demographic items and GPA are in the managed tier). Files are named by data type, e.g. `release/public/composite/2018/prepost.csv` or `release/public/item-level/2018/morning.csv`. Per-administration EMA files are labeled by survey time (morning, midday1, midday2, evening, and sun/wed for the twice-weekly surveys); 2018 twice-weekly responses are stored in a single weekly file, as day-of-week was not recorded separately that year.

3.2 Accompanying Documentation

We provide the following documentation alongside the dataset. The codebook and academic calendar are publicly available at [this link](#); the cross-year ID mapping is held in the managed tier:

Codebook. A complete codebook (`UWEXP_codebook.xlsx`) documenting all variable names, item text, response options, scoring procedures, access tier, and year-by-year availability for each measure across all data types.

Academic calendar. A file documenting the academic calendar for each cohort year, including quarter start and end dates, midterm weeks, finals weeks, and break periods, is provided to support temporal alignment of EMA data with the academic cycle. This file is publicly available.

Cross-year ID mapping. A secured cross-year participant identifier mapping file (UWEXP_pid_mapping.xlsx) linking randomized participant identifiers across cohort years is maintained by the research team and available to qualified researchers by request under the managed tier (DUA + current CITI certification). This file also identifies the SpecialPost 2020 cohort by their PIDs.

4 Technical Validation

This section provides evidence for the completeness, reliability, and cross-year comparability of the deposited dataset. We conducted all validation analyses on the cleaned data prior to deposit.

4.1 Survey Completion Rates

Completion rates for pre- and post-quarter surveys, EMAs, and GPA records are summarized in Figure 3. All rates are calculated as a proportion of participants enrolled via the demographic survey. For the SpecialPost 2020 cohort, completion rates are calculated as a proportion of the 49 enrolled participants.

Demo, Pre, and Post Surveys. Demographic survey completion was high across all cohort years, averaging 100%. Pre survey completion rates averaged 95.9% (range: 92.0%–100.0%), and post survey completion averaged 91.7% (range: 84.2%–100.0%). Note that the SpecialPost 2020 cohort (N = 49) completed a post-equivalent survey only, with no pre-quarter survey administered.

EMAs. Overall EMA completion averaged 85.8% (range: 79.1%–91.4%) across all cohort years and survey types. Daily EMAs averaged 84.2% (range: 76.6%–90.7%), and twice-weekly EMAs averaged 89.3% (range: 85.0%–92.8%).

Academic Performance Records. GPA records were obtained for an average of 81.9% of enrolled participants per year (range: 73.4%–95.2%).

4.2 Missing Data

Missing data in the UWEXP dataset arise from two distinct sources. The first source is *planned measure omissions*, where a measure was not administered in a given cohort year due to protocol changes or the pilot nature of the 2018 data collection, as documented in Table 2. These should not be treated as missing at random. The second source is *participant non-response*, where an administered measure was not completed by a given participant.

Pre and Post Surveys. We assessed missingness at the composite score level and summarized it in Figure 4. Core psychological measures—comprising 14–20 measures administered to all participants in a given cohort year—showed near-complete composite availability among participants who submitted the survey: pre-survey composite completion averaged 99.8% across cohort years and measures (minimum across any scale–cohort combination: 94.9%), and post-survey composite completion averaged 99.6% (minimum: 98.3%). In other words, participants who submitted the survey almost always provided sufficient item responses to yield a valid composite score. Post-survey composite-level missingness in the broader dataset primarily reflects non-retention—participants who did not return for the post-survey are missing all post-survey composites—resulting in composite missingness of 0–10% across measures and cohort years (mean: 6.2%; maximum: 10.4%).

We administered four measures only to eligible subgroups rather than all participants: the Everyday Discrimination Scale (EDS), the Chronic Education Discrimination and Harassment scale (CEDH), the Brief Young Adult Alcohol Consequences Questionnaire (B-YAACQ), and the Major Experiences of Discrimination (MED). These measures show

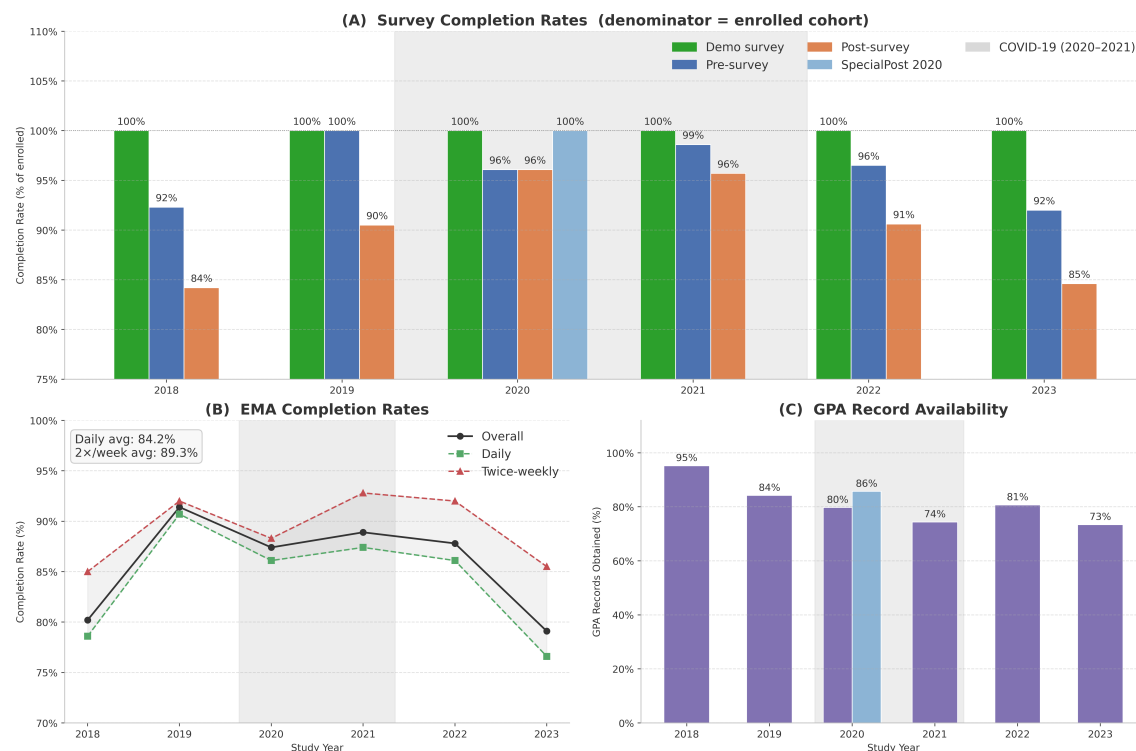


Fig. 3. **Survey Completion Rate by Data Type and Cohort Year.** All rates are expressed as percentage of participants enrolled via the demographic survey. (A) Completion rates for demographic (green), pre (dark blue), post (orange), and SpecialPost post-equivalent (light blue) surveys across cohort years 2018–2023. (B) EMA completion rates by cohort year, shown separately for daily surveys (square markers), twice-weekly surveys (triangle markers), and overall average (circle markers). (C) GPA record availability by cohort (dark purple) and SpecialPost 2020 cohort (light blue).

substantially lower apparent completion by design—averaging 33–51% at pre-survey and 24–40% at post-survey—reflecting eligibility-based routing rather than non-response among the general sample. For example, B-YAACQ was administered only to participants reporting alcohol use, and MED only to those reporting discrimination experiences.

EMAs. For EMAs, we report item-level missingness, which is defined as the proportion of non-conditional items not answered within a submitted EMA response, and is visualized in Figure 5. Among submitted EMA responses, the proportion of non-conditional items *not* completed was near-zero in 2018 and 2019 across all survey types: morning (0.2–0.5%), midday1 and midday2 (0–0.1%), evening (0.2–2.4%), and weekly (0.2–5.9%).

From 2020 onward, missingness increased for surveys that expanded substantially in length. The morning EMA showed the sharpest increase, rising from 0.2% in 2019 to 4.4% in 2020, 49.0% in 2021, and 49.8% in 2022, before recovering to 9.5% in 2023—consistent with its item count growing from approximately 29 items in 2018–2019 to approximately 90 items in 2021–2022 and being shortened again in 2023. The midday surveys (midday1 and midday2) showed a more moderate increase, rising from near-zero in 2018–2019 to 10–12% in 2021–2022 and 8–11% in 2023. The evening and weekly surveys increased to 21–25% and 22–30% respectively from 2020 onward, reflecting similar survey length

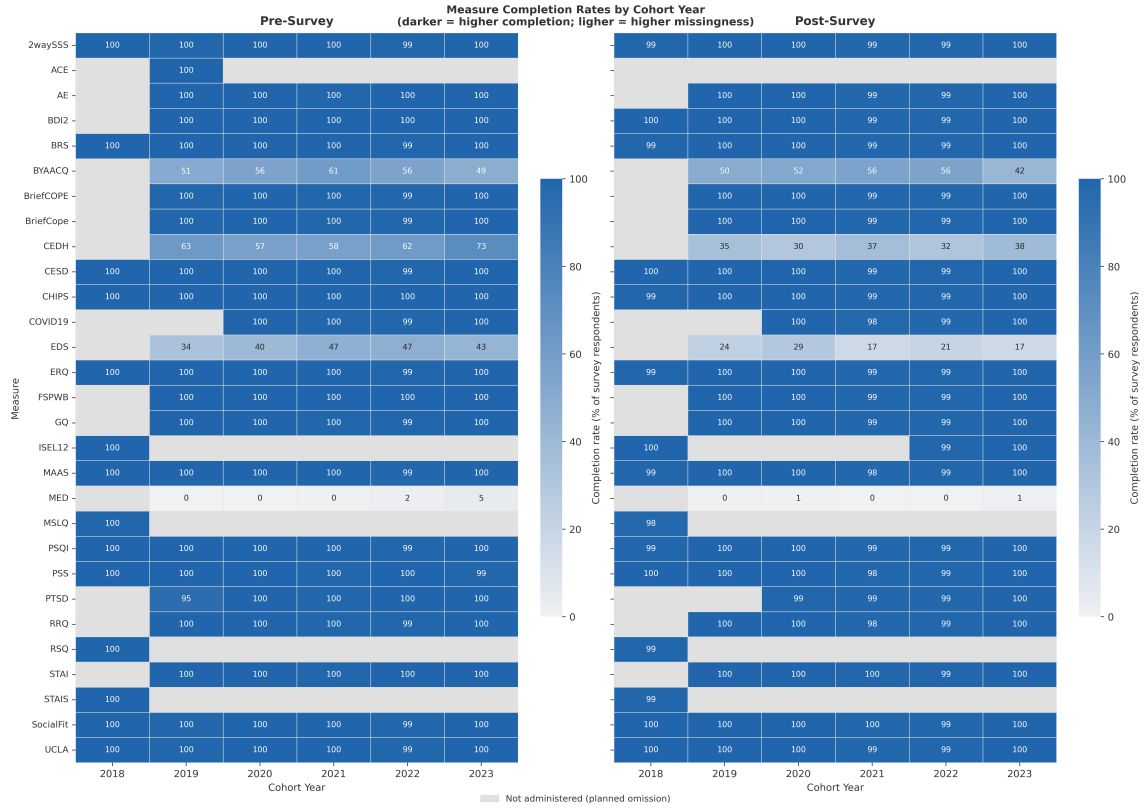


Fig. 4. **Measure Completion Rates by Cohort Year.** Heatmap showing the percentage of survey respondents with a valid composite score for each measure, by cohort year and survey type (pre-quarter, left; post-quarter, right). Darker cells indicate higher completion; lighter cells indicate higher missingness. Gray cells indicate planned omissions—measures not administered in that cohort year. Subgroup-administered measures (EDS, CEDH, B-YAACQ, MED) show lower completion by design, reflecting eligibility-based routing rather than non-response.

expansion. These patterns indicate that participants frequently did not complete the full survey after submitting a response, particularly as item counts increased in later cohorts.

Items with high apparent missingness in the EMA data are predominantly conditional skip-logic items — for example, stress-type subtypes shown only when a stressor is endorsed, or alcohol and drug follow-up items shown only to participants who reported use. These items do not reflect true non-response and are excluded from the missingness metric reported above. Researchers should distinguish between routing-based non-administration and genuine non-response when interpreting item-level missingness in the EMA files.

4.3 Internal Consistency

We computed Cronbach's α for every multi-item scale, by cohort year and pre/post administration, on the item responses with reverse-keyed items scored in a common direction; scales that changed length are reported per version and multi-subscale measures per subscale. Internal consistency was good to excellent throughout ($\alpha = 0.69$ – 0.96 across scales, cohorts, and administrations), with the lowest values for the four-item ERQ suppression subscale. Formative indices



Fig. 5. **EMA Item-Level Missingness Rate by Cohort Year.** Mean percentage of non-conditional items not completed within submitted EMA responses, shown separately for each cohort year (2018–2023). The solid line with square markers represents the daily EMA average (mean of morning, midday1, midday2, and evening surveys); the dashed line with circle markers represents weekly EMA completion. The shaded band reflects the min–max range across the four daily survey types on each survey day.

Internal consistency (Cronbach's α) by cohort year												
	Pre-survey						Post-survey					
	2018	2019	2020	2021	2022	2023	2018	2019	2020	2021	2022	2023
BDI-II		0.90	0.90	0.91	0.91	0.93	0.93	0.92	0.91	0.93	0.93	0.94
CES-D (9-item)	0.81	0.76	0.77	0.78	0.82	0.82	0.76	0.79	0.77	0.82	0.81	0.81
CES-D (10-item)	0.83		0.78	0.80	0.84	0.85	0.79		0.79	0.85	0.83	0.83
CES-D (20-item)	0.92						0.90					
STAI (State)	0.95						0.93					
STAI (Trait)		0.91	0.90	0.93	0.92	0.93		0.92	0.91	0.93	0.93	0.93
PSS (10-item)	0.84	0.86	0.83	0.88	0.86	0.84	0.82	0.88	0.86	0.88	0.88	0.86
PSS (14-item)		0.84	0.81	0.87	0.84	0.82		0.87	0.84	0.87	0.88	0.82
UCLA (10-item)	0.86	0.88	0.87	0.89	0.90	0.90	0.88	0.88	0.86	0.91	0.90	0.89
UCLA (20-item)	0.91						0.93					
2-Way SSS (Giving)	0.89	0.89	0.87	0.90	0.91	0.87	0.90	0.90	0.91	0.91	0.90	0.90
2-Way SSS (Receiving)	0.93	0.92	0.93	0.93	0.94	0.94	0.95	0.93	0.93	0.94	0.95	0.95
ISEL-12	0.83						0.85				0.87	0.86
Social Fit	0.90	0.92	0.87	0.89	0.91	0.89	0.88	0.90	0.91	0.90	0.90	0.90
ERQ (Reappraisal)	0.84	0.85	0.81	0.85	0.83	0.85	0.89	0.84	0.84	0.85	0.86	0.88
ERQ (Suppression)	0.72	0.69	0.76	0.76	0.80	0.79	0.81	0.73	0.71	0.80	0.79	0.80
RRQ (Rumination)		0.89	0.92	0.93	0.94	0.91		0.90	0.90	0.93	0.91	0.93
RRQ (Reflection)		0.89	0.92	0.93	0.94	0.92		0.90	0.92	0.93	0.92	0.91
RSQ	0.84						0.87					
BRS	0.85	0.86	0.88	0.88	0.89	0.89	0.86	0.84	0.86	0.90	0.89	0.89
Brief COPE (Adaptive)		0.85	0.84	0.87	0.87	0.89		0.87	0.85	0.88	0.87	0.87
Brief COPE (Maladaptive)		0.74	0.73	0.79	0.79	0.81		0.82	0.76	0.83	0.83	0.83
MAAS (7-item)	0.89	0.84	0.87	0.90	0.87	0.86	0.87	0.90	0.88	0.90	0.88	0.91
MAAS (15-item)		0.88	0.90	0.92	0.89	0.90		0.91	0.92	0.92	0.90	0.93
FS-PWB		0.90	0.89	0.92	0.93	0.92		0.90	0.92	0.93	0.94	0.94
GQ-6		0.71	0.76	0.80	0.79	0.81		0.78	0.76	0.80	0.75	0.79
CHIPS	0.93	0.92	0.89	0.92	0.91	0.94	0.92	0.93	0.90	0.95	0.94	0.94
Academic Self-Efficacy		0.88	0.89	0.88	0.90	0.88		0.87	0.89	0.90	0.91	0.90
Self-Regulated Learning		0.87	0.91	0.91	0.92	0.89		0.89	0.90	0.92	0.92	0.88
PTSD		0.94	0.93	0.94	0.95	0.95			0.94	0.96	0.95	0.96
EDS		0.86	0.87	0.88	0.87	0.87		0.86	0.89	0.87	0.86	0.86
CEDH		0.85	0.77	0.83	0.80	0.81		0.79	0.75	0.77	0.82	0.84
ACE		0.79										
B-YAACQ		0.92	0.82	0.82	0.77	0.89		0.90	0.87	0.84	0.85	0.83

Fig. 6. **Internal consistency (Cronbach's α) by cohort year and administration.** α for each multi-item scale (rows) at each cohort-year $\times$ pre/post administration (columns, 2018–2023), computed with reverse-keyed items scored in a common direction. Scale versions (e.g., PSS 10- vs. 14-item) and subscales (e.g., ERQ reappraisal/suppression) are shown as separate rows; darker cells indicate higher α , blank cells indicate the scale/version was not administered that year. Reliability was good to excellent throughout ($\alpha = 0.69$ – 0.96).

(e.g., PSQI, MLE), for which internal consistency is not the appropriate metric, are excluded. Full per-scale, per-cohort values are provided in the codebook (“Scale Reliability” sheet); Figure 6 summarizes them.

5 Usage Notes

This section provides practical guidance for researchers accessing and working with the UWEXP dataset, including notes on data access, longitudinal linkage, and key analytical considerations specific to this dataset’s design.

5.1 Getting Started

Researchers should begin with the codebook (UWEXP_codebook.xlsx), which is the single complete reference for the study and dataset: every variable, the exact item asked of participants, composite-score definitions, availability by cohort year, and access tier. Researchers wishing to review or reproduce any measure, or to check how an adapted measure differs across years, will find the necessary detail there. Open-access files are available for immediate download from [this link](#). Researchers requiring access to restricted files should complete a DUA (through the ICPSR access request system after transfer). Researchers requiring access to demographics, GPAs, and cross-cohort pid mapping file should provide institutional affiliation, a brief description of the intended research use, and a current CITI human-subjects certification to the research team. Their requests should also specify the minimal set of variables needed for the proposed analysis; releases are held to a $k \geq 5$ disclosure-risk floor on the specific combination of demographic variables requested.

5.2 Longitudinal Data Linkage

A participant identifier (PID) is included in all data files. PIDs are randomized per cohort year, so a returning participant appears under a different PID in each cohort; within a cohort the PID is consistent across data types. Cross-year linking of returning participants is supported through the cross-year participant identifier mapping file (UWEXP_pid_mapping.xlsx), available by request under the managed tier. Researchers should note that not all participants contributed data in every study year; participant-level missingness across years is documented in the cross-year participant identifier mapping file.

5.3 Handling the 2018 Pilot Year

The 2018 data collection served as a pilot year and has two distinct characteristics that researchers should consider. First, 2018 is the only cohort year spanning two academic quarters (winter and spring), making it uniquely valuable for within-year longitudinal analyses of psychological and behavioral change across two consecutive quarters. Researchers interested in examining intra-year trajectories, seasonal effects, or the transition from winter to spring quarter are encouraged to leverage this feature of the 2018 data.

Second, certain measure versions, item sets, and EMA file structures differ from the 2019–2023 protocol, and researchers should exercise caution when including 2018 data in cross-year longitudinal analyses. Specific caveats include: PANAS-SF negative affect items in 2018 are not directly comparable to 2019–2023 due to scale and item differences; twice-weekly EMA responses in 2018 are combined into a single weekly file rather than separate Sunday and Wednesday files; and several measures were administered in different versions or not at all in 2018 relative to subsequent years. Year-by-year item availability and version differences are documented in detail in the accompanying codebook.

5.4 Cross-Year Score Comparability

Several measures varied in version or length across cohort years, including the CESD (9-, 10-, and 20-item versions), PSS (10- and 14-item), UCLA Loneliness Scale (10- and 20-item), MAAS (7- and 15-item), and 2WaySSS (21- and 22-item). We calculated composite scores for these measures using consistent item subsets to ensure cross-year comparability, as described in Section 2.6. Researchers should nonetheless be aware of these version differences when conducting analyses that span multiple cohort years and should report them as a methodological consideration. The specific item subsets used for each version and the scoring procedures applied are documented in detail in the accompanying codebook.

5.5 The SpecialPost 2020 Cohort

The SpecialPost 2020 cohort (N = 49), recruited in June 2020 specifically to supplement COVID-19-related assessments and differs from the primary longitudinal cohort in two important ways. First, these participants entered the study mid-cycle during the spring 2020 post-quarter survey window and therefore do not have pre-quarter survey data for 2020. Second, their initial data collection focused specifically on COVID-19 measures; from 2021 onward, their data collection was structurally identical to that of returning participants from the primary cohort. Researchers conducting longitudinal analyses should carefully consider whether to include the SpecialPost 2020 cohort depending on the research question. For analyses requiring pre-quarter 2020 data or focused on pre-pandemic to pandemic trajectories, this cohort should be excluded or treated as a separate group. The SpecialPost 2020 cohort can be identified via the cross-year mapping file (managed tier).

6 Data Availability

The UWEXP dataset is available at [this link](#). Open-access files are available for download without registration; upon acceptance, all files will be transferred to ICPSR and this link replaced with the accession number. Restricted files containing sensitive variables are available to qualified researchers upon completion of a DUA, submitted through the ICPSR access request system. Demographics, GPA, and the cross-year participant identifier mapping are held in a by-request (managed) tier, available to qualified researchers under a DUA and current CITI certification. The complete codebook (UWEXP_codebook.xlsx) and academic calendar are included in the open deposit.

[The linked folder contains the deposit-ready open and restricted tiers, complete codebook, academic calendar, and de-identification documentation, mirroring the ICPSR file structure exactly. Managed-tier files (demographics, GPA, cross-year PID mapping) are excluded from the linked copy consistent with the access model described above; their structure is documented in the codebook]

7 Code Availability

All data processing and quality control scripts used to clean and harmonize the raw survey data prior to deposit are publicly available at [this repo](#). We provide scripts for response exclusion, out-of-range value flagging, score recoding and harmonization, composite-score generation, the access-tier split and de-identification (k-anonymity) reporting, and internal-consistency (Cronbach's α) computation.

References

- [1] Nancy E Adler, Elissa S Epel, Grace Castellazzo, and Jeannette R Ickovics. 2000. Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health psychology* 19, 6 (2000), 586.

- [2] Jeffrey Jensen Arnett. 2000. Emerging adulthood: A theory of development from the late teens through the twenties. *American psychologist* 55, 5 (2000), 469.
- [3] Jeffrey Jensen Arnett. 2005. The developmental context of substance use in emerging adulthood. *Journal of drug issues* 35, 2 (2005), 235–254.
- [4] Randy P Auerbach, Philippe Mortier, Ronny Bruffaerts, Jordi Alonso, Corina Benjet, Pim Cuijpers, Koen Demyttenaere, David D Ebert, Jennifer Greif Green, Penelope Hasking, et al. 2018. WHO world mental health surveys international college student project: Prevalence and distribution of mental disorders. *Journal of abnormal psychology* 127, 7 (2018), 623.
- [5] Aaron T Beck, Robert A Steer, Roberta Ball, and William F Ranieri. 1996. Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of personality assessment* 67, 3 (1996), 588–597.
- [6] Peter J Bieling, Martin M Antony, and Richard P Swinson. 1998. The State–Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour research and therapy* 36, 7-8 (1998), 777–788.
- [7] Kirk Warren Brown and Richard M Ryan. 2003. The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology* 84, 4 (2003), 822.
- [8] Ronny Bruffaerts, Philippe Mortier, Glenn Kiekens, Randy P Auerbach, Pim Cuijpers, Koen Demyttenaere, Jennifer G Green, Matthew K Nock, and Ronald C Kessler. 2018. Mental health problems in college freshmen: Prevalence and academic functioning. *Journal of affective disorders* 225 (2018), 97–103.
- [9] Daniel J Buysse, Charles F Reynolds III, Timothy H Monk, Susan R Berman, and David J Kupfer. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28, 2 (1989), 193–213.
- [10] Charles S Carver. 1997. You want to measure coping but your protocol’s too long: Consider the brief cope. *International journal of behavioral medicine* 4, 1 (1997), 92–100.
- [11] Martin M Chemers, Li-tze Hu, and Ben F Garcia. 2001. Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational psychology* 93, 1 (2001), 55.
- [12] Sheldon Cohen and Harry M Hoberman. 1983. Positive events and social supports as buffers of life change stress 1. *Journal of applied social psychology* 13, 2 (1983), 99–125.
- [13] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A global measure of perceived stress. *Journal of health and social behavior* (1983), 385–396.
- [14] Sheldon Cohen, Robin Mermelstein, Tom Kamarck, and Harry M Hoberman. 1985. Measuring the functional components of social support. In *Social support: Theory, research and applications*. Springer, 73–94.
- [15] Sheldon Cohen and Gail M Williamson. 1988. Perceived stress in a probability sample of the United States. In *The Social Psychology of Health: Claremont Symposium on Applied Social Psychology*, Shirley Spacapan and Stuart Oskamp (Eds.). Sage, Newbury Park, CA, 31–67.
- [16] Jason C Cole, Adele S Rabin, Tom L Smith, and Alan S Kaufman. 2004. Development and validation of a Rasch-derived CES-D short form. *Psychological assessment* 16, 4 (2004), 360.
- [17] Colleen S Conley, Alexandra C Kirsch, Daniel A Dickson, and Fred B Bryant. 2014. Negotiating the transition to college: Developmental trajectories and gender differences in psychological functioning, cognitive-affective strategies, and social well-being. *Emerging adulthood* 2, 3 (2014), 195–210.
- [18] Ed Diener, Derrick Wirtz, William Tov, Chu Kim-Prieto, Dong-won Choi, Shigehiro Oishi, and Robert Biswas-Diener. 2010. New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social indicators research* 97, 2 (2010), 143–156.
- [19] ESDA. 2012. U.S. Adult Food Security Survey Module: Three-stage Design, with Screeners, Economic Research Service. <https://www.ers.usda.gov/media/6424/us-adult-food-security-survey-module.pdf>.
- [20] Vincent J Felitti, Robert F Anda, Dale Nordenberg, David F Williamson, Alison M Spitz, Valerie Edwards, and James S Marks. 1998. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American journal of preventive medicine* 14, 4 (1998), 245–258.
- [21] James J Gross and Oliver P John. 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology* 85, 2 (2003), 348.
- [22] Robert I Kabacoff, Daniel L Segal, Michel Hersen, and Vincent B Van Hasselt. 1997. Psychometric properties and diagnostic utility of the Beck Anxiety Inventory and the State-Trait Anxiety Inventory with older adult psychiatric outpatients. *Journal of anxiety disorders* 11, 1 (1997), 33–47.
- [23] Christopher W Kahler, David R Strong, and Jennifer P Read. 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 (2005), 1180–1189.
- [24] Kurt Kroenke, Robert L Spitzer, Janet BW Williams, and Bernd Löwe. 2009. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics* 50, 6 (2009), 613–621.
- [25] Sarah Ketchen Lipson, Sasha Zhou, Sara Abelson, Justin Heinze, Matthew Jirsa, Jasmine Morigney, Akilah Patterson, Meghna Singh, and Daniel Eisenberg. 2022. Trends in college student mental health and help-seeking by race/ethnicity: Findings from the national healthy minds study, 2013–2021. *Journal of affective disorders* 306 (2022), 138–147.
- [26] Michael E McCullough, Robert A Emmons, and Jo-Ann Tsang. 2002. The grateful disposition: a conceptual and empirical topography. *Journal of personality and social psychology* 82, 1 (2002), 112.
- [27] Margaret E Morris, Kevin S Kuehn, Jennifer Brown, Paula S Nurius, Han Zhang, Yasaman S Sefidgar, Xuhai Xu, Eve A Riskin, Anind K Dey, Sunny Consolvo, et al. 2021. College from home during COVID-19: A mixed-methods study of heterogeneous experiences. *PLoS one* 16, 6 (2021), e0251580.

- [28] Subigya Nepal, Wenjun Liu, Arvind Pillai, Weichen Wang, Vlado Vojdanovski, Jeremy F Huckins, Courtney Rogers, Meghan L Meyer, and Andrew T Campbell. 2024. Capturing the college experience: a four-year mobile sensing study of mental health, resilience and behavior of college students during the pandemic. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 8, 1 (2024), 1–37.
- [29] Paula S Nurius, Yasaman S Sefidgar, Kevin S Kuehn, Jake Jung, Han Zhang, Olivia Figueira, Eve A Riskin, Anind K Dey, and Jennifer C Mankoff. 2023. Distress among undergraduates: Marginality, stressors and resilience resources. *Journal of American college health* 71, 5 (2023), 1445–1453.
- [30] Beatrice Rammstedt and Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of research in Personality* 41, 1 (2007), 203–212.
- [31] Daniel W Russell. 1996. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of personality assessment* 66, 1 (1996), 20–40.
- [32] John E Schulenberg and Jennifer L Maggs. 2002. A developmental perspective on alcohol use and heavy drinking during adolescence and the transition to young adulthood. *Journal of Studies on Alcohol, Supplement* 14 (2002), 54–70.
- [33] Yasaman S Sefidgar, Woosuk Seo, Kevin S Kuehn, Tim Althoff, Anne Browning, Eve Riskin, Paula S Nurius, Anind K Dey, and Jennifer Mankoff. 2019. Passively-sensed behavioral correlates of discrimination events in college students. *Proceedings of the ACM on Human-computer Interaction* 3, CSCW (2019), 1–29.
- [34] Jane Shakespeare-Finch and Patricia L Obst. 2011. The development of the 2-way social support scale: A measure of giving and receiving emotional and instrumental support. *Journal of personality assessment* 93, 5 (2011), 483–490.
- [35] Bruce W Smith, Jeanne Dalen, Kathryn Wiggins, Erin Tooley, Paulette Christopher, and Jennifer Bernard. 2008. The brief resilience scale: assessing the ability to bounce back. *International journal of behavioral medicine* 15, 3 (2008), 194–200.
- [36] Paul D Trapnell and Jennifer D Campbell. 1999. Private self-consciousness and the five-factor model of personality: distinguishing rumination from reflection. *Journal of personality and social psychology* 76, 2 (1999), 284.
- [37] Wendy Treynor, Richard Gonzalez, and Susan Nolen-Hoeksema. 2003. Rumination reconsidered: A psychometric analysis. *Cognitive therapy and research* 27, 3 (2003), 247–259.
- [38] Gregory M Walton and Geoffrey L Cohen. 2007. A question of belonging: race, social fit, and achievement. *Journal of personality and social psychology* 92, 1 (2007), 82.
- [39] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2024. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 3–14.
- [40] John E Ware, Mark Kosinski, and Susan D Keller. 1996. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 34, 3 (1996), 220–233.
- [41] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.
- [42] Frank W Weathers, Brett T Litz, Terence M Keane, Patrick A Palmieri, Brian P Marx, Paula P Schnurr, et al. 2013. The ptsd checklist for dsm-5 (pcl-5). (2013).
- [43] David R Williams, Hector M Gonzalez, Stacey Williams, Selina A Mohammed, Hashim Moomal, and Dan J Stein. 2008. Perceived discrimination, race and health in South Africa. *Social science & medicine* 67, 3 (2008), 441–452.
- [44] David R Williams, Yan Yu, James S Jackson, and Norman B Anderson. 1997. Racial differences in physical and mental health: Socio-economic status, stress and discrimination. *Journal of health psychology* 2, 3 (1997), 335–351.
- [45] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K Villalba, Janine M Dutcher, Michael J Tumminia, Tim Althoff, Sheldon Cohen, Kasey G Creswell, J David Creswell, et al. 2019. Leveraging routine behavior and contextually-filtered features for depression detection among college students. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–33.
- [46] Xuhai Xu, Prerna Chikersal, Janine M Dutcher, Yasaman S Sefidgar, Woosuk Seo, Michael J Tumminia, Daniella K Villalba, Sheldon Cohen, Kasey G Creswell, J David Creswell, et al. 2021. Leveraging collaborative-filtering for personalized behavior modeling: a case study of depression detection among college students. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–27.
- [47] Xuhai Xu, Xin Liu, Han Zhang, Weichen Wang, Subigya Nepal, Yasaman Sefidgar, Woosuk Seo, Kevin S Kuehn, Jeremy F Huckins, Margaret E Morris, et al. 2023. GLOBEM: cross-dataset generalization of longitudinal human behavior modeling. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–34.
- [48] Xuhai Xu, Han Zhang, Yasaman Sefidgar, Yiyi Ren, Xin Liu, Woosuk Seo, Jennifer Brown, Kevin Kuehn, Mike Merrill, Paula Nurius, et al. 2022. GLOBEM dataset: multi-year datasets for longitudinal human behavior modeling generalization. *Advances in neural information processing systems* 35 (2022), 24655–24692.
- [49] Han Zhang, Margaret Morris, Paula Nurius, Kelly Mack, Jennifer Brown, Kevin Kuehn, Yasaman Sefidgar, Xuhai Xu, Eve Riskin, Anind Dey, et al. 2022. Impact of online learning in the context of COVID-19 on undergraduates with disabilities and mental health concerns. *ACM Transactions on Accessible Computing* 15, 4 (2022), 1–27.
- [50] Han Zhang, Yiyi Ren, Paula S Nurius, Jennifer Mankoff, and Anind K Dey. 2025. Towards Human-Centered Early Prediction Models for Academic Performance in Real-World Contexts. *Proceedings of the ACM on Human-Computer Interaction* 9, 7 (2025), 1–41.
- [51] Barry J Zimmerman, Albert Bandura, and Manuel Martinez-Pons. 1992. Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American educational research journal* 29, 3 (1992), 663–676.

Author Contribution

The following describes each author's contribution to the presented dataset.

- Han Zhang: Led the development and refinement of data cleaning and processing infrastructure, technical validation of data quality, generation and maintenance of the codebook, paper writing and organization, and design of the data sharing process. Also assisted with study design and data collection from 2020 onward.
- Yasaman Sefidgar: Led the design of the data cleaning and processing infrastructure and codebook, as well as planning for the 2019 data collection. Also assisted with data collection in 2018 and 2019, and with planning for the 2020 data collection.
- Xuhai Xu: Led data collection planning and execution for the 2022 cohort and assisted with data collection across other cohort years from 2019 onward. Reviewed and commented on the manuscript.
- Yiyi Ren: Led data collection planning and execution for the 2023 cohort, assisted with the data collection for the 2022 cohort.
- Kevin Kuehn: Led data collection planning and execution for the 2019 cohort.
- Woosuk Seo: Led data collection planning and execution for the 2018 cohort.
- Paula Nurius: Supervised study material design and high-level planning throughout the full study period.
- Margaret E. Morris: Supervised study material design and high-level planning from 2019 onward. Reviewed and commented on the manuscript.
- Eve Riskin: Supervised study material design and high-level planning throughout the full study period. Reviewed and commented on the manuscript.
- Anind Dey: Supervised study material design and high-level planning throughout the full study period. Supervised data analysis, data use agreement, and paper writing.
- Jennifer Mankoff: Supervised study material design and high-level planning throughout the full study period. Supervised data analysis, data use agreement, and paper writing. Reviewed, edited, and commented on the manuscript.

Competing Interests

The authors declare no competing interests that could have influenced the work reported in this paper.