Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Social Media Identity Statements and Study Burnout Among University Students: A Secondary Analysis of the Helsinki COVID-19 Survey

Light Onyoche Victor

lightonyochevictor@gmail.com

Benue State University

Research Article

Keywords: social media, study burnout, university students, secondary data analysis, COVID-19, cross-sectional study

Posted Date: October 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11515943/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background: Associations between social media and well-being depend on what is measured. A public University of Helsinki survey contains six statements about social-media identity, similarity, and trust, alongside study-burnout items. This secondary analysis examined their cross-sectional association.

Methods: The Autumn 2020 public CSV contained 1,603 records. We calculated the mean of the six documented social-media statements and a nine-item study-burnout score following the source team's OSB9 scoring syntax. Complete cases for both scores, age group, and gender formed the analysis sample (n = 1,399). We fitted ordinary least squares regression with HC3 robust standard errors, adjusting for age group and gender.

Results: Mean social-media statement score was 3.29 (SD = 1.12; response range 1–7), and mean study burnout was 3.29 (SD = 1.34; item range 1–6). The social-media score had a small negative bivariate correlation with burnout (r = −0.026). In the adjusted model, a one-point higher statement score was associated with 0.074 points lower burnout (B = −0.074, HC3 95% CI −0.138 to −0.009); model R^2 was 0.070.

Conclusions: The observed association was small and cross-sectional. The six statements do not measure time spent, frequency of use, or problematic use, and their construct validity is not established in the public documentation. The result should therefore be interpreted as an exploratory association between a documented statement index and study burnout, not as evidence that social-media use changes student well-being.

1. Introduction

Social media use is not a single exposure. Time spent on platforms, frequency of checking, problematic use, social comparison, online identity, and trust in information are related but distinct constructs. Reviews have reported heterogeneous associations between broad social-media use and well-being, while problematic-use measures more consistently show adverse associations. This makes it important to describe the exact exposure items instead of using "social media use" as if it were a uniform dose [3,4].

Study burnout is a relevant student outcome because it captures study-related exhaustion, cynicism, and perceived inadequacy. During the COVID-19 pandemic, University of Helsinki survey research documented changes in study burnout and engagement across repeated cross-sectional waves [2]. The open survey archive includes additional items, among them six statements on social-media identity, similarity, and trust [1]. These statements provide an opportunity for a focused secondary analysis, although the archive's scale documentation does not cite a validation study for this item set.

The present analysis estimates the association between the mean of those six statements and a study-burnout composite in the Autumn 2020 public file. The analysis is explicitly exploratory and cross-

sectional. It does not test the effects of screen time, platform frequency, problematic social-media use, or a validated social-media exposure scale. The question is whether the documented statement score is associated with study burnout after adjustment for age group and gender.

2. Methods

2.1. Design and data source

This study is a secondary analysis of the openly available “University Student Well-being During Covid-19” dataset, deposited on Zenodo by the University of Helsinki research team [1]. The archive describes six voluntary, anonymous online surveys conducted in spring and autumn 2020, 2021, and 2022 among University of Helsinki students. The waves are repeated cross-sectional snapshots; respondents cannot be linked between waves. The analysis used only the public Autumn 2020 CSV file. The dataset archive provides the source files, a bilingual sample report, and item-level scale documentation. The dataset is licensed under Creative Commons Attribution 4.0 (CC BY 4.0).

2.2. Participants and analytic sample

The Autumn 2020 raw public CSV had 1,603 rows. We retained records with non-missing responses to all six social-media statements, all nine burnout items used for scoring, and both age group and gender. First, 163 rows lacked at least one social-media item, leaving 1,440. Of those, 11 lacked at least one burnout item, leaving 1,429. A further 30 lacked age group or gender, leaving 1,399 complete cases. These are counts for this secondary analysis, not PRISMA study-selection counts. The raw public file did not contain all screening variables used in the original research-team preparation syntax; accordingly, this analysis uses a clearly defined complete-case set from the deposited raw file and does not claim to reproduce the original article’s analytic sample.

Table 1
Record flow for the secondary analysis. The figures arise from the deposited participant-level file and are not literature-search or PRISMA review counts.

Stage	Records
Autumn 2020 public raw file	1,603
Excluded: ≥1 missing social-media statement	163
Remaining with all six social-media items	1,440
Excluded: ≥1 missing burnout item	11
Remaining with complete social-media and burnout scores	1,429
Excluded: age group or gender missing	30
Complete-case analytic sample	1,399

2.3. Measures

Social-media statements. The codebook lists six items under a shared prompt asking what respondents think of statements about social media. The statements concern whether online communities are important to identity, whether respondents are proud of those communities, whether they interact only with similar people or people sharing their interests, whether they trust information shared with them, and whether people on social media think similarly to them. Responses range from 1 (“Does not describe me at all”) to 7 (“Describes me completely”). We calculated the arithmetic mean when all six items were present; higher scores indicate greater endorsement. The item set has no source or validation reference in the archive’s scale-documentation row. Its internal consistency in the item-complete raw sample was Cronbach’s α = 0.807 (n = 1,440), which does not establish construct validity.

Study burnout. We calculated the mean of burnout items 1–6 and 8–10, producing the source team’s nine-item OSB9 composite on the recorded 1–6 item scale. Item 7 was omitted to match the original preparation syntax available from the research team’s OSF materials [5]. Higher values indicate greater study burnout. Internal consistency for these nine items in records with complete item responses was α = 0.920 (n = 1,589).

Covariates. Age was documented in four groups: 20 years or younger, 21–25, 26–30, and 31 years or older. Gender was coded in the codebook as female, male, or other. In regression, the youngest age group and female category served as reference levels.

2.4. Statistical analysis

We report means and standard deviations, the Pearson correlation between the two composite scores, and Cronbach’s alpha as descriptive internal-consistency estimates. The primary model regressed study burnout on the social-media statement mean, age-group indicators, and gender indicators. We report unstandardized coefficients and HC3 heteroskedasticity-robust standard errors with 95% normal-approximation confidence intervals. The coefficient for the social-media statement score is the conditional mean difference in burnout associated with a one-point higher score, holding age group and gender constant. Analysis was complete-case. No imputation, causal model, or multiplicity-adjusted family of hypotheses was used. R^2 is reported as a model-fit description, not as a measure of causal explanation.

2.5. Ethics and transparency

The original study publication reports that the ethical standards of the Declaration of Helsinki were followed and that informed consent was obtained from all participants [2]. The Zenodo archive describes the released responses as anonymous. This manuscript reports secondary analysis of the public data and did not recruit or contact participants. The author must confirm with the institution and data custodians whether any additional local determination or permission is required before posting. No ethics committee approval number for this secondary analysis is claimed.

3. Results

3.1. Sample and descriptive statistics

The analysis included 1,399 complete cases. The score means were close to the midpoint of their respective response ranges: social-media statement mean 3.29 (SD = 1.12) on the 1–7 response scale, and study burnout mean 3.29 (SD = 1.34) on the 1–6 item scale. The bivariate correlation between the scores was − 0.026, indicating little unadjusted linear association.

Table 2
Descriptive statistics and analytic-sample composition.

Measure	n	Mean	SD
Social-media statement mean (1–7)	1,399	3.29	1.12
Study burnout OSB9 mean (1–6)	1,399	3.29	1.34
Age group: ≤20 / 21–25 / 26–30 / ≥31	1,399	203 / 641 / 285 / 270	—
Gender: female / male / other	1,399	1,127 / 234 / 38	—

3.2. Adjusted association

In the model adjusted for age group and gender (n = 1,399; R^2 = 0.070; adjusted R^2 = 0.066), a one-point increase in the social-media statement score was associated with 0.074 points lower study burnout (B = − 0.074, HC3 SE = 0.033, 95% CI − 0.138 to − 0.009). The estimate is small relative to the six-point burnout item range. The association is observational and should not be interpreted as a protective causal effect.

Table 3
Complete-case ordinary least squares coefficients with HC3 robust standard errors. The model also includes the listed categorical covariates; R^2 = 0.070 and adjusted R^2 = 0.066. Confidence intervals use the normal approximation.

Predictor	B	HC3 SE	95% CI
Social-media statement mean	−0.074	0.033	−0.138 to − 0.009
Age 21–25 (ref ≤ 20)	−0.168	0.101	−0.365 to 0.029
Age 26–30 (ref ≤ 20)	−0.599	0.120	−0.833 to − 0.365
Age ≥ 31 (ref ≤ 20)	−0.958	0.120	−1.193 to − 0.724
Male (ref female)	−0.183	0.095	−0.369 to 0.003
Other (ref female)	0.260	0.196	−0.124 to 0.645

4. Discussion

In this secondary analysis of the public Autumn 2020 Helsinki student survey, the six social-media identity, similarity, and trust statements had a small negative adjusted association with study burnout. The bivariate correlation was close to zero, and the estimated conditional association was modest. The result does not show that social media reduces burnout: the exposure was not time or frequency of use, the design was cross-sectional, and unmeasured confounding and reverse direction remain plausible.

This distinction matters when comparing the finding with reviews of social media and well-being. Existing reviews combine measures ranging from general use to problematic use and show that associations vary by exposure and outcome definition [3,4]. The present six-item score reflects endorsement of particular statements about online identity and social similarity, plus trust in shared information. It cannot be treated as a proxy for hours spent online, addiction, or clinically meaningful wellbeing. A validation study should test whether the items form a coherent scale and whether that score predicts outcomes independently of use intensity and problematic-use symptoms.

4.1. Strengths and limitations

The analysis uses a real, openly documented data release, reports its complete-case record flow, follows the source team's documented OSB9 scoring, and provides reproducible item definitions and code. The archive's scale documentation includes English item wordings and response anchors, and the original research publication documents consent and ethical standards [1,2]. These features make the data provenance more traceable than a synthetic or undocumented dataset.

Several limitations constrain interpretation. First, the social-media item set has no validation citation in the codebook and mixes identity, homophily, pride, and trust statements; internal consistency alone cannot prove that it measures one construct. Second, no item asks about frequency, duration, platform, or problematic use. Third, the survey is cross-sectional and was collected during the exceptional pandemic period, so temporal ordering and generalization to current student life are uncertain. Fourth, the archive describes voluntary convenience sampling from one university, and the sample is disproportionately female. Fifth, the public raw data file includes 1,603 rows, while the published study reports different wave-specific participant counts and uses preparation filters; this analysis should not be described as a reproduction of that publication. Finally, the model's modest R^2 indicates that most observed variation in burnout remains outside the model.

4.2. Implications for future research

Future studies should use validated measures that distinguish social-media duration, active versus passive use, social comparison, online support, and problematic use. The six statements could be evaluated through exploratory and confirmatory factor analysis, cognitive interviewing, and convergent and discriminant validity tests before being treated as a scale. Longitudinal data with repeated exposure and burnout measurements would be needed to assess temporal ordering. Replication outside a single university and outside pandemic conditions would help establish generalizability.

5. Conclusions

A small inverse association was observed between study burnout and a six-item index of social-media identity, similarity, and trust statements in complete cases from the Helsinki Autumn 2020 public survey. Because the item set does not measure time or frequency of use and lacks documented validation, the finding should remain exploratory. It neither establishes that social media improves student well-being nor supports causal or population-wide conclusions.

Declarations

Ethics approval:

The original study publication reports adherence to the Declaration of Helsinki and participant informed consent. No separate approval number for this secondary analysis is claimed; the author must verify any applicable institutional determination before submission.

Consent to participate:

Obtained in the original study, as reported by the original investigators.

Consent for publication:

Not applicable to this secondary analysis of anonymous public data.

Funding:

No funding was received.

Competing interests: There were no competing interests.

Author contributions:

Light Onyoche Victor, AI assisted with drafting and analysis-code preparation for this manuscript

Acknowledgments:

I acknowledge the University of Helsinki data creators and my friend Queen Ochanya Esa, who helped in reviewing the article

Data availability:

The dataset is openly available from Zenodo, “University Student Well-being During Covid-19,” DOI: 10.5281/zenodo.19492338. Source preparation materials are linked from the original paper’s OSF record at https://osf.io/6d35r/.

Code availability: The analysis script and machine-readable result file accompany this draft. The script uses the public Autumn 2020 CSV and Python with pandas and NumPy.

References

1. Niikkonen I, Hietajärvi L, Upadyaya K, Salmela-Aro K (2026) University Student Well-being During Covid-19 [dataset]. Zenodo; https://doi.org/10.5281/zenodo.19492338
2. Salmela-Aro K, Upadyaya K, Ronkainen I, Hietajärvi L (2022) Study burnout and engagement during COVID-19 among university students: the role of demands, resources, and psychological needs. J Happiness Stud 23:2685–2702. https://doi.org/10.1007/s10902-022-00518-1
3. Ansari S, Iqbal N, Asif R, Hashim M, Farooqi SR, Alimoradi Z (2024) Social media use and well-being: a systematic review and meta-analysis. Cyberpsychology Behav Social Netw 27(10):704–719. https://doi.org/10.1089/cyber.2024.0001
4. Ahmed O, Walsh EI, Dawel A, Alateeq K, Espinoza Oyarce DA, Cherbuin N (2024) Social media use, mental health and sleep: a systematic review with meta-analyses. J Affect Disord 367:701–712. https://doi.org/10.1016/j.jad.2024.08.193
5. Salmela-Aro K, Kiuru N, Leskinen E, Nurmi JE (2009) School Burnout Inventory (SBI): reliability and validity. Eur J Psychol Assess 25(1):48–57. https://doi.org/10.1027/1015-5759.25.1.48
6. University of Helsinki research team (2026) Source data preparation and analysis materials for Study Burnout and Engagement During COVID-19 Among University Students. Open Science Framework. https://osf.io/6d35r/. Accessed 26 October