Girardini et al. *EPJ Data Science* (2023) 12:55
https://doi.org/10.1140/epjds/s13688-023-00429-y

EPJ Data Science
a SpringerOpen Journal

REGULAR ARTICLE Open Access

Adaptation of student behavioural routines during Covid-19: a multimodal approach

Nicolò Alessandro Girardini[1,2*], Simone Centellegher[1], Andrea Passerini[2], Ivano Bison[3], Fausto Giunchiglia[2] and Bruno Lepri[1]

*Correspondence:
ngirardini@fbk.eu;
nicolo.girardini@unitn.it
[1]Fondazione Bruno Kessler (FBK), Via Sommarive 18, 38123, Trento, Italy
[2]Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 9, 38123, Trento, Italy
Full list of author information is available at the end of the article

Abstract

One population group that had to significantly adapt and change their behaviour during the COVID-19 pandemic is students. While previous studies have extensively investigated the impact of the pandemic on their psychological well-being and academic performance, limited attention has been given to their activity routines. In this work, we analyze students' behavioural changes by examining qualitative and quantitative differences in their daily routines between two distinct periods (2018 and 2020). Using an Experience Sampling Method (ESM) that captures multimodal self-reported data on students' *activity*, *locations* and *sociality*, we apply Non-Negative Matrix Factorization (NMF) to extract meaningful behavioural components, and quantify the variations in behaviour between students in 2018 and 2020. Surprisingly, despite the presence of COVID-19 restrictions, we find minimal changes in the activities performed by students, and the diversity of activities also remains largely unaffected. Leveraging the richness of the data at our disposal, we discover that activities adaptation to the pandemic primarily occurred in the *location* and *sociality* dimensions.

Keywords: Human behaviour; Behavioural Change; Activity Routines; Non-Negative Matrix Factorization; COVID-19

1 Introduction

The COVID-19 pandemic has had a wide-ranging impact on various components of our daily lives. The implementation of government restrictions such as travel bans, business and school closures, stay-at-home orders and physical distancing mandates aimed at preventing the spread of the virus [11, 19, 57]. These measures resulted in substantial changes in how we live, work, and socialize with long lasting impacts, among others, on human mobility and encounters [19, 48, 68], employment [1, 6], education [7, 65, 69], mental health and well-being [13, 23, 53].

Traditionally, social scientists have relied on qualitative techniques such as direct observation and fieldwork to study various aspects of our daily lives and gain insights into these social phenomena [14, 37]. However, these methods have limitations as they can be time-consuming, resource-intensive, and challenging to replicate, making them less suitable for large-scale studies.

In the last decades, the emergence of the digital sensing revolution has provided researchers with vast amounts of quantitative data generated by, among others, social media platforms, credit cards, and mobile phones. Digital data sources collected with these technologies offer unprecedented opportunities to gain insights into individual and group behaviours with a breadth and depth that was previously inconceivable [45, 46]. This wealth of digital data has become particularly valuable during the COVID-19 pandemic, enabling researchers to uncover patterns and trends in human behaviour and to understand the changes that have occurred in our daily lives and society as a whole [10, 41, 58].

While these data sources allow us to track the behaviour of thousands or even millions of individuals, it's important to note that the temporal resolution of the collected data is not continuous but rather tied to specific events (e.g., phone calls or traffic data for mobile phone data, app usage for GPS location data, credit card usage for spending data, etc.). Furthermore, these data sources often represent only one aspect of human behaviour at a time such as communication, spending, and mobility.

To overcome these limitations researchers have implemented "living labs" studies, which collect multiple sources of data [2, 18, 27, 34, 49, 51, 62]. In these studies, participants are equipped with mobile phones and/or sensing apps installed on their devices to collect data from diverse sources and sensors, including phone and SMS logs, Bluetooth proximity interactions, GPS location data, accelerometer data, etc. Additionally, surveys and experience sampling methodologies are employed to gather information on individual characteristics (e.g., attitudes, personality traits, etc.), on states (e.g., mood, stress levels, etc.) as well as situational and activity data. By integrating data from multiple sources, living labs deployments offer the advantage of capturing a richer representation of the individuals under study.

In our work, we leverage an Experience Sampling Method (ESM) [22] approach on smartphones to capture self-reported information on the daily activities of university students. ESM is a methodology aiming at collecting information on the behaviours and emotions of study participants throughout their daily activities and routines [55, 60, 64]. As in traditional time diary studies, ESM collects data by means of study participants' self-reports. However, participants, unlike in diary studies, are proactively triggered at various moments during the day. In our work, we ask the study participants to report information on their *activities* with a prompt every 30 minutes, and we collect not only the activity of the students but also their *location* and *sociality* (i.e., with whom they were). Hence, this approach produces a unique understanding of students' behaviour from the point of view of these three different behavioural dimensions. Moreover, the ESM-based data collection happened in two periods, in 2018 and 2020, and we leverage these data to analyse how students' routines changed across all three behavioural dimensions during the COVID-19 pandemic.

As initial step, we delve into the shifts in students' behaviours by examining the raw frequencies of activities. By comparing the distributions of the 2018 and 2020 samples, we identify the most noteworthy changes. Our analysis reveals that common activities, such as moving and engaging in social activities, experienced a decrease in their frequency, while others, like watching shows, saw an increase in prominence. Concurrently, there was a general decrease in location frequencies, compensated with a significant increase in the frequency of the home location. A similar trend is observed in the dimension of social interactions, where even interactions within the same household decreased, leading to an increase in alone time.

To place these results within the context of daily time allocation, we further investigate how students' routines were impacted. To this end, we use *Non-Negative Matrix Factorization* (NMF) [20, 61] to extract meaningful behavioural components (i.e., routines) from self-reported data. Indeed, NMF unveils underlying behavioral patterns not readily apparent in the data while also reducing irrelevant or noisy information. This results in the extraction of more interpretable behavioral components when compared to alternative methods based on Principal Component Analysis (PCA). These alternative techniques often face issues related to interpretability due to their lack of a non-negativity constraint. As a result, they may extract less meaningful routines (see Sect. S3 in the Additional file 1 for the routines obtained using Principal Component Analysis and Robust Principal Component Analysis).

Hence, by analyzing the extracted components and their associated weights, we are able to quantify the differences in behaviour between students in 2018 and students in 2020. Our findings indicate that despite the impact of the pandemic, the *activity* routines of students in 2018 and 2020 remained largely similar with minimal differences observed. However, with the multimodal information collected through the time diaries, we discover that the behavioural dimensions of *location* and *sociality* exhibit stronger signals of adaptation. To delve deeper into this shift of behaviour, we first examine the relationship between *activities* and *locations*, then the relationship between *activities* and *sociality*. This approach allows us to identify the activities that were associated with diverse *locations* and *sociality* routines before and during the pandemic. Notably, activities such as studying and attending lessons, which had a social component before the pandemic, transitioned to solitary activities, done in private homes, even if students could go to the university. We find that most of the extracted routines in 2020 follow this pattern, and take place at home and alone, while in 2018 students visited more diverse locations and spent more time with friends. For example, social life is more present in 2018, with students spending time with friends, while in 2020 this activity is done with roommates, thus people residing in the same household.

2 Related work

2.1 Behaviours and routines modelling

In recent years, the field of modelling human behaviour and routines has witnessed significant growth, thanks to advancements in mobile sensing [44] and experience sampling approaches [60].

Various techniques have been employed to understand individual behaviour. Principal Component Analysis (PCA) was used to extract students' routines (*eigenbehaviours*) [28], from the Reality Mining dataset [27]. Techniques typically used for Natural Language Processing tasks [71] have been used to extract regular behavioural patterns from sequential data, such as Latent Dirichlet Allocation [8], which was also used to group individual activities into routines [29, 42]. Other sequence analysis techniques have been used as well [12, 63], from modelling purchasing behaviour [26] to the analysis of individual mobility [48].

Additional methods include more complex tools, such as Hidden Markov Models [59], Conditional Random Fields [30] and Bayesian Networks [67]. These approaches are commonly used to build predictive and generative models, and can be used to detect the activities from sensor data and images [30, 67] or to model cyclic behaviour and its anomalies [59]. While these methods have their advantages, in our study, we aim to identify easily interpretable and human-readable patterns to compare and measure behavioural differences.

In recent years, *Non-Negative Matrix Factorization* (NMF) [61] has gained popularity, thanks to its interpretability. NMF has proven to be effective in identifying repeating patterns in the daily behaviours of people, both at an individual and a collective level. NMF has been successfully applied in various domains, such as in analyzing social interactions in networks [32], understanding of chronotypes [3], and identifying routines in urban environments [26, 70]. Inspired by these studies, we use NMF in our methodology, as it allows us to associate the resulting patterns with realistic individual behaviour, and the weights associated with these patterns provide a quantitative measure of behavioural differences. Additionally, following [73], we include in our methodology, the use of multiple modalities to analyse subjects' behaviour, which includes (i) activity, (ii) location, and (iii) sociality dimensions.

2.2 Behavioural change and the pandemic

There has been an upsurge of scientific research in response to the COVID-19 crisis aimed at understanding and mitigating the spread of the virus [4, 19]. The various consequences and effects of COVID-19 restrictions on the population, spanning economic, social, and psychological domains have been tackled in diverse disciplines.

From an economic perspective, studies have demonstrated that the effects of restrictions differ across individuals with varying income levels, often increasing segregation. For instance, in Italy, the mobility reduction was higher in high fiscal capacity municipalities, which also display high income inequality [9]. Similar findings have emerged in other countries such as France and the United Kingdom, where the mobility network became more fragmented on a national scale but maintained more connections on a small scale [31]. In the United States, researchers have observed a decrease in encounters between citizens of different income levels in urban areas, hindering economic recovery and city growth [68], despite mobility returning to pre-pandemic levels.

Another area of interest has been the well-being of the population, with a particular focus on students, enhancing research that was present even before the pandemic [66]. The impact of the pandemic on students' educational paths has been substantial [7], as shown in studies on academic performance and mental well-being [38]. Research has indicated a significant increase in stress and anxiety levels among younger students during lockdowns [39], particularly college and university students [52, 69].

Previous studies on students primarily rely on surveys and self-reported data, often neglecting the impact on behavioural routines. Furthermore, only a few studies consider multiple dimensions of behaviour and well-being with high-resolution data. For instance, the study conducted in [51] combines surveys with mobile sensing data to explore how students' perspectives and concerns regarding the pandemic lead to diverse responses to restrictions and shifts in behaviour.

In contrast, our work demonstrates the utility of Experience Sampling Methods on smartphones, which provide a wealth of information on activities, locations, and social interactions in a real-world setting. These methodologies offer a level of granularity, particularly in activity definition, that cannot be achieved through passive sensor data collected by mobile phones. Moreover, we highlight how these methodologies can be harnessed to understand multiple dimensions of students' behaviour, both independently and in conjunction with one another.

3 Data

The data utilized in this work are derived from two separate datasets collected during two distinct living lab studies: the *SmartUniTn* study [33, 35, 36] conducted in 2018, and the subsequent *WeNet* study [5, 34, 50] conducted in 2020.

The *SmartUniTn* study gathered both sensor data and questionnaire responses from students at the University of Trento, employing the *iLog* [72] smartphone application. Over a period of two weeks, with a frequency of 30 minutes, participants were requested to complete a *Time Diary* via the Experience Sampling Method (ESM) application. Each entry in the diary required students to report their *activity* (e.g., studying, eating, etc.), *location* (e.g., home, university, bar, etc.), and *sociality*, indicating whether they were alone or in the company of others (e.g., alone, friends, relatives, etc.). The response options were predefined to ensure uniformity and minimize user biases in activity descriptions (refer to [33] for the available sets of responses). This comprehensive dataset, which was collected from May 10th to May 23rd, 2018, provides a multimodal representation of student behaviour in a period when students attended classes and prepared for the summer exam session.

The subsequent *WeNet* study [34] expanded upon the *SmartUniTn* study by extending the data collection to include seven universities worldwide, including the University of Trento. To ensure consistency with the previous study, we only focus on the data from the latter university. The data collection process closely followed that of the *SmartUniTn* study, with refinements made to the sets of activities and locations (for detailed information, see [34]). Notably, the data collection occurred during the COVID-19 pandemic, and as such, specific government restrictions were in place at the time. The Time Diaries were collected between November 14th and November 30th, 2020, a period during which the government of the province of Trento implemented several restrictions [21]. Certain activities, such as gyms, were closed, while others, like shops, had limited access. Distancing measures were enforced in various contexts, including bars and restaurants, where in addition to physical distancing, time restrictions were imposed on the duration people could spend at these locations. A curfew was in place from 10 PM. Moreover, while schools were generally closed, first-year students participating in our study could attend lessons in person.

After completing the data collection and pre-processing phases, our dataset comprises samples of *128* students in the 2018 study and *119* students in the 2020 study, resulting in an almost balanced dataset with a total of *247* students.

4 Methods

4.1 Data processing

Since the data collection of the two studies was performed in different iterations, the student samples are different. However, the study settings remained consistent, and the population under examination was similar as both studies focused on students. To facilitate a more robust comparison between the groups of students in 2018 and 2020, we performed a mapping between the categories in the *Time Diaries* to ensure the comparability of activities, locations, and sociality categories. Given that the data collected in 2020 were more refined, the mapping primarily involved the alignment of certain activities and locations, where multiple entries from 2020 were mapped into a single entry from 2018. For instance, activities like *Free Time Study*, *Arts*, *Hobbies*, and *Games* in 2020 were all mapped to *Hobbies* as in 2018. A comprehensive list of activity and location mappings can be found in the Additional file 1, Sect. S1.

Subsequently, we organized the data to follow circadian cycles starting at 5 AM, ensuring that the daily Time Diaries of each student were appropriately aligned. We excluded days and students with insufficient reported data. Additionally, as our focus was on typical behaviour during a regular student week, we excluded weekends from our analysis.

Given that Time Diaries provide insights into three different facets of behaviour (*activity*, *location* and *sociality*), we created three distinct matrices, each corresponding to the type of information collected. In these matrices, the rows represent students, while the columns represent their behaviour throughout the day. To achieve this, we allocated a number of columns (or features) equal to $c = |A| * 48$, where A represents the set of possible choices for each data type (e.g., the possible activities for the activity matrix), and *48* represents the number of time intervals within a day (given that the data was collected every 30 minutes). Consequently, the matrices have dimensions of $N \times c$, with N denoting the *247* students from both datasets. Each matrix contains the normalized counts of observed behaviour (*activity*, *location* and *sociality*) within specific time slots for each student.

4.2 Methods for analyzing differences in students' behaviours

As a first step in understanding students' behavioural differences between 2018 and 2020, we examined the distributions of frequencies of activities, locations, and social interactions, to identify the behavioural facets that exhibited the most significant changes. To assess the significance of these differences, we employed three statistical tests, namely *t-test*, *Analysis of Variance (ANOVA)* and *Kolmogorov–Smirnov (KS) test*, to evaluate the distributions' characteristics across the 2018 and 2020 samples.

More specifically, for each behavioural dimension (i.e., activities, locations and sociality) we computed the frequencies of each of the possible choices in that particular dimension (e.g., for activities: *Break, Eating, Hobbies,* etc.), splitting the data into the 2018 and 2020 samples and comparing them with the aforementioned tests.

4.3 Methods for detecting students' daily behavioural routines

The focus of our paper is to analyse and investigate the behavioural routines of students, taking into account the temporal aspect of their daily routines, while simultaneously looking at different facets of students' behaviour (activities, location and sociality). We explored established methods for modelling and extracting behavioural routines from diary data. We tested the widely used matrix decomposition techniques, *Principal Component Analysis (PCA)* [43] and its extension, *Robust PCA (RPCA)* [16], which is more robust to the presence of outliers. These methods decompose data matrices into repeating patterns, with assigned weights for reconstructing initial observations. However, the presence of negative values in the extracted components and weights complicates their interpretability. Furthermore, routines produced by these methods often proved challenging to interpret, resulting in noisy patterns and redundant information. We thoroughly tested these methods on all the behavioural dimensions (activities, location and sociality) and we report the extracted components in the Additional file 1 Sect. S3.

As an example, both PCA and RPCA, in Component 2 of the location dimension (Figs. S2 and S5 in the Additional file 1 Sect. S3.1 and S3.2 respectively), display the redundant information that students are not at *Home* when they are at the *University*. Moreover, the larger the number of items in the original data, the more noisy and less interpretable the extracted components are (e.g., activity components are harder to interpret than location components, see Figs. S1–S2 in the Additional file 1 Sect. S3.1).

Therefore, to extract meaningful daily routines from our data, we applied *Non-Negative Matrix Factorization (NMF)* [20, 61], which follows the same principle of the aforementioned methods. However, NMF operates on non-negative matrices and decomposes them into two non-negative matrices, making the results easier to interpret, as previous applications in different fields, such as image analysis [47], text mining [54] and also in behavioural and routine analysis [3, 68], have shown. Moreover, NMF has been proven successful in analysing temporal data [3, 17, 40]. In our scenario, the use of this method allows for the exploitation of its clustering properties, for which the algorithm extracts coherent routines from similar users with similar behaviours.

4.3.1 Non-negative matrix factorization

The Non-Negative Matrix Factorization (NMF) algorithm,[1] identifies recurring patterns within the input data and approximates the original data by representing it as a linear combination of these patterns. In more detail, NMF decomposes each of our behavioural matrices, denoted as X, into an approximation $X \approx WK$. Here, K represents the $r \times c$ matrix of extracted components, where c denotes the number of features defining the behaviour (as defined in Sect. 4.1), and r represents the number of extracted components or *rank*. Consequently, the matrix K contains the weights of each feature for each component, while the matrix W has dimensions $N \times r$ and contains the weights of the components for each student. The algorithm leverages its inherent clustering property to identify common observations within X, storing the discovered repeating patterns in K.

For example, if we consider the *location* dimension of students' behaviour, the $N \times c$ *location* matrix X, contains in each row i the normalized counts of each location at each time of the day for a single student (e.g., *Home* at 8.30 AM, *Home* at 9.00 AM, etc.). The matrix of extracted components K captures common *location* routines among students, while the matrix W contains the weights of the extracted components for each student. The original behaviour of a student i can thus be reconstructed as: $x_i = w_{i1}k_1 + w_{i2}k_2 + \cdots + w_{ir}k_r$.

The approximation $X \approx WK$ is achieved by minimizing the reconstruction error, computed using the *Frobenius distance* between the original matrix and its approximation:

$$\text{Err} = \|X - WK\|_F = \sqrt{\sum_{i,j}\left(x_{i,j} - \sum_r w_{i,r}k_{r,j}\right)^2}, \tag{1}$$

where $x_{i,j}$ are items of the original data matrix X, $k_{r,j}$ are the values of the components and $w_{i,r}$ the weights of the components for every observation.

The NMF algorithm takes the rank r as a parameter, indicating the number of components to extract. To determine the optimal number of components, we used the *cophenetic correlation coefficient* [15], a common practice in the literature [3, 68]. The coefficient

[1] The algorithm was implemented using the *scikit-learn* Python package.

measures the preservation of the distance between different observations during the NMF transformation, indicating the reliability of the components. Typically, the rank is chosen where the cophenetic coefficient is maximized and begins to decline. In our analysis, we selected the number of components by considering the cophenetic coefficient, and when the choice was not evident, we prioritized the rank r that yielded the most interpretable components.

5 Results

5.1 Investigating students' behaviour in 2018 and 2020

The initial step in understanding the shifts in students' behaviours between 2018 and 2020, was to investigate possible differences in the frequencies of activities, locations, and sociality behavioural dimensions.

As described in Sect. 4.2, we first computed these frequencies for each student and then compared the distributions of students in the 2018 and 2020 samples. As previously anticipated, we employed three different statistical tests namely t-test, Analysis of Variance (ANOVA) and Kolmogorov–Smirnov (KS) test to evaluate the differences in the distributions' characteristics across the samples. Given that several distributions were skewed, we report here the significant differences measured with the KS test, while the similar results that we found using the other tests are reported in the Additional file 1 Sect. S2. The statistics and the test outcome for the three behavioural dimensions are presented in Table 1, Table 2, and Table 3 respectively.

Looking at the *activity* behavioural dimension, displayed in Table 1, we see several activity items exhibit variations across the 2018 and 2020 samples. Activities that show noteworthy differences include *Sleeping*, which slightly increased in 2020, *Moving* and *Social Life*, which decreased from 2018, and *Watching TV/Shows*, which increased in 2020. These changes are not surprising, considering the restrictions in place during the COVID-19 pandemic. It is interesting to note that some main activities, such as *Eating*, *Hobbies*, *Lesson* and *Study*, do not show significant changes between the two years.

Table 2 illustrates differences in students' locations. As expected, given the limitations on places, there are significant differences for all location items. The frequencies of all locations decreased, with the *Home* location more than doubling its frequency during the pandemic.

Similarly, social interactions underwent significant changes. As shown in Table 3, only interactions with *Colleagues* and *Others* do not exhibit a significant change, most likely because of their low frequency. It is noteworthy that interactions with *Classmates*, *Friends*, *Relatives*, and even *Roommates* and *Partners* all decreased from 2018 to 2020. Instead, as a result of restrictions and location changes, students spent significantly more time *Alone*.

It is important to note that for some of the items with significant change from 2018 to 2020 (e.g. *Reading*, *Gym* and *Other Library*), the fraction of students that engaged in those behaviours is not relevant. Another observation is that these significant differences pertain to activity frequency and do not provide a comprehensive understanding of students' time allocation throughout the day. For example, the activities that increased in frequency might account for the decrease in other activities. Therefore, it is essential to conduct an analysis that considers students' actual daily routines, which will be presented in the subsequent sections.

Table 1 Activity Frequencies. Mean frequency and variance of the students' activities during 2018 and 2020. The KS test measures whether the difference between the distributions is significant

Activity	Mean 2018	Mean 2020	Var 2018	Var 2020	KS stat	KS p-value
Break	0.01495	0.00835	0.00024	0.00012	0.28499	6.217e-05***
Eating	0.06623	0.06471	0.00035	0.00025	0.16636	5.657e-02
Hobbies	0.02957	0.02670	0.00197	0.00143	0.08384	7.347e-01
Housework	0.01262	0.03789	0.00016	0.00060	0.53164	5.551e-16***
Lesson	0.04606	0.05428	0.00130	0.00157	0.13892	1.636e-01
Movie...	0.00329	0.00104	0.00003	0.00001	0.23142	2.109e-03 **
Moving	0.06369	0.02339	0.00081	0.00032	0.63235	5.551e-16 ***
Other	0.02426	0.01189	0.00097	0.00026	0.31775	4.964e-06 ***
Phone	0.01430	0.02076	0.00042	0.00066	0.16295	6.521e-02
Reading/Music	0.00652	0.01083	0.00014	0.00034	0.19584	1.462e-02 *
Rest/Nap	0.01904	0.04562	0.00047	0.00144	0.45647	3.853e-12 ***
Selfcare	0.02941	0.03161	0.00021	0.00056	0.10471	4.660e-01
Shopping	0.00424	0.00584	0.00002	0.00003	0.27928	9.458e-05 ***
Sleeping	0.34371	0.35381	0.00079	0.00164	0.25361	5.379e-04 ***
Social life	0.05453	0.02453	0.00167	0.00056	0.39752	3.033e-09 ***
Social Media	0.01149	0.02580	0.00018	0.00091	0.25670	4.397e-04 ***
Sport	0.01162	0.00972	0.00031	0.00017	0.11417	3.604e-01
Study	0.13214	0.13305	0.00629	0.00612	0.07491	8.430e-01
TV-Shows/Youtube	0.04360	0.06529	0.00125	0.00191	0.25906	3.764e-04 ***
Work	0.01486	0.01019	0.00086	0.00099	0.22801	2.592e-03 **

Significance level: '***' $\leq$ 0.001, '**' $\leq$ 0.01, '*' $\leq$ 0.05.

Table 2 Location Frequencies. Mean frequency and variance of the students' locations during 2018 and 2020. The KS test measures whether the difference between the distributions is significant

Location	Mean 2018	Mean 2020	Var 2018	Var 2020	KS stat	KS p-value
Friends' Home	0.04642	0.02874	0.00509	0.00320	0.34959	3.191e-07 ***
Gym	0.00539	0.00074	0.00014	0.00003	0.33357	1.297e-06 ***
Home	0.31914	0.66480	0.07100	0.10840	0.63203	5.551e-16 ***
Moving	0.06587	0.02207	0.00087	0.00031	0.67260	5.551e-16 ***
Other Indoor	0.01846	0.00693	0.00218	0.00086	0.37126	4.167e-08 ***
Other Library	0.00305	0.00002	0.00011	0.00003	0.20313	1.006e-02 *
Outdoors	0.04189	0.00999	0.00101	0.00024	0.66964	5.551e-16 ***
Bar/Restaurant	0.01418	0.00433	0.00017	0.00004	0.51543	5.551e-16 ***
Relatives' Home	0.30275	0.19535	0.07393	0.11185	0.40060	2.186e-09 ***
Supermarket/Shop	0.00450	0.00588	0.00003	0.00008	0.26425	2.686e-04 ***
University	0.10523	0.01734	0.00464	0.00199	0.73523	5.551e-16 ***
Work Place	0.01331	0.00361	0.00082	0.00029	0.20444	9.406e-03 ***

Significance level: '***' $\leq$ 0.001, '**' $\leq$ 0.01, '*' $\leq$ 0.05.

Table 3 Sociality Frequencies. Mean frequency and variance of the students' company during 2018 and 2020. The KS test measures whether the difference between the distributions is significant

With Whom	Mean 2018	Mean 2020	Var 2018	Var 2020	KS stat	KS p-value
Alone	0.49799	0.63072	0.02913	0.05023	0.38157	1.501e-08 ***
Classmate(s)	0.05904	0.01998	0.00313	0.00181	0.52797	5.551e-16 ***
Colleague(s)	0.00953	0.00441	0.00057	0.00035	0.16124	6.983e-02
Friend(s)	0.12447	0.03745	0.00936	0.00282	0.53460	5.551e-16 ***
Other	0.01173	0.00628	0.00150	0.00010	0.07957	7.891e-01
Partner	0.07974	0.07412	0.01756	0.02626	0.23241	2.009e-03 **
Relative(s)	0.09628	0.15208	0.01587	0.03974	0.18284	2.725e-02 *
Roommate(s)	0.06139	0.03475	0.00925	0.00867	0.34769	3.822e-07 ***

Significance level: '***' $\leq$ 0.001, '**' $\leq$ 0.01, '*' $\leq$ 0.05.

5.2 Students' routines changes

Given the differences found in the previous section, it is still unclear how the actual daily routines of students changed during the pandemic. In fact, routines contextualise the raw frequency of behaviours, accounting for the temporal dimension of students' day.

To do so, we applied NMF, as described in Sect. 4, on the matrices with students from both the 2018 and 2020 data collection campaigns. Including both samples allowed us to analyze the behaviour of the two groups of students and to extract the most dominant students' routines in both 2018 and 2020. By using the three separate matrices that represented *activity*, *location*, and *sociality* behaviours, we identified 6 *activity* routines, 4 *location* routines, and 7 *sociality* routines.

In Fig. 1 and Fig. 2, we observe the 6 *activity* components (or routines) extracted by the algorithm. Note that the order of the components is not related to their significance, unlike other techniques such as Principal Component Analysis. The heatmaps (Fig. 1(a) and Fig. 2(a)) illustrate the weight of each activity at specific times of the day, with higher weights indicating a higher occurrence of the activity. As it is possible to observe, during the day, certain routine patterns are more prominent than others, with essential activities like *Eating* and *Sleeping* being influenced by other activities of the day. For example, this can be seen in Component 5, where students go to sleep later than other components (e.g., Component 6) because they watch YouTube and TV shows. Furthermore, NMF correctly extracts routines related to a student's academic life, and activities such as attending lessons and studying exhibit high weights in three components (3, 4, and 6). However, other behavioural patterns are less clear. Components 1 and 2 (in Fig. 1(a)) include several activities that share time allocations, indicating more complex daily routines. Component 1 is the routine that mostly captures social life, with higher weights for commuting (Moving) and social activities during the late afternoon and evening (Social Life).

A natural question that follows after looking at the most dominant *activity* routines extracted by the NMF algorithm, is whether the behaviour described by a component is more associated with typical behaviours of students in 2018 or 2020. We can investigate such association by looking at the weights extracted by the NMF algorithm present in matrix W (see Sect. 4.3.1).

The weights assigned to students on the extracted components provide a measure of the presence of each component in their behaviour. Higher weights indicate a stronger association with the respective NMF component. Figure 1(b) and Fig. 2(b) display the distributions of component weights for students in 2018 and 2020 respectively. A noticeable difference is observed in the first component, characterized by students *Moving* and engaging in evening social activities, which is more closely associated with a behaviour typical of 2018, as indicated by the higher values of the weight distribution. The distribution for students in 2020 is much more skewed towards 0, meaning that this kind of behaviour was not present, probably due to the COVID-19 restrictions imposed during 2020 (see Sect. 3). Also, Component 6 (Fig. 2) displays a (small) distinction between the 2018 and 2020 students groups, with typical behaviours in which students study during the day and watch YouTube, TV shows, or similar activities in the evening. Here, the weight distribution for students in 2020 exhibits a heavier tail and a second saddle point, with larger values compared to the weights of 2018 students, thus associating this more sedentary routine more with 2020 than 2018.

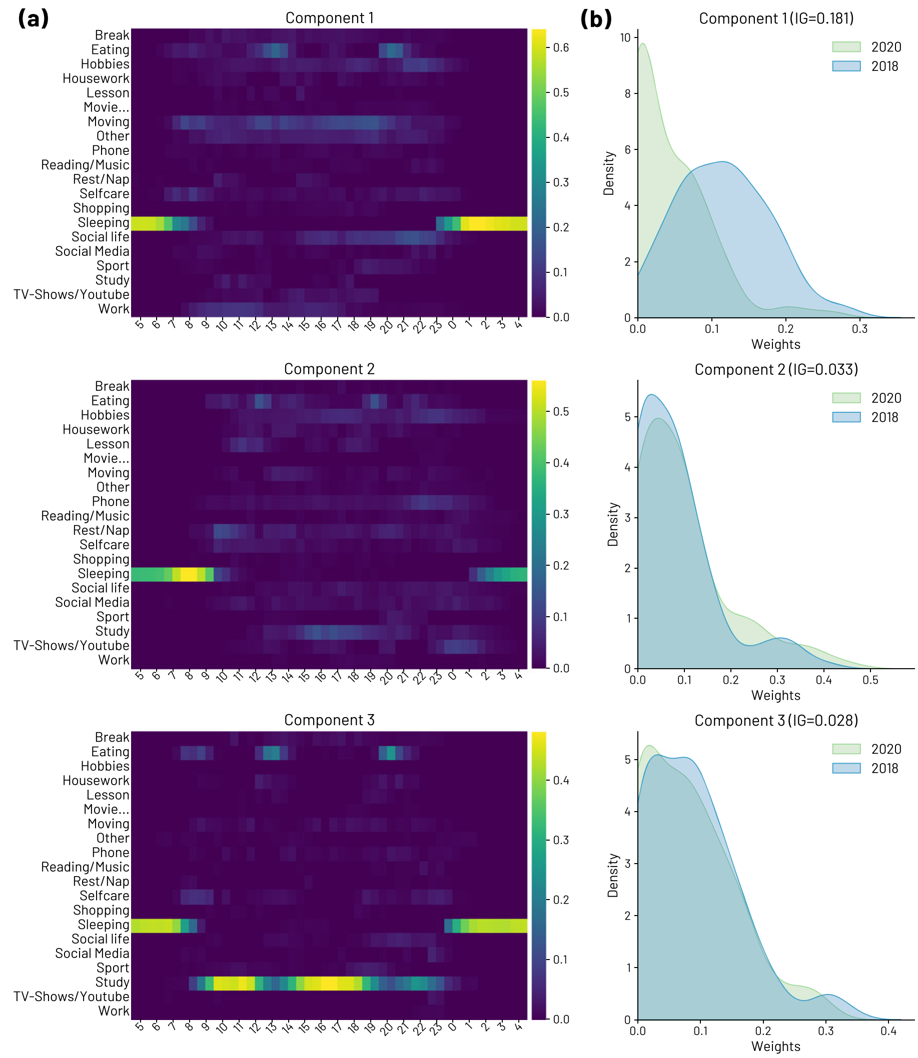

Figure 1 Activity Routines 1–3. In **(a)** we can see the components 1–3 extracted by the NMF algorithm for the *activity* dimension. The vectors representing each component in the matrix *K* have been visually rearranged in a matrix format, so that the rows represent the different possible activities, while columns are the 48 bins in which the day has been divided (i.e., every 30 minutes). Thus, the heatmaps shows how present an activity is in a specific time slot. In **(b)** we can see the distributions of the weights each student (divided by year) has for the components, showing how much that routine is important for the student's behaviour. We see that only the first component has different distributions for 2018 and 2020. The Information Gain (IG) shown quantifies how much knowing the component can tell apart 2018 behaviours from 2020 behaviours, and thus quantifies the weights distributions difference

Examining the weight distributions of the other components, it becomes evident that there are just small differences between the two groups of students. This observation is quantitatively supported by the Information Gain (IG), computed for each of the components, which quantifies the amount of information that knowing the weight of a component gives in discriminating whether the described behaviour is typical of a student in 2018 or 2020 (higher IG values indicate a larger separation in the distributions). As reported in Fig. 1(b) and Fig. 2(b), we can see that the first component carries the most informative weight, while the remaining components do not exhibit notable differences between 2018 and 2020.

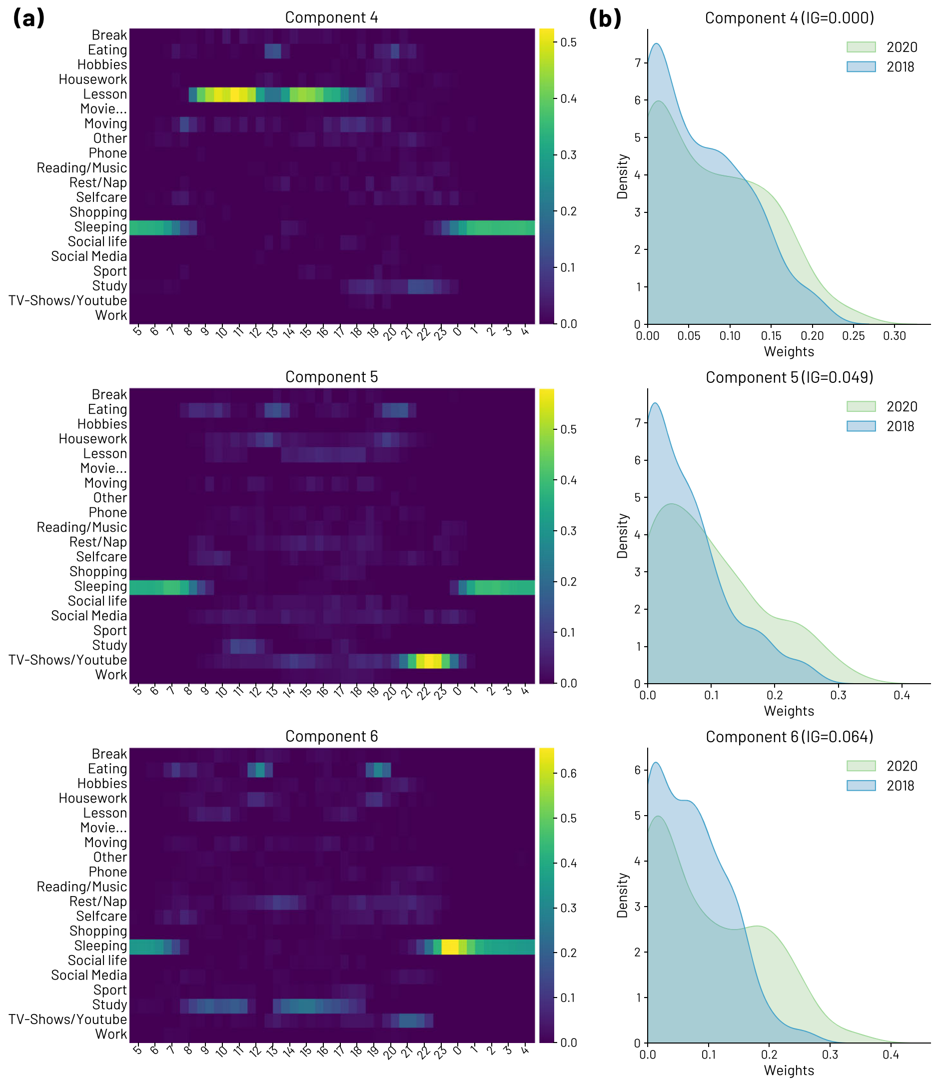

Figure 2 Activity Routines 4–6. Similar to Fig. 1, **(a)** shows the NMF extracted components 4–6 for the *activity* dimension, while **(b)** shows the students' weights distribution associated with them. We see that in this case the distributions do not differ much between 2018 and 2020. In fact, the IG is very low for each component

Interestingly, the extracted *activity* routines displayed minimal changes between 2018 and 2020. Despite the pandemic restrictions, students in Trento maintained their activity behaviours focused on lessons and studying, results which are consistent with our findings in Sect. 5.1.

As detailed in the Additional file 1 (the Additional file 1 Sect. S4), more pronounced differences were observed in the locations visited and in the social interactions between the two time periods. These disparities were expected due to the imposed restrictions, which limited mobility and reduced interpersonal interactions in 2020. Among the four *location* routines identified through the NMF decomposition, staying at *Home* throughout the day exhibited a significantly stronger association with students in 2020, while going to the *University* was less common in 2020 compared to 2018 (Components 1 and 4, see

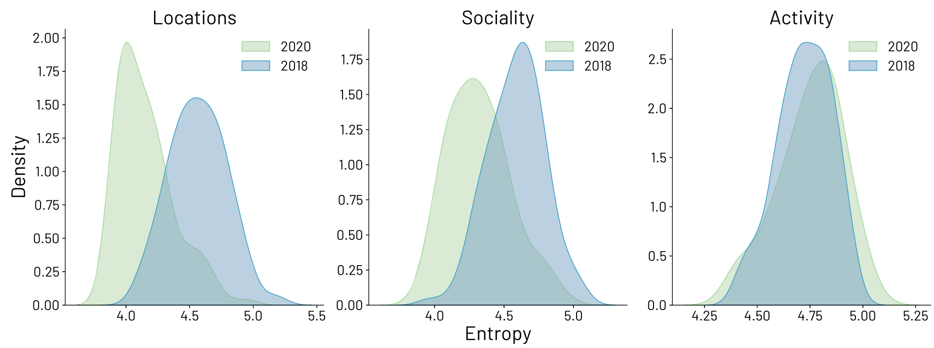

Figure 3 Entropy for the Types of Information. The three plots show the distribution of the entropy, i.e. diversity, computed on 2018 and 2020 students' reported behaviour for *location*, *sociality* and *activity* respectively. We see that the variety of activities remains the same, but students visited less diverse locations and have less diverse social interactions during the COVID-19 pandemic

Fig. S7 in the Additional file 1 Sect. S4). These differences are supported by higher values of Information Gain.

Furthermore, the imposed restrictions also affected social gatherings, which, along with limitations on location visits, influenced students' social behaviour. Our analysis revealed differences in three out of seven extracted *sociality* components. These three routines encompassed spending a day *Alone*, spending a day with *Friends*, and a working day with *Classmates* (Components 5, 6, 7, see Fig. S8 in the Additional file 1 Sect. S4). As expected, higher weights were observed for a day spent alone among 2020 students, while routines involving friends and classmates showed higher weights for 2018 students.

To support our findings, we conducted a quantitative evaluation of the diversity of behaviours both in 2018 and 2020. To measure this diversity, we calculated the Shannon's entropy [24–26] of students' behaviour using the *activity*, *location*, and *sociality* matrices. The distributions depicted in Fig. 3 represent the entropy computed for each student's average daily behaviour (i.e., using the normalized counts of each matrix) across the three types of information captured in the Time Diaries. This results in a measure of how diverse, on average, the students' routines are across the measurement period. Intuitively, lower values of entropy indicate a lower diversity in a specific behavioural dimension.

The results clearly indicate that in the *activity* dimension, the diversity of activities performed by students in 2018 and 2020 is similar (KS-test: stat = 0.1463, p-value = 0.1252; T-test: t = –1.483, p-value = 0.139), suggesting that the variety of activities did not decrease significantly in 2020. However, in the *location* (KS-test: stat = 0.6769, p-value = $5.551e^{-16}$; T-test: t = 14.403, p-value = $1.734e^{-34}$), and *sociality* (KS-test: stat = 0.4938, p-value = $3.220e^{-14}$; T-test: t = 9.748, p-value = $3.498e^{-19}$) dimensions, the entropy for 2020 is significantly lower than the entropy of 2018. This indicates that the COVID-19 restrictions had a noticeable impact on the diversity of behaviours in these dimensions, with a reduction in the range of locations visited and the social interactions engaged by students in 2020. The lower entropy values suggest a more limited range of options and routines in terms of locations and social behaviours due to the pandemic restrictions.

5.3 Multimodal activity routines: where and with whom

In the previous sections, we observed that despite the COVID-19 pandemic's government restrictions, there were minimal changes in the *activities* performed by students in 2020 compared to those in 2018. However, as expected, the *location* and *sociality* dimensions of their behaviour exhibited larger variations. This observation raised the question of how these two dimensions changed in conjunction with the activities, specifically where and with whom these activities were performed.

To gain insights into these variations, we applied the NMF decomposition once again, but this time we concatenated the behavioural matrices. Initially, to examine changes in locations for the activities, we concatenated the *activity* behaviour matrix with the *location* behaviour matrix. Subsequently, we explored the social dimension of activities by concatenating the *activity* and *sociality* behaviour matrices. Additionally, to enable a more focused analysis, we separated the groups of students in 2018 and 2020. By doing so, the NMF algorithm can extract the most dominant behaviours specific to each group. This approach allowed us to look into routines specific to each year and subsequently compare them, to potentially reveal shifts in *location* and *sociality* of the activities.

5.3.1 Where: activities and locations

We begin our exploration of the variation in multimodal students' routines by examining the relationship between activities and their respective locations. Firstly, we observe that the optimal number of components of the NMF algorithm, determined using the cophenetic coefficient, differs between students in 2018 and 2020, with values of 7 and 5, respectively. This difference suggests that the extracted behaviour in 2018 exhibits greater diversity compared to 2020, which aligns with our previous findings regarding the entropy of students in the *location* behavioural matrix, where the diversity of locations is lower in 2020 (Sect. 5.1). In 2020, all *activity* routines predominantly occur either at *Home* or at *Relatives' Homes* except Component 6, which includes days at the *University* and Component 3, which does not retain much information about the location. Conversely, in 2018, we observe that *activity* routines, in addition to the *Home* or at *Relatives' Homes*, took place at *University*, *Friends' Home*, and even *Workplace* (Components 3 and 6).

The differentiation in students' behaviour can be observed in Fig. 4 (the complete components can be found in the Additional file 1 Sect. S5). When examining the activities, we find that the main routines extracted by NMF for both groups of students primarily revolve around activities such as *Lesson* and *Study*. In 2018 (Fig. 4(a)), these activities are closely associated with being at *Uiversity* (as seen in Components 3 and 6). When, instead, students are at *Home* (either alone or with their relatives), *Study* and *Lesson* are not as present, and their routine is more varied. Conversely, in 2020 students carry out the majority of their activities, including lessons and studying, at *Home*. It is interesting that in 2020, there is still a component associated with being at the *University*, namely Component 6. In fact, it was still possible for students to go to the library and follow lessons in-person (if they were of the first year). Moreover, we see that when students go to the university, their routine gets more complicated and they do more activities.

There were also other minor differences, such as having *Social Life* being a unique activity in 2018, while *Housework* is more distinctive among students in 2020, further highlighting the impact of the pandemic (see Figs. S9–S10 in the Additional file 1 Sect. S5). These results indicate that while the main activities remained consistent between the two groups of students, the pandemic restriction forced these activities to be done in different locations.

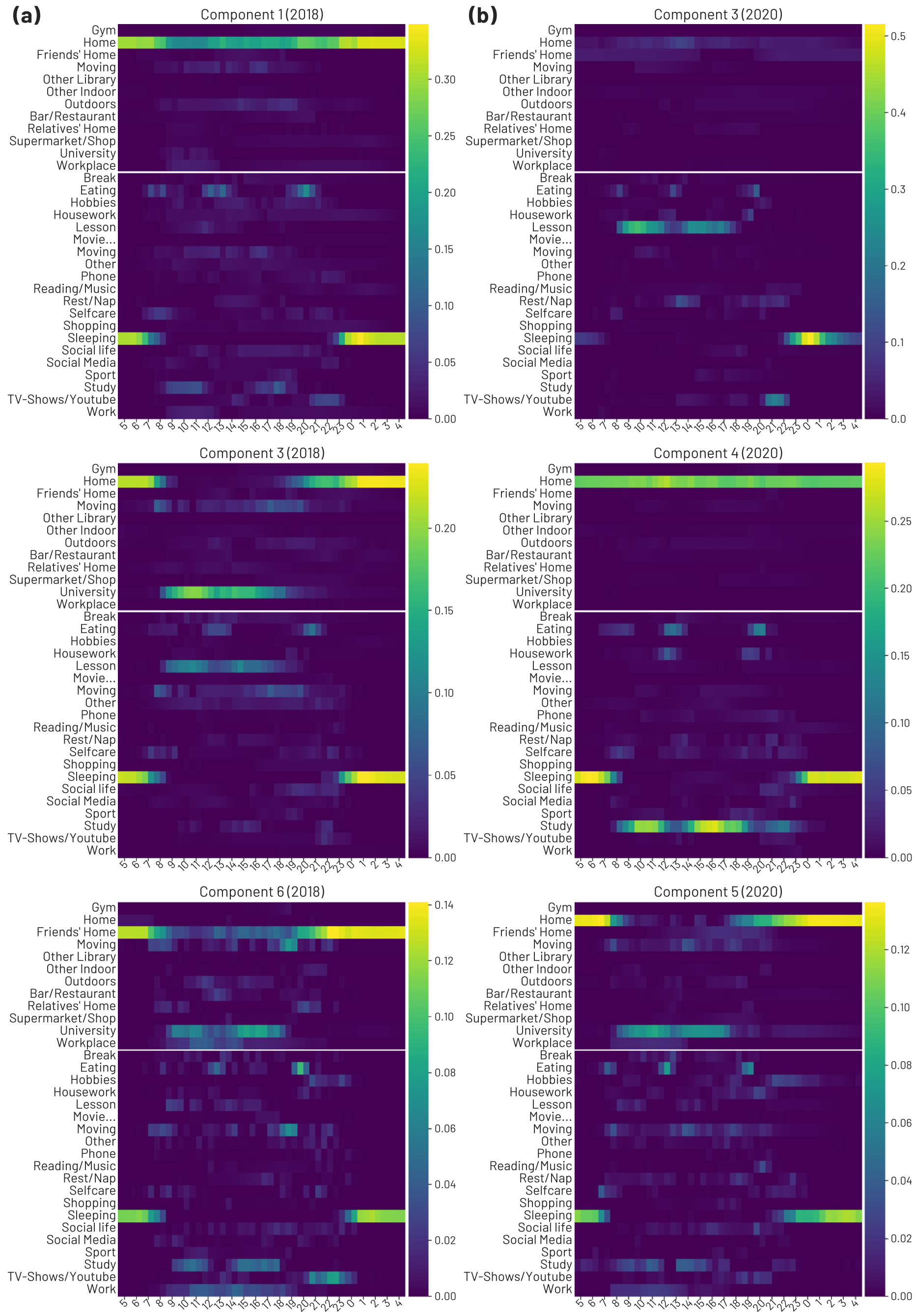

Figure 4 Location and Activity Joint Components. This figure shows the joint *location* and *activity* (separated by the white line) components that most differentiate students behaviour before and during the pandemic. **(a)** includes routines from 2018, while **(b)** includes routines from 2020. We can see the shift from conducting activities such as studying and having lesson from the university to private homes

5.3.2 *With whom*: *activities and sociality*

Turning our attention to the social component of activities, we observe that in 2020, the majority of the extracted components primarily involve activities performed *Alone*, see Fig. 5(b) (the complete components can be found in the Additional file 1 Sect. S5). Alternatively, these components involve a small number of individuals who likely reside in

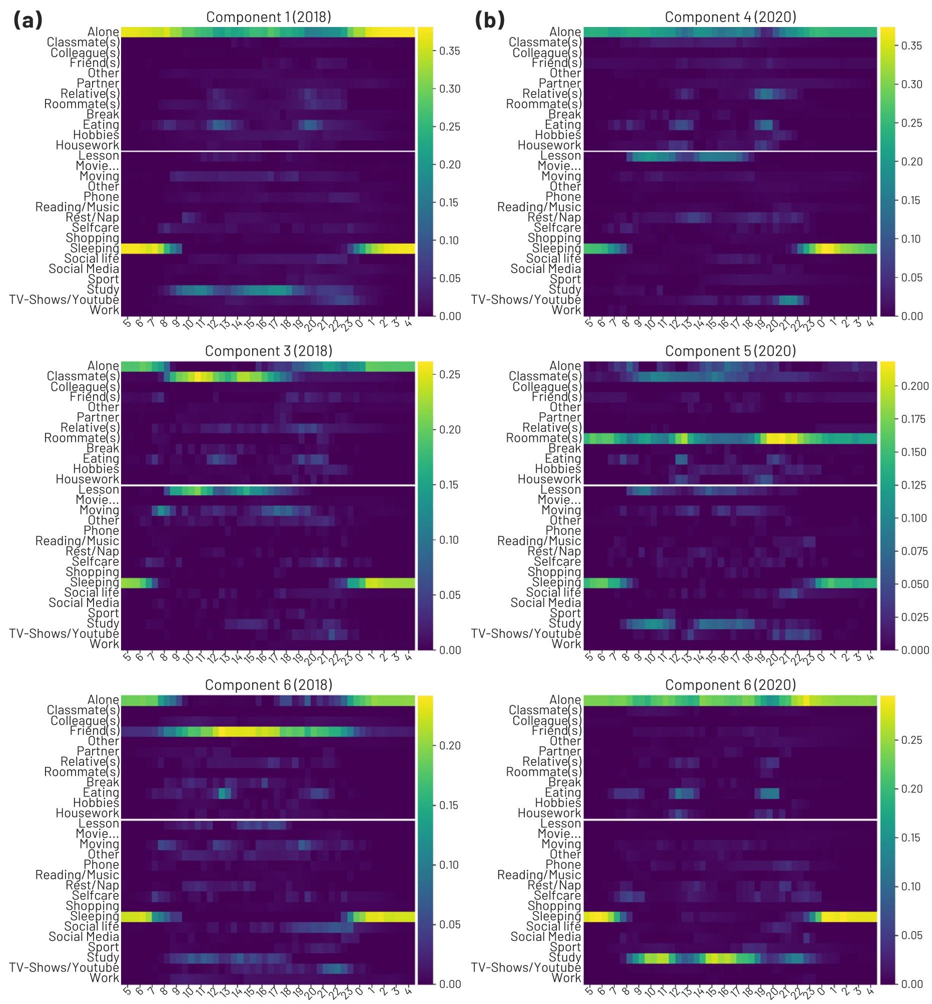

Figure 5 Sociality and Activity Joint Components. This figure shows the joint *sociality* and *activity* components that most differentiate students behaviour before and during the pandemic. **(a)** includes routines from 2018, while **(b)** includes routines from 2020. The main activities are still *Lesson* and *Study*, which are social activities during 2018, while in 2020 students do them mostly alone (or with people in the same house)

the same household, such as *Relatives*, *Partners*, or *Roommates*. This finding is consistent with the COVID-19 restrictions that limited the variety of social interactions during that period.

In contrast, as shown in Fig. 5(a), the social component of activities in 2018 exhibits a greater variation, which involved also interactions with *Friends* and *Classmates* (e.g., Components 3 and 6), while *Study* is the only activity strongly associated with being *Alone* (Component 1).

Although the NMF algorithm extracted six dominant components for both years, in 2018, there is a greater diversity of activities compared to students in 2020, where the weights in the heatmaps are more concentrated on specific activities. The activity patterns observed in the resulting routines are similar to those detected when considering the combined *locations* and *activities* matrices, which again involve as main activities studying and having lessons.

Notably, in 2018, there is a presence of mobility activities (e.g., *Moving*), which are almost absent in 2020. Additionally, in Component 5 of 2020 (Fig. 5(b)), we can observe the presence of *Social Life* in the evening, primarily involving interactions with *Roommates*. In contrast, in 2018, without government restrictions, social activities are more consistently carried out in the company of *Friends*.

6 Discussion and conclusions

In this work, we have analyzed the changes in the behavioural routines of students of the University of Trento due to the COVID-19 pandemic, focusing on three dimensions of behaviour: their *activity*, their *location* and their *sociality*. Few studies have considered students' behaviour, and generally, they tend to examine just one facet of behaviour [51]. Moreover, existing studies are often limited to the analysis of data that comes from either surveys, where data is collected manually, or mobile phones, where data is generally passively collected when triggered by specific events (e.g., phone calls, app usage, etc.). By leveraging an Experience Sampling Method (ESM) approach on smartphones, we collected self-reported information on the daily activities of university students by means of Time Diaries, which allowed for high-resolution and multimodal experience sampling data collection in a real-world setting, with a high level of control over participants' responses.

Our investigation first focused on understanding how specific behaviours evolved from 2018 to 2020. We accomplished this by conducting several statistical tests, including the Kolmogorov–Smirnov test, to compare the frequency distributions of each behaviour in the three dimensions, enabling us to pinpoint behaviours that exhibited significant changes. Notably, among the most prevalent activities, we observed a decrease in engagement with activities like *Moving* and *Social Life*, coupled with an increase in sedentary activities such as watching shows and videos. Interestingly, main activities like *Eating*, *Lesson* and *Study* did not show significant changes. This shift is mirrored in the *location* dimension, where the majority of places experienced a reduction in frequency, while the *Home* location registered a substantial increase. Furthermore, the trend towards more solitary activities and locations is reaffirmed by the overall decrease in social interactions, even with people residing in the same household, with students spending more time alone. These changes can most likely be attributed to the impact of the restrictions in place during the 2020 data collection.

These findings served as a basis for a more comprehensive analysis of our dataset, which extended to the study of daily behavioural patterns. We used Non-Negative Matrix Factorization (NMF) to incorporate the temporal dimension of daily routines, which is neglected in the analysis of activity frequencies. Thanks to NMF interpretability and clustering properties, this technique is particularly useful in identifying repeating patterns in the daily behaviours of people, both at an individual and a collective level. It has been successfully used to study behavioural routines in different contexts such as urban mobility [70], people's chronotypes [3] and interactions [32]. Our findings suggest that despite the COVID-19 restrictions, which have significantly altered various aspects of our lives [1, 48, 53, 65, 68], there is surprisingly little change in the main routine activities performed by students between 2018 (pre-pandemic) and 2020 (during the pandemic). Additionally, as shown in our analysis, the diversity of activities did not change significantly.

However, leveraging the richness of the data at our disposal, which enables a comprehensive multimodal analysis of students' behaviour, we discovered that their adaptation to the pandemic circumstances primarily occurred in the *location* and *sociality* dimensions. We observed differences in these dimensions for the dominant activities that constitute typical student routines, such as studying and attending classes. For these routines, our results indicate that, while the main activities remained consistent between the two groups of students, the pandemic restriction forced these activities to be done in different locations, and, thus, with different people. These results, indeed, are in line with previous works on the impact of COVID-19 on students' lives [51], and they are also reflected in the general population. Previous studies have shown that, during the pandemic, there is a reduction in the number of places visited and the number of social encounters [48, 56].

More in detail, in 2018, students' activities such as studying and having lessons were often located at the university. In contrast, in 2020, due to the COVID-19 pandemic, students predominantly engaged in activities, including studying and lessons, from home, although there was the possibility for them to follow lessons at the university. Notably, when students went to university, their routines became more complex and involved a broader range of activities. From a sociality point of view, in 2020, the majority of social activities were performed alone or with a small number of individuals from the same household, reflecting the impact of COVID-19 restrictions on limiting social interactions. In contrast, the *sociality* components of activities in 2018 exhibited greater variation, including interactions with friends and classmates, indicating a broader range of social engagements before the pandemic. This could be concerning, given that the lack and the reduction of social interactions, along with more time spent alone, has a negative effect on students' mental health, stress levels, sleeping patterns and general well-being [39, 51, 52, 69].

While our results stem from data collected from two living lab initiatives, in which the sample of students is different for 2018 and 2020, we believe that the outcomes are still reliable. In fact, this limitation is mitigated by the fact that the samples are from similar populations (students), and the data is collected in similar settings: they attend the same university and faculty in the same city.

Concluding, in applying Non-Negative Matrix Factorization, we have shown its usefulness in modelling students' behaviour and in extracting meaningful patterns, i.e., routines. We believe that our contribution can be valuable for further understanding the impact of the COVID-19 pandemic on students. Moreover, we show that using a multimodal approach can be useful in identifying patterns of joint behavioural dimensions, and, thus, be used to understand the correlations among these dimensions. Finally, this methodology can also help correlate the behavioural dimensions with other facets of human activity, enriching the literature on well-being [51], stress levels [69], and mood [50] of students who were significantly affected by the pandemic.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1140/epjds/s13688-023-00429-y.

Additional file 1. Additional file contains Supplementary Information 1—Data Mapping, Supplementary Information 2—Statistical Behavioural Differences, Supplementary Information 3—Alternative Routine Extraction Methods, Supplementary Information 4—Location and Sociality Routines, Supplementary Information 5—Joint Activities, Locations and Sociality. It includes Supplementary Tables S1–S5 and Supplementary Figs. S1 to S12. (PDF 11.4 MB)

Acknowledgements
Not applicable.

Funding
BL and SC are partly supported by the PNRR ICSC National Research Centre for High Performance Computing, Big Data and Quantum Computing (CN00000013), under the NRRP MUR program funded by the NextGenerationEU. BL and SC received also support from the project "AIAAA—Applicazioni di Intelligenza Artificiale e Analisi Avanzata", funded by the Autonomous Province of Trento.
The research by IB and FG has received funding from the European Union's Horizon 2020 FET Proactive project "WeNet—The Internet of us", grant agreement No 823783.

List of Abbreviations
ESM, Experience Sampling Method; NMF, Non-Negative Matrix Factorization; KS, Kolmogorov–Smirnov; PCA, Principal Component Analysis; RPCA, Robust Principal Component Analysis; IG, Information Gain.

Availability of data and materials
The datasets generated and analysed during the current study are not publicly available due to GDPR regulation and pending full approval from UniTN DPO but may be available from Fausto Giunchiglia (fausto.giunchiglia@unitn.it).

Declarations

Competing interests
The authors declare that they have no competing interests.

Author contributions
NAG, SC and BL designed the research. IB and FG designed the data collection experiments and provided the data after cleaning and preparation. NAG processed the data and performed the analysis. All the authors contributed to the writing of the article. All authors read and approved the final manuscript.

Author details
[1]Fondazione Bruno Kessler (FBK), Via Sommarive 18, 38123, Trento, Italy. [2]Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 9, 38123, Trento, Italy. [3]Department of Sociology and Social Research, University of Trento, via Verdi 26, 38122, Trento, Italy.

Received: 13 June 2023 Accepted: 20 November 2023 Published online: 05 December 2023

References

1. Adams-Prassl A, Boneva T, Golin M, Rauh C (2020) Inequality in the impact of the coronavirus shock: evidence from real time surveys. J Public Econ 189:104245. https://doi.org/10.1016/j.jpubeco.2020.104245
2. Aharony N, Pan W, Ip C, Khayal I, Pentland A (2011) Social fMRI: investigating and shaping social mechanisms in the real world. Pervasive Mob Comput 7(6):643–659. https://doi.org/10.1016/j.pmcj.2011.09.004
3. Aledavood T, Kivimäki I, Lehmann S, Saramäki J (2022) Quantifying daily rhythms with non-negative matrix factorization applied to mobile phone data. Sci Rep 12(1):5544. https://doi.org/10.1038/s41598-022-09273-y
4. Aleta A, Martin-Corral D, Pastore y Piontti A, Ajelli M, Litvinova M, Chinazzi M, Dean NE, Halloran ME, Longini IM Jr, Merler S et al (2020) Modelling the impact of testing, contact tracing and household quarantine on second waves of Covid-19. Nat Hum Behav 4(9):964–971. https://doi.org/10.1038/s41562-020-0931-9
5. Assi K, Meegahapola L, Droz W, Kun P, De Götzen A, Bidoglia M, Stares S, Gaskell G, Chagnaa A, Ganbold A et al (2023) Complex daily activities, country-level diversity, and smartphone sensing: a study in Denmark, Italy, Mongolia, Paraguay, and UK. In: Proceedings of the 2023 CHI conference on human factors in computing systems, pp 1–23. https://doi.org/10.1145/3544548.3581190
6. Baek C, McCrory PB, Messer T, Mui P (2021) Unemployment effects of stay-at-home orders: evidence from high-frequency claims data. Rev Econ Stat 103(5):979–993. https://doi.org/10.1162/rest_a_00996
7. Betthäuser BA, Bach-Mortensen AM, Engzell P (2023) A systematic review and meta-analysis of the evidence on learning during the Covid-19 pandemic. Nat Hum Behav 7(3):375–385. https://doi.org/10.1038/s41562-022-01506-4
8. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3:993–1022. https://doi.org/10.5555/944919.944937
9. Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, Schmidt AL, Valensise CM, Scala A, Quattrociocchi W et al (2020) Economic and social consequences of human mobility restrictions under Covid-19. Proc Natl Acad Sci 117(27):15530–15535. https://doi.org/10.1073/pnas.2007658117
10. Bonaccorsi G, Pierri F, Scotti F, Flori A, Manaresi F, Ceri S, Pammolli F (2021) Socioeconomic differences and persistent segregation of Italian territories during Covid-19 pandemic. Sci Rep 11(1):21174. https://doi.org/10.1038/s41598-021-99548-7
11. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, Stephenson AB, Leech G, Altman G, Mikulik V et al (2021) Inferring the effectiveness of government interventions against Covid-19. Science 371(6531):9338. https://doi.org/10.1126/science.abd9338
12. Brdiczka O, Su NM, Begole JB (2010) Temporal task footprinting: identifying routine tasks by their temporal patterns. In: Proceedings of the 15th international conference on intelligent user interfaces. IUI '10. Association for Computing Machinery, pp 281–284. https://doi.org/10.1145/1719970.1720011
13. Brodeur A, Clark AE, Fleche S, Powdthavee N (2021) Covid-19, lockdowns and well-being: evidence from Google trends. J Public Econ 193:104346. https://doi.org/10.1016/j.jpubeco.2020.104346
14. Brown E, Shortell T (2014) Walking in the European city: quotidian mobility and urban ethnography. Ashgate Publishing, Ltd, Farnham. https://doi.org/10.4324/9781315547770
15. Brunet J-P, Tamayo P, Golub TR, Mesirov JP (2004) Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci 101(12):4164–4169. https://doi.org/10.1073/pnas.0308531101

16. Candès EJ, Li X, Ma Y, Wright J (2011) Robust principal component analysis? J ACM 58(3):1–37. https://doi.org/10.1145/1970392.1970395
17. Cazabet R, Jensen P, Borgnat P (2018) Tracking the evolution of temporal patterns of usage in bicycle-sharing systems using nonnegative matrix factorization on multiple sliding windows. Int J Urban Sci 22(2):147–161. https://doi.org/10.1080/12265934.2017.1336468
18. Centellegher S, De Nadai M, Caraviello M, Leonardi C, Vescovi M, Ramadian Y, Oliver N, Pianesi F, Pentland A, Antonelli F et al (2016) The mobile territorial lab: a multilayered and dynamic view on parents' daily lives. EPJ Data Sci 5:1. https://doi.org/10.1140/epjds/s13688-016-0064-6
19. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, Pastore y Piontti A, Mu K, Rossi L, Sun K et al (2020) The effect of travel restrictions on the spread of the 2019 novel coronavirus (Covid-19) outbreak. Science 368(6489):395–400. https://doi.org/10.1126/science.aba9757
20. Cichocki A, Phan A-H (2009) Fast local algorithms for large scale nonnegative matrix and tensor factorizations. IEICE Trans Fundam Electron Commun Comput Sci 92(3):708–721. https://doi.org/10.1587/transfun.E92.A.708
21. Covid Restrictions for the Trento Province in the Sampling Period (in Italian) (2023) Last Accessed May 2nd, 2023. https://www.ufficiostampa.provincia.tn.it/content/download/174818/3026422/file/ORDINANZA_N_54_prot_726529_del_14_novembre_2020.pdf
22. Csikszentmihalyi M, Larson R, Csikszentmihalyi M (2014) The experience sampling method. In: Flow and the foundations of positive psychology: the collected works of Mihaly Csikszentmihalyi, 21–34. https://doi.org/10.1007/978-94-017-9088-8_2
23. Delussu F, Tizzoni M, Gauvin L (2022) Evidence of pandemic fatigue associated with stricter tiered Covid-19 restrictions. PLOS Digit Health 1(5):0000035. https://doi.org/10.1371/journal.pdig.0000035
24. Dhifaoui Z, Khalfaoui R, Abedin MZ, Shi B (2022) Quantifying information transfer among clean energy, carbon, oil, and precious metals: a novel transfer entropy-based approach. Finance Res Lett 49:103138. https://doi.org/10.1016/j.frl.2022.103138
25. Dhifaoui Z, Khalfaoui R, Jabeur SB, Abedin MZ (2023) Exploring the effect of climate risk on agricultural and food stock prices: fresh evidence from EMD-based variable-lag transfer entropy analysis. J Environ Manag 326:116789. https://doi.org/10.1016/j.jenvman.2022.116789
26. Di Clemente R, Luengo-Oroz M, Travizano M, Xu S, Vaitla B, González MC (2018) Sequences of purchases in credit card data reveal lifestyles in urban populations. Nat Commun 9(1):1–8. https://doi.org/10.1038/s41467-018-05690-8
27. Eagle N, Pentland AS (2006) Reality mining: sensing complex social systems. Pers Ubiquitous Comput 10(4):255–268. https://doi.org/10.1007/s00779-005-0046-3
28. Eagle N, Pentland AS (2009) Eigenbehaviors: identifying structure in routine. Behav Ecol Sociobiol 63:1057–1066. https://doi.org/10.1007/s00265-009-0739-0
29. Farrahi K, Gatica-Perez D (2010) Probabilistic mining of socio-geographic routines from mobile phone data. IEEE J Sel Top Signal Process 4(4):746–755. https://doi.org/10.1109/JSTSP.2010.2049513
30. Fatima I, Fahim M, Lee Y-K, Lee S (2013) A unified framework for activity recognition-based behavior analysis and action prediction in smart homes. Sensors 13(2):2682–2699. https://doi.org/10.3390/s130202682
31. Galeazzi A, Cinelli M, Bonaccorsi G, Pierri F, Schmidt AL, Scala A, Pammolli F, Quattrociocchi W (2021) Human mobility in response to Covid-19 in France, Italy and UK. Sci Rep 11(1):13141. https://doi.org/10.1038/s41598-021-92399-2
32. Gauvin L, Panisson A, Cattuto C (2014) Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach. PLoS ONE 9(1):86028. https://doi.org/10.1371/journal.pone.0086028
33. Giunchiglia F, Bignotti E, Zeni M (2017) Personal context modelling and annotation. In: 2017 IEEE international conference on pervasive computing and communications workshops (PerCom workshops), pp 117–122. https://doi.org/10.1109/PERCOMW.2017.7917543
34. Giunchiglia F, Bison I, Busso M et al (2021) A worldwide diversity pilot on daily routines and social practices (2020-2021). University of Trento Technical Report (#DISI-2001-DS-01). https://doi.org/10.13140/RG.2.2.35261.54244
35. Giunchiglia F, Zeni M, Gobbi E, Bignotti E, Bison I (2017) Mobile social media and academic performance. In: Social informatics: 9th international conference, SocInfo 2017, proceedings, part II 9, Oxford, UK, September 13–15, 2017. Springer, pp 3–13. https://doi.org/10.1007/978-3-319-67256-4_1
36. Giunchiglia F, Zeni M, Gobbi E, Bignotti E, Bison I (2018) Mobile social media usage and academic performance. Comput Hum Behav 82:177–185. https://doi.org/10.1016/j.chb.2017.12.041
37. Goffman E (1959) The presentation of self in everyday life. Doubleday, New York
38. Gonzalez T, De La Rubia MA, Hincz KP, Comas-Lopez M, Subirats L, Fort S, Sacha GM (2020) Influence of Covid-19 confinement on students' performance in higher education. PLoS ONE 15(10):0239490. https://doi.org/10.1371/journal.pone.0239490
39. Guo Y-F, Liao M-Q, Cai W-L, Yu X-X, Li S-N, Ke X-Y, Tan S-X, Luo Z-Y, Cui Y-F, Wang Q et al (2021) Physical activity, screen exposure and sleep among students during the pandemic of Covid-19. Sci Rep 11(1):8529. https://doi.org/10.1038/s41598-021-88071-4
40. Hamon R, Borgnat P, Flandrin P, Robardet C (2014) Nonnegative matrix factorization to find features in temporal networks. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, New York, pp 1065–1069. https://doi.org/10.1109/ICASSP.2014.6853760
41. Hou X, Gao S, Li Q, Kang Y, Chen N, Chen K, Rao J, Ellenberg JS, Patz JA (2021) Intracounty modeling of Covid-19 infection with human mobility: assessing spatial heterogeneity with business traffic, age, and race. Proc Natl Acad Sci 118(24):2020524118. https://doi.org/10.1073/pnas.2020524118
42. Huynh T, Fritz M, Schiele B (2008) Discovery of activity patterns using topic models. In: Proceedings of the 10th international conference on ubiquitous computing, pp 10–19. https://doi.org/10.1145/1409635.1409638
43. Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans R Soc Lond A, Math Phys Eng Sci 374(2065):20150202. https://doi.org/10.1098/rsta.2015.0202
44. Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT (2010) A survey of mobile phone sensing. IEEE Commun Mag 48(9):140–150. https://doi.org/10.1109/MCOM.2010.5560598
45. Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M et al (2009) Computational social science. Science 323(5915):721–723. https://doi.org/10.1126/science.1167742

46. Lazer DMJ, Pentland A, Watts DJ, Aral S, Athey S, Contractor N, Freelon D, Gonzalez-Bailon S, King G, Margetts H, Nelson A, Salganik MJ, Strohmaier M, Vespignani A, Wagner C (2020) Computational social science: obstacles and opportunities. Science 369(6507):1060–1062. https://doi.org/10.1126/science.aaz8170
47. Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401(6755):788–791. https://doi.org/10.1038/44565
48. Lucchini L, Centellegher S, Pappalardo L, Gallotti R, Privitera F, Lepri B, De Nadai M (2021) Living in a pandemic: changes in mobility routines, social activity and adherence to Covid-19 protective measures. Sci Rep 11(1):24452. https://doi.org/10.1038/s41598-021-04139-1
49. Madan A, Cebrian M, Moturu S, Farrahi K et al (2011) Sensing the "health state" of a community. IEEE Pervasive Comput 11(4):36–45. https://doi.org/10.1109/MPRV.2011.79
50. Meegahapola L, Droz W, Kun P, de Götzen A, Nutakki C, Diwakar S, Correa SR, Song D, Xu H, Bidoglia M et al (2023) Generalization and personalization of mobile sensing-based mood inference models: an analysis of college students in eight countries. In: Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 6(4), pp 1–32. https://doi.org/10.1145/3569483
51. Nepal S, Wang W, Vojdanovski V, Huckins JF, daSilva A, Meyer M, Campbell A (2022) COVID student study: a year in the life of college students during the Covid-19 pandemic through the lens of mobile phone sensing. In: Proceedings of the 2022 CHI conference on human factors in computing systems, pp 1–19. https://doi.org/10.1145/3491102.3502043
52. Ochnik D, Rogowska AM, Kuśnierz C, Jakubiak M, Schütz A, Held MJ, Arzenšek A, Benatov J, Berger R, Korchagina EV et al (2021) Mental health prevalence and predictors among university students in nine countries during the Covid-19 pandemic: a cross-national study. Sci Rep 11(1):18644. https://doi.org/10.1038/s41598-021-97697-3
53. Oved S, Mofaz M, Lan A, Einat H, Kronfeld-Schor N, Yamin D, Shmueli E (2021) Differential effects of Covid-19 lockdowns on well-being: interaction between age, gender and chronotype. J R Soc Interface 18(179):20210078. https://doi.org/10.1098/rsif.2021.0078
54. Pauca VP, Shahnaz F, Berry MW, Plemmons RJ (2004) Text mining using non-negative matrix factorizations. In: Proceedings of the 2004 SIAM International Conference on Data Mining. SIAM, Philadelphia, pp 452–456. https://doi.org/10.1137/1.9781611972740.45
55. Pejovic V, Lathia N, Mascolo C, Musolesi M (2016) Mobile-based experience sampling for behaviour research. In: Emotions and personality in personalized services: models, evaluation and applications. Springer, Berlin, pp 141–161. https://doi.org/10.1007/978-3-319-31413-6_8
56. Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C, Tizzoni M (2020) Covid-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown. Sci Data 7(1):230. https://doi.org/10.1038/s41597-020-00575-2
57. Perra N (2021) Non-pharmaceutical interventions during the Covid-19 pandemic: a review. Phys Rep 913:1–52. https://doi.org/10.1016/j.physrep.2021.02.001
58. Persson J, Parie JF, Feuerriegel S (2021) Monitoring the Covid-19 epidemic with nationwide telecommunication data. Proc Natl Acad Sci 118(26):2100664118. https://doi.org/10.1073/pnas.2100664118
59. Pierson E, Althoff T, Leskovec J (2018) Modeling individual cyclic variation in human behavior. In: Proceedings of the 2018 world wide web conference, pp 107–116. https://doi.org/10.1145/3178876.3186052
60. Raento M, Oulasvirta A, Eagle N (2009) Smartphones: an emerging tool for social scientists. Sociol Methods Res 37(3):426–454. https://doi.org/10.1177/0049124108330005
61. Sra S, Dhillon I (2005) Generalized nonnegative matrix approximations with Bregman divergences. In: Weiss Y, Schölkopf B, Platt J (eds) Advances in neural information processing systems, vol 18. MIT Press, Cambridge
62. Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen MM, Larsen JE, Lehmann S (2014) Measuring large-scale social networks with high resolution. PLoS ONE 9(4):95978. https://doi.org/10.1371/journal.pone.0095978
63. Vagni G, Cornwell B (2018) Patterns of everyday activities across social contexts. Proc Natl Acad Sci 115(24):6183–6188. https://doi.org/10.1073/pnas.1718020115
64. Van Berkel N, Ferreira D, Kostakos V (2017) The experience sampling method on mobile devices. ACM Comput Surv 50(6):1–40. https://doi.org/10.1145/3123988
65. Van Lancker W, Parolin Z (2020) Covid-19, school closures, and child poverty: a social crisis in the making. Lancet Public Health 5(5):243–244. https://doi.org/10.1016/S2468-2667(20)30084-0
66. Xu X, Chikersal P, Doryab A, Villalba DK, Dutcher JM, Tumminia MJ, Althoff T, Cohen S, Creswell KG, Creswell JD et al (2019) Leveraging routine behavior and contextually-filtered features for depression detection among college students. Proc ACM Interact Mob Wearable Ubiquitous Technol 3(3):1–33. https://doi.org/10.1145/3351274
67. Xu Y, Bull D, Damen D (2017) Unsupervised long-term routine modelling using dynamic Bayesian networks. In: 2017 international conference on digital image computing: techniques and applications (DICTA), pp 1–8. https://doi.org/10.1109/DICTA.2017.8227502
68. Yabe T, Bueno BGB, Dong X, Pentland A, Moro E (2023) Behavioral changes during the Covid-19 pandemic decreased income diversity of urban encounters. Nat Commun 14(1):2310. https://doi.org/10.1038/s41467-023-37913-y
69. Yang C, Chen A, Chen Y (2021) College students' stress and health in the Covid-19 pandemic: the role of academic workload, separation from school, and fears of contagion. PLoS ONE 16(2):0246676. https://doi.org/10.1371/journal.pone.0246676
70. Yang Y, Pentland A, Moro E (2023) Identifying latent activity behaviors and lifestyles using mobility data to describe urban dynamics. EPJ Data Sci 12(1):1. https://doi.org/10.1140/epjds/s13688-023-00390-w
71. Yao L, Mimno D, McCallum A (2009) Efficient methods for topic model inference on streaming document collections. In: Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining, pp 937–946. https://doi.org/10.1145/1557019.1557121
72. Zeni M, Zaihrayeu I, Giunchiglia F (2014) Multi-device activity logging. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: adjunct publication, pp 299–302. https://doi.org/10.1145/2638728.2638756
73. Zhang W, Shen Q, Teso S, Lepri B, Passerini A, Bison I, Giunchiglia F (2021) Putting human behavior predictability in context. EPJ Data Sci 10(1):42. https://doi.org/10.1140/epjds/s13688-021-00299-2

Publisher's Note