Beyond Surveys: Adolescent Profiling via Ecological Momentary Assessment and Mobile Sensing

Jasminka Dobša
University of Zagreb
Faculty of Organization and Informatics
Varaždin, Croatia
jasminka.dobsa@foi.hr

Simona Korenjak-Černe
University of Ljubljana
School of Economics and Business, and IMFM
Ljubljana, Slovenia
simona.cerne@ef.uni-lj.si

Miranda Novak
University of Zagreb
Faculty of Education and Rehabilitation Sciences
Zagreb, Croatia
miranda.novak@erf.unizg.hr

Maja Buhin Pandur
University of Zagreb
Faculty of Organization and Informatics
Varaždin, Croatia
mbuhin@foi.unizg.hr

Lucija Šutić
University of Zagreb
Faculty of Education and Rehabilitation Sciences
Zagreb, Croatia
lucija.sutic@erf.unizg.hr

Abstract

The aim of this study is to identify profiles of adolescents using survey data and data collected via mobile phones, which included ecological momentary assessment (EMA) and passive mobile sensing. EMA involved responses to short questionnaires delivered seven times per day over one week, while mobile sensing captured time spent using different categories of mobile applications. The study was conducted on a sample of 77 secondary school students. Profiling was performed through clustering of EMA data aggregated into six composite variables reflecting confidence, attentiveness, positive and negative emotions related to friends, and overall positive and negative affect. Based on the interpretability of the results, four adolescent profiles were identified. These profiles are further explained using survey data and passive data on mobile application usage patterns.

Keywords

Adolescents, clustering, mobile sensing, ecological momentary assessment, well-being

1 Introduction

This study was conducted using the Effortless Assessment of Risk States (EARS) application developed by Ksana Health in collaboration with the University of Oregon (https://ksanahealth.com/ears/) [6]. The EARS application was originally launched in 2018 to facilitate the collection of high-quality passive mobile sensing data and to support the development of predictive machine learning algorithms capable of identifying risk states for human well-being before they escalate into crises. In 2023 [7], the platform was reintroduced with significant improvements, enabling the collection of behavioral and interpersonal data through natural smartphone use which enabled collection of reported self-ratings known as ecological momentary assessments used in this research.

Previous research using EARS has explored various applications. For instance, one study examined the use of mobile sensing data to assess stress by analyzing affective language captured via smartphone keyboards [4]. Another study investigated the role of friendship quality and well-being in adolescence [9], concluding that adolescents who experienced more positive affect also reported more positive characteristics of close friendships two hours later.

In the present study, profiles of adolescents were identified using EMA variables, resulting in four distinct groups. These profiles were subsequently analyzed with respect to survey data and passive mobile sensing data. The study was guided by the following research questions:

- What distinct adolescent profiles emerge from EMA-based composite variables?
- How are these profiles associated with demographic and psychosocial survey measurements (gender, academic achievement, perceived overuse of social media, level of depression, anxiety, and stress symptoms)?
- What patterns of mobile application use characterize the identified profiles?

The rest of the paper is organized in the following way: in the second section materials and methods are described, the third section presents the results of data analysis, and the fourth section offers a discussion of results and conclusion.

2 Materials and methods

A sample of 77 Croatian high school students participated in this study. We employed three types of data: (1) survey data, (2) EMA data aggregated into six composite variables (confidence, attentiveness, positive and negative emotions related to friends, and overall positive and negative affect), and (3) passive mobile sensing data related to mobile applications usage. The survey

Information Society 2025, 6–10 October 2025, Ljubljana, Slovenia

http://doi.org/10.70314/is.2025.sikdd.29

data included respondents' gender, academic achievement (final grades of 3, 4, or 5), self-reported perceptions of overuse of social media (measured on a scale from 14 to 70), and symptoms of depression, anxiety, and stress (determined by DASS-21 scale, each measured on a scale from 0 to 21 [1]). EMA data and passive mobile data were collected using the EARS application.

Within the framework of ecological momentary assessment (EMA), respondents reported on the quality of their close friendships and their affect, seven times per day over the course of one week (i.e., up to 49 assessments). The assessment schedule followed a semi-random structure: respondents received questions at random intervals within 2-hour windows between 7 a.m. and 9 p.m. Only respondents who completed at least 10 out of 49 assessments were included in the analyses.

Friendship quality was measured with five items rated on a scale from 1 (not at all like me) to 7 (completely like me). All items were adapted from prior studies on close relationships [3, 5, 8]. Two composite variables were derived: *PosFriendEmo*, calculated as the average of three items related to positive friendship-related emotions, and *NegFriendEmo*, calculated as the average of items reflecting negative friendship-related emotions. Items related to positive friendship-related emotions were following:

- "I feel that I can share some worries or secrets with my close friends."
- "I enjoy being with my close friends."
- "I have fun with my close friends."

Items related to negative friendship-related emotions included following statements:

- "I feel that my close friends criticize me."
- "My close friends get on my nerves."

Affect was measured with ten items on the same 7-point scale, adapted from [3]. Two composite variables were created: *PosAffects* (joyful, cheerful, happy, lively, proud) and *NegAffects* (guilty, angry, insecure, scared, sad, worried, ashamed), representing the mean values of the respective items. In addition, a composite variable *Confident* was formed from three items related to peer popularity, self-satisfaction, and body satisfaction, while a composite variable *Attentive* was formed from five items reflecting responsibility, caring for others, perceived adult support, readiness for schoolwork, and perceived teacher support.

Regarding passive data, respondents used a total of 927 applications, which were categorized into 16 groups. Of these, 11 categories were included in the analysis, while the remaining five were excluded due to their negligible usage time. Initial categorization was performed using generative AI tools (Google Bard and ChatGPT) based on app functionality. Each app's classification was then manually verified through its official website to confirm its primary function. Beside variables related to usage of 11 observed categories of mobile apps, variable reflecting the total time spent on the mobile phone (*Total passive*) was also included into the analysis. The analyzed categories included: *Tools and productivity*, *Social media*, *Music and audio*, *Games*, *Communication*, *Multimedia*, *Education and learning*, *Online shopping and services*, *Travel*, *Device management*, and *Entertainment*.

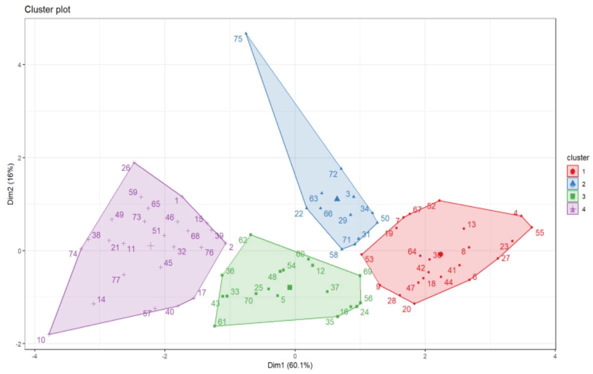

Figure 1. Groups obtained by k-means algorithm projected to the first two principal components of composite EMA variables.

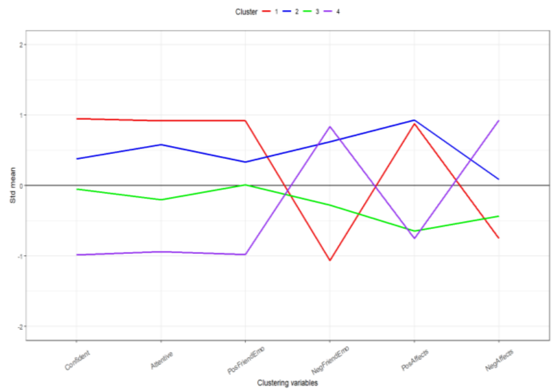

Figure 2. Mean values of standardized composite variables by groups.

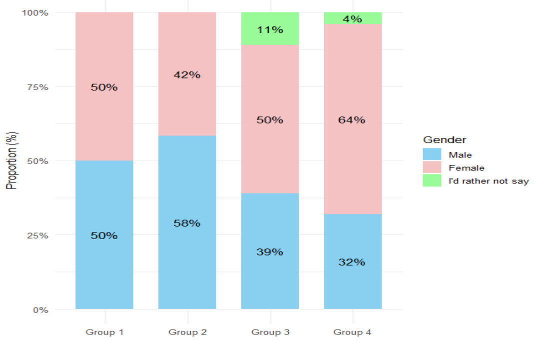

Figure 3. Proportion of respondents by group and gender (male, female, I'd rather not say).

Profiles of adolescents were identified using k-means clustering applied to standardized composite EMA variables. Based on the interpretability of the resulting clusters, the model with four groups was selected.

Data analysis was conducted using **R** statistical software. Group differences were tested using the non-parametric Kruskal–Wallis test, followed by Dunn's post hoc test. Non-parametric tests were applied because analyzed variables were not normally distributed. For the analysis of dependency between groups and their school success it was used chi-square test.

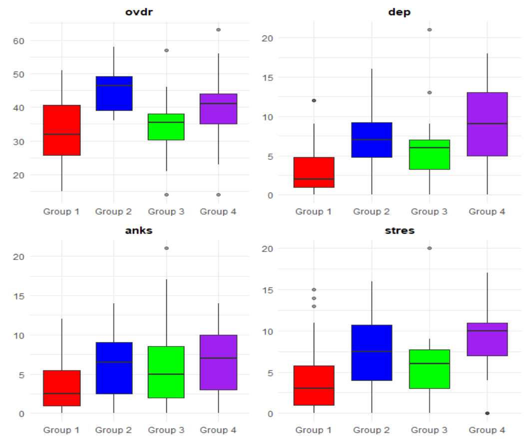

Figure 4. Box-plots for variables of self-assessment of overuse of social media (*ovdr,* 14-70), level of symptoms of depression (*dep*, 0-21), level of symptoms of anxiety (*anks*, 0-21), and level of symptoms of stress (*stres*, 0-21).

3 Results

Figure 1 shows groups of respondents obtained by k-means algorithm projected to the first two principal components of composite EMA variables. Figure 2 illustrates the mean values of the composite variables across groups. Two related pairs of groups can be observed: Groups 1 and 4, and Groups 2 and 3. Groups 1 and 4 display nearly mirror-image profiles with respect to the x-axis. For Group 1, the composite variables *Confident*, *Attentive*, *PosFriendEmo*, and *PosAffects* are above average, whereas in Group 4, these same variables fall below average. Conversely, *NegFriendEmo* and *NegAffects* are below average for Group 1 but above average for Group 4. A similar pattern emerges for Groups 2 and 3, which also show mirror-image profiles, though shifted slightly toward above-average values. Group 3 is characterized by nearly average levels of *Confident*, *Attentive*, and *PosFriendEmo*, while *NegFriendEmo*, *PosAffects*, and *NegAffects* are slightly below average. In contrast, Group 2 demonstrates above-average mean values across all variables. Overall, emotions related to friendships and affective states are less pronounced in Groups 2 and 3 compared to Groups 1 and 4.

Figure 3 shows that female respondents predominate in Groups 3 and 4, in Group 1 there is approximately an equal proportion of male and female respondents, while in Group 2 predominate male respondents. Figure 4 presents the distribution of survey-based variables: self-assessment of overuse of social media (*ovdr,* 14-70), level of symptoms of depression (*dep*, 0-21), anxiety (*anks*, 0-21), and stress (*stres*, 0-21). Group 4 exhibits the highest levels of symptoms of depression, anxiety, and stress. According to the non-parametric Kruskal-Wallis test, there is a significant difference between the groups in symptoms of depression ($p=0.0045$) and stress ($p=0.0162$). The Dunn's post hoc test indicated that Group 4 has statistically significant higher levels of symptoms of depression ($p=0.0015$) and stress ($p=0.0090$) compared to Group 1. The Kruskal-Wallis test shows that there is a difference in the perception of overuse of social media between the groups ($p=0.0024$). The highest perceived overuse was reported by Group 2, with a significant difference compared to Group 3 ($p=0.0021$) and Group 1 ($p=0.0040$). Results indicate that respondents' perceptions of their social media use did not correspond to the actual time spent on social media ($r = 0.0741$).

Figure 5 presents the distribution of daily time (in seconds) that respondents spent using different categories of mobile applications across groups. No statistically significant differences were found in the median time spent on social media or in the total time spent across all application categories. Group 1, which showed the highest median values for the composite variables *Confidence*, *Attentiveness*, and positive friendship-related emotions, also reported spending the most time on social media; however, their perception of social media overuse was the lowest among all groups. Group 3, characterized by near-average median values of *Confidence*, *Attentiveness*, positive and negative friendship-related emotions, and affect, demonstrated the highest median usage across most application categories (*Tools and productivity*, *Music and audio*, *Games*, *Communication*, *Education and learning*, *Travel*, *Device management*, and *Entertainment).* The Kruskal-Wallis test revealed a significant difference in application use only for the *Education and learning* category, although Dunn's post hoc test did not confirm differences between specific group pairs. Respondents in Group 4 had the highest median usage of *Multimedia* applications, while those in Group 2 spent the most time on applications related to *Shopping and services*. Notably, respondents in Group 2 were predominantly male and reported the highest perceived overuse of social media among all groups. School success was measured by average grade point, which was 4.05 for Group 1, 4.33 for Group 2, 4.61 for Group 3, and 4.20 for Group 4. The chi-square test indicated a borderline non-significant difference in school success across the groups ($p=0.0501$). Group 3, which showed the highest median time of application use across most categories, also achieved the highest average grade point (4.61). In contrast, Group 1, which reported the highest levels of confidence and attentiveness in EMA (including perceived readiness for school tasks), had the lowest average grade point.

4 Discussion and conclusion

This study identified four adolescent profiles based on data collected from 77 Croatian high-school students using EMA. Data collected from EMA was aggregated across respondents in the form of 6 composite variables representing their self-reported confidence, attentiveness, positive and negative friendship-related emotions, and positive and negative affect. Two pairs of mirror-image profiles were observed: Groups 2 and 3, and Groups 1 and 4. Emotional states related to friendships and affective states are less pronounced in Groups 2 and 3 compared to other pair of groups, and these groups are characterized by better academic success.

Mobile sensing revealed that respondents used a total of 927 apps, which were categorized into 16 categories, out of which 11 were analyzed in this study. Although social media accounted for the largest share of usage time, no significant group differences were found either in social media use or in total application use. Group 1, according to self-perception, exhibited the most confident and attentive and has lowest median levels of depression, anxiety and stress, spent the most time on social media, but perceived its overuse the least. This group contains approximately an equal proportion of male and female respondents. Group 2, which was predominantly male, spent the most time on *Online shopping and services* and reported the highest perceived overuse of social media, with significant differences compared to Group 1 ($p=0.0040$) and Group 3 ($p=0.0021$).

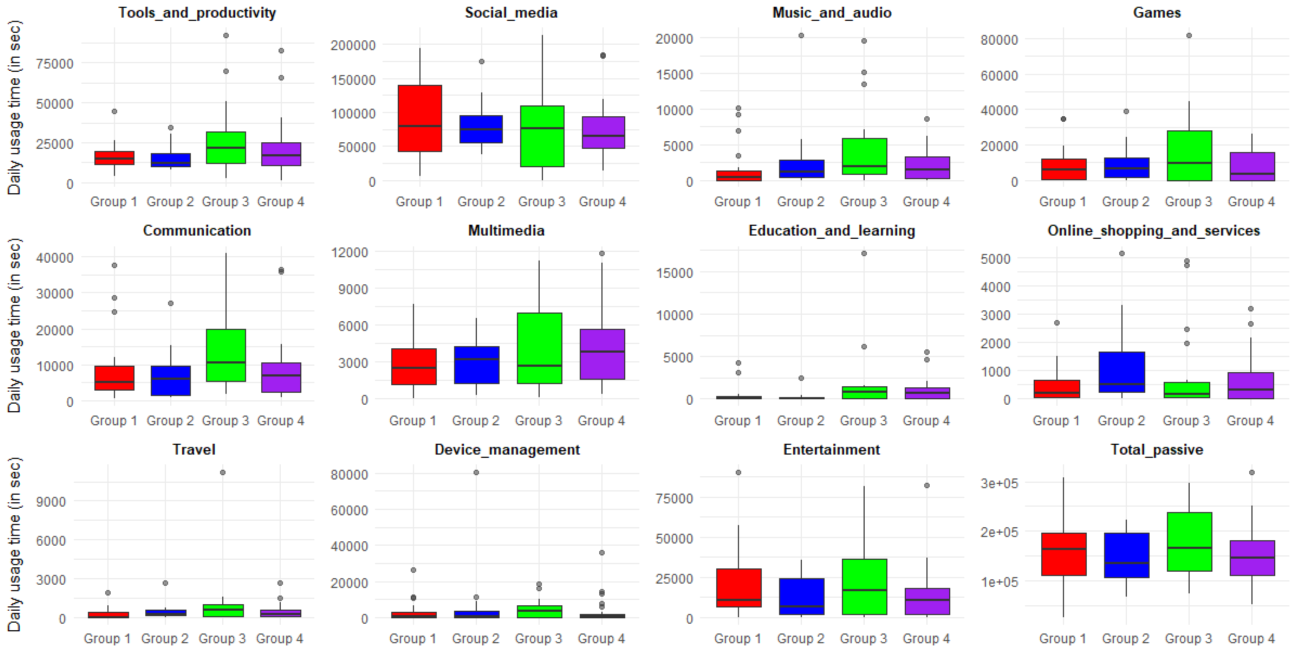

Figure 5. Box-plots for variables of daily usage of categories of mobile applications by groups (in seconds). Note the different ordinal scales due to the large differences in the use of apps.

Group 3, which had the highest academic achievements and the majority of female respondents, had the highest usage of applications in categories *Tools and productivity*, *Music and audio*, *Games*, *Communication*, *Education and learning*, *Travel*, *Device management*, and *Entertainment.* Group 4, also predominantly female, exhibited the highest levels of depression, anxiety, and stress symptoms, spent the least time on social media, used *Multimedia* applications more than other groups, and ranked second in the use of *Education and learning* applications. Importantly, there was no significant correlation between perceived overuse of social media by respondents and their actual time spent using it, as measured by passive sensing. This finding highlights the added value of combining mobile sensing with survey data, as it provides insights that would not be captured through self-report alone. While symptoms of depression, anxiety, and stress were assessed on a 0-21 scale, all median values were below 10, reflecting the general population sample in which the prevalence of psychological problems is low. Future research could therefore focus on adolescents with higher levels of depression, anxiety, and stress symptoms.

In addition, future work will explore the application of symbolic data analysis for clustering based on both EMA and mobile sensing data. Symbolic data analysis, developed for the study of complex and large-scale datasets, incorporates variability directly into the aggregation process [2]. This approach would allow us to account for the stability of emotional states and behavioral patterns at the individual level, potentially offering more refined indicators for defining adolescent profiles.

Acknowledgments

This study was conducted as a part of the Testing the 5C framework of positive youth development: traditional and digital mobile assessment (P.R.O.T.E.C.T.) research project, founded by the Croatian Science Foundation (UIP-2020-02-2852).

References

[1] Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. 1998. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, 10(2), 176–181. https://doi.org/10.1037/1040-3590.10.2.176

[2] Billard, L., Diday, E. 2007. Symbolic Data Analysis: Conceptual Statistics and Data Mining, 1st edition, Wiley

[3] Bülow, A., van Roekel, E., Boele, S., Denissen, J.J.A. and Keijsers, L.. 2022. Parent –adolescent interaction quality and adolescent affect: An experience sampling study on effect heterogeneity. *Child Development*, 93(3), 315-331, DOI: https://doi.org/10.1111/cdev.13733

[4] Byrne, M.L., Lind, M.N., Horn, S.R., Mills, K.L., Nelson, B.W., Barnes, M.L., Slavich, G.M. and Allen, N.B. 2012. Using mobile sensing data to assess stress: Associations with perceived and lifetime stress, mental health, sleep, and inflammation, *Digital Health*. 2021:7, DOI: 10.1177/20552076211037227

[5] Li, L.M.W., Chen, Q., Gao, H., Li, W.Q. and Ito, K.2021. Online/offline self-disclosure to offline friends and relational outcomes in a diary school: The moderating role of self-esteem and relational closeness. *International Journal of Psychology*, 56(1), 129-137, DOI: https://doi.org/10.1002/ijop.12684

[6] Lind, M.N., Byrne, M.L., Wicks, G., Smidt, A.M., Allen, N.B., 2018. The Effortless Assessment of Risk States (EARS) Tool: An Interpersonal Approach to Mobile Sensing, *JMIR Ment. Health*, 2018; 5(3):e10334, DOI: 10.2196/10334.

[7] Lind, M. N., Kahn, L. E., Crowley, R., Read, W., Wicks, G., Allen, N. B., 2023. Reintroducing the Effortless Assessment Research System (EARS), *JMIR Ment. Health*, 2023; 10:e38920, DOI: 10.21196/38920.

[8] Ng, Y.T., Huo, M., Gleason, M.E., Neff, L.A., Charles, S.T. and Fingerman, K.L. 2021. Friendship in old age: Daily encounters and emotional well-being. *The Journals of Gerontology: Series B*, 76(3), 551-562, DOI: https://doi.org/10.1093/geronb/gbaa007

[9] Šutić, L., van Roekel, E. and Novak, M. 2025. Quality of friendships and well-being in adolescence: daily life study. *International Journal of Adolescence and Youth*, 30(1), DOI: https://doi.org/10.1080/02673843.2025.2467112