Computers in Human Behavior 150 (2024) 107991

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

# Group-level analysis of engagement poorly reflects individual students' processes: Why we need idiographic learning analytics

Mohammed Saqr

*University of Eastern Finland, School of Computing, P.O. Box 111, Joensuu campus, Joensuu, Yliopistokatu 2, fi-80100, Joensuu, Finland*

ARTICLE INFO

Handling editor: Nicolae Nistor

*Keywords:*
Learning analytics
Idiographic engagement
Within-person
Psychological networks
Precision education

ABSTRACT

A central assumption of the scientific method is that inferences derived from group-level analysis align with and generalize to the individual level. This study was conducted to put this assumption to the test to examine if and to what extent our analysis, inferences, and assumptions hold, and which variables generalize from the group to the individual level. We use engagement as the underpinning of this study. However, the same methods and questions apply elsewhere. The study included 238 students over six courses and applied the latest advances in psychological networks. Two networks were estimated using the same data: a between-person model that captures the group-level engagement and a within-person model that captures the within-person processes. The results showed that there were significant differences between the two networks and a lack of generalizability regarding regularity, academic achievement, and online disengagement. Such findings cast doubts on inferences drawn from group-level data about our understanding of learners' performance or engagement, or to design personalized interventions. More attention and efforts are needed to further model within-person processes to understand, and possibly deliver precise personalized support and interventions that are more generalizable and truly personalized.

## 1. Introduction

Despite decades of research, practical experience, and wealth of insights, educational institutions are struggling to understand students' success, improve retention rates, or offer personalized support (Romero & Ventura, 2020; Shafiq, Marjani, Habeeb, & Asirvatham, 2022). A wealth of research on effective intervention regarding academic achievement, behavior, and attitude already exists (Scammacca, Roberts, Vaughn, & Stuebing, 2015; Taylor, Oberle, Durlak, & Weissberg, 2017). However, real-life implementation of such interventions has fallen short of promise (Cook, Kilgus, & Burns, 2018; Scammacca et al., 2015). Reasons for the disappointing intervention relate to how intervention research is conducted, and the way intervention is applied. Research is always conducted by collecting cross-sectional group-level data from a sample of students to generate an aggregate average or derive "standard laws". The central assumption is that the average (and the variance thereof) represent the individual students (and their variance). However, the "aggregate average" rarely transmutes to individual students or even to the majority of students (Richters, 2021). As Winne elaborated in great detail in a seminal paper, the average of a group poorly describes any particular individual student (Winne, 2017). As such, a more holistic approach based on a nuanced understanding of the educational and psychological theories is therefore needed. The said approach needs to generate adaptive insights based on true person-specific processes that can be generalizable and trustworthy (Sailer, Ninaus, Huber, Bauer, & Greiff, 2023).

In contrast to group-level methods, person-specific (idiographic) methods rely on within-person variance to analyze the individual processes (where the process happens) to produce insights that are more aligned with the individual processes, as the name implies. This is achieved through collecting several data points over time from an individual subject that suffice for a robust statistical analysis of the individual (Beltz, Wright, Sprague, & Molenaar, 2016; Molenaar & Campbell, 2009; Saqr, 2023). Given that psychological phenomena and, in particular, engagement —the focus of this study— unfold at the individual level, person-specific methods are therefore more appropriate for capturing the individual variability and peculiarities (Fisher, Medaglia, & Jeronimus, 2018; Richters, 2021; Saqr, 2023). Relying on person-specific methods could enable a paradigm shift in research, which has been described as a "brink of a major reorientation" that is "no longer an option, but a necessity" (Molenaar & Campbell, 2009). Yet, educational research has so far not sufficiently examined to what

*E-mail address:* mohammed.saqr@uef.fi.

https://doi.org/10.1016/j.chb.2023.107991
Received 19 August 2023; Received in revised form 2 October 2023; Accepted 8 October 2023
Available online 16 October 2023
0747-5632/

extent our traditional methods —group-level analysis— align with and generalize to the individual level.

Therefore, the main aim of this study is to examine how group-level analysis compares, transfers, or represents individual-level analysis of engagement as a process. In other words, we aim to investigate how far inferences generated from the group-level analysis generalize to a person-specific level and can therefore be used as a basis for intervention, personalization or understanding of the individual student engagement?

The next section discusses the conceptual foundations: what group-level analysis is, how it is performed, and where it falls short. Then, a discussion of individual-level analysis, the related concepts and how it captures the individual processes. Engagement is introduced with a reflection on the theoretical role of the "person". Then, a discussion on methods that can capture the individual engagement processes with a focus on modern network methods which are used in this paper. The section concludes with examples of the literature and the motivation of the paper.

## 2. Background

### 2.1. Theoretical underpinning

Online engagement reflects students' *energy in action,* investment in learning and participation in educational activities (Reschly & Christenson, 2022). Researchers widely agree that engagement is a multi-dimensional construct that encompasses a behavioral, an affective, and a cognitive dimension (Fredricks, Blumenfeld, & Paris, 2004; Reschly & Christenson, 2022; Sinatra, Heddy, & Lombardi, 2015). In online learning —the context of our study— engagement is commonly captured through the observable involvement in learning activities that reflect both behavioral and cognitive dimensions (Fredricks et al., 2004; Reschly & Christenson, 2022; Sinatra et al., 2015; Wang, Degol, & Henry, 2019). Engagement theoreticians have always recognized the centrality of *"self"* as a fundamental concept since the early days (Tinto, 1975) where the "person-role" was deemed fundamental in maintaining engagement of the individual students. To that date, the self (or the individual) occupies a central place in most engagement theories (Connell & Wellborn, 1991; Reschly & Christenson, 2012; Tinto, 2022) and remains at the core of almost every motivation framework (Skinner & Raine, 2022; Urdan & Kaplan, 2020; Wigfield & Eccles, 2020). Yet, studies examining engagement have used group-level analysis to reflect, account for, and theorize about the individual engagement processes. Given that engagement as a process unfolds at the individual-level, it is imperative to capture it at the level where it unfolds, i.e., at the individual-level, which is an objective this study aims to achieve.

### 2.2. Between and within-person

Group-level (*nomothetic*) research is commonly performed through collecting data from a group of individuals to calculate statistics that represent the "state of affairs" as a basis for generalizable laws or rules; see Fig. 1 for an illustration. In other words, group-level methods model the between-person average behavior based on a large sample (Fisher et al., 2018; Molenaar, 2004; Richters, 2021). For instance, a researcher would collect data from a school through a survey that measures engagement; analysis of the data would result in, e.g., conclusions that higher engagement predicts higher academic achievement. Based on this finding, the researcher assumes that there is an association between engagement and achievement; such association is a "rule" that should generalize to most students (Beck & Jackson, 2021). That is, we expect that all engaged students will be high achievers.

However, for group-level insights to be generalizable to individuals, the process under study, e.g., engagement, must be *ergodic* (Fisher et al., 2018). An ergodic process fulfills two conditions: homogeneity (lack of inter-individual variations) and stationarity (stability over-time) (Molenaar & Campbell, 2009). By assuming that a process is ergodic, we imply that all people have a stable average engagement process over time and have an identical mean to the cross-sectional mean of the population (Fisher et al., 2018). In other words, ergodicity implies that every person has a similar degree of within-person variance, which is in turn identical to the degree of between-person average (Fisher et al., 2018). These assumptions rarely hold in real life, and recent empirical research has repeatedly shown that psychological processes, including —engagement— are non-ergodic (Bakker, Sanz Vergel, & Kuntze, 2015; Dirk & Schmiedek, 2016; Fisher et al., 2018; Martin et al., 2015). In other words, learners' engagement shows significant heterogeneity between and within-person and varies by time (Jovanovic et al., 2019; Martin et al., 2015; Saqr, López-Pernas, Helske, & Hrastinski, 2023; Saqr & Lopez-Pernas, 2021). Therefore, person-specific methods are needed to account for the within-person variability of engagement (Richters, 2021).

Person-specific (idiographic) methods aim at modeling the within-person variance, i.e., the variations or fluctuations within an individual from their own average (Beck & Jackson, 2021; Hamaker, 2012); for instance, how much a student is engaged compared to his/her average. Fig. 1 shows the concepts of between-person and within-person statistics. In the example, data is sampled from a group of hypothetical students at time point 4 (Fig. 1A). Statistics are cross-sectional group-level averages of engagement level. In Fig. 1B data is sampled from each

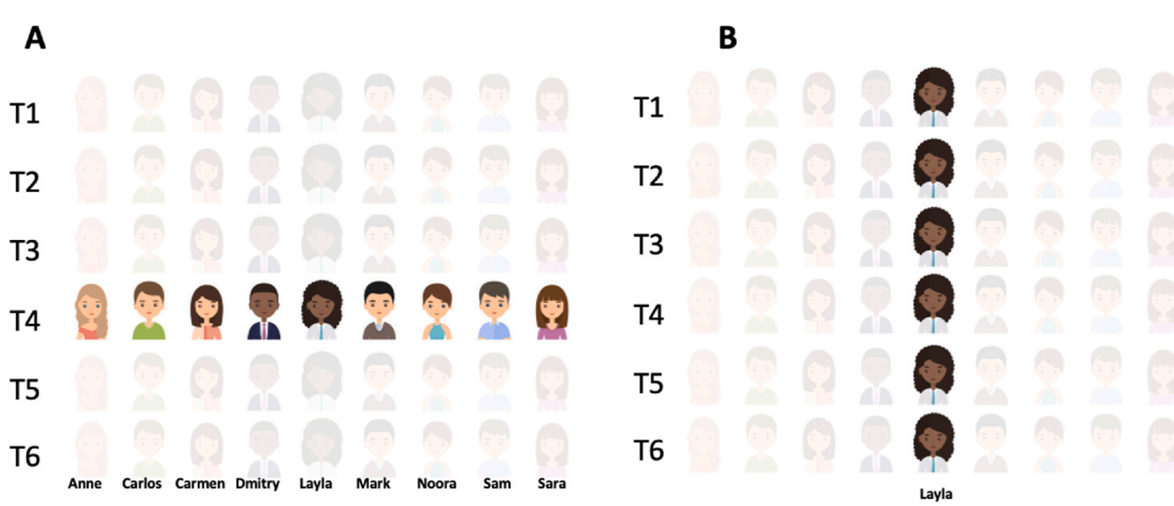


**Fig. 1.** **A**: cross section group-level data of nine students at T4. **B:** multiple data points collected for within-person analysis of *Layla.*

individual student across several time points (e.g., Layla) and so, representing the within-person variation of her engagement.

### 2.3. Modeling the within-person processes

A wide variety of methods have been developed to understand learning and learners' processes (Romero & Ventura, 2020). Regression models and in particular multilevel models (MLM) have been used to model within-person processes and have offered several important insights. An example of MLM is the work of Trautwein, Schnyder, Niggli, Neumann, and Lüdtke (2009) who studied the relationship between homework and performance; the findings of the study have shown the beneficial effects of homework on achievement at both levels of analysis (group and individual). Another example is the work of Martin et al. (2015) who studied real-time motivation and engagement and emphasized the importance of modeling the temporal variations of within-student engagement. Similarly, Bakker et al. (2015) studied the weekly variations of student engagement using a diary, and the authors reported a positive relationship between learning activities and grades. In the same vein, Webster and Hadwin (2015) studied emotional experiences using MLM, the studied variables included goal attainment, intensity of emotions, and emotion regulation. The authors reported that boredom was an emotional predictor of goal attainment on the group-level but not on the individual level. Dirk and Schmiedek (2016) used multilevel models to study memory variations across possible *bad* and *good* days; the authors reported significant variability in the daily working memory of the included children. Nevertheless, MLM are not true idiographic methods because the estimates in MLM are informed by and shrunk towards group averages (Beck & Jackson, 2021).

### 2.4. Psychological networks

Recently, probabilistic networks – the methods of this study– have emerged as a powerful framework for the modeling of complex phenomena. In contrast to the reductionist view of human processes, networks allow the modeling of learners' engagement as a complex, multidimensional, and multicausal process (Glaser, 1989; Reimann, 2019). Networks offer the advantage of modeling engagement as a collective system of interactions and allow us to understand the interdependence of components, their organization, and the interactions between them (Borsboom et al., 2021). In this study, we will take advantage of probabilistic networks and in particular, psychological networks to model the interplay between engagement indicators. In psychological networks, the *nodes* are variables, e.g., constructs, emotions, or behaviors. The relationships between the nodes or the *edges* are regularized partial correlations (Epskamp, Waldorp, Mõttus, & Borsboom, 2018). The networks are commonly undirected, signed (positive or negative) and weighted according to the magnitude of the partial correlation. What is more, several types of network models have been successfully used to capture within-person processes that include cross-sectional, dynamic, and idiographic models (Beck & Jackson, 2020; Saqr & López-Pernas, 2021).

While psychological network research in education is so far very rare, there are some emerging examples for modeling the within-person processes. The work of Saqr and López-Pernas (2021) is relevant to our study; the authors studied the interplay between engagement, motivation and self-regulation over 30 days using idiographic networks; the authors were able to identify gaps in the studied student's self-regulation (e.g., planning not connected to task execution). Similarly, idiographic networks have been used to capture the interplay of self-regulated learning on the individual-level using intensive longitudinal video data (Malmberg, Saqr, Järvenoja, Haataja, et al., 2022a; Malmberg, Saqr, Järvenoja, & Järvelä, 2022b). Other applications of psychological networks – not idiographic networks – include modeling the complex interactions of academic writing and self-regulation (Saqr, Viberg, & Peteers, 2021). To that end, what is not known or examined in education is to what extent group-level insights gained from group-level data can guide or inform us about an individual's engagement.

To illustrate the concept of psychological networks, a simple network based on a subset of experimental data is presented in Fig. 2, where we see a network of associations between the frequency of course browsing (FRQ), lecture reading (LEC), forum reading (RFM) and posting (CFM). We see there is a strong association between reading and posting forums (RFM and CFM) (shown in dark blue), this correlation is computed after controlling for the two other variables in the network. There was also a weak negative correlation between lecture views (LEC) and reading forums (RFM) after controlling for the other two variables.

### 2.5. Motivation and research questions

Given the importance of individuality in learner's engagement, and the need for personalized and generalizable learning interventions, the worth of studying the within-person process of engagement cannot be overstated (Sailer et al., 2023). Indeed, an understanding of the individual processes and how modern learning analytics can create adaptive insights that can work for every individual is an open challenge for learning analytics that need to be solved (Sailer et al., 2023). We need to investigate if and to what extent the common approach to analysis using between-person data captures, reflects, or provides enough information about the within-person engagement processes. We do so by taking advantage of the rigor of the psychological networks. RQ1 aims at mapping the complex relationships between engagement indicators (derived from the learning management system (LMS)) through studying the topology of the network, and the interplay between different indicators. Most importantly, we are interested in the extent to which these variables predict or explain the two relevant variables: academic achievement operationalized as a final grade or LMS inactive days (gap) which reflects disengagement. In RQ1a, we use the aggregated average of a group (between-person) while in RQ1b we use the within-person variance. RQ2 is related to how the insights gained from either level (RQ1a and RQ1b) are compared to each other. RQ3 aims to find what the network can tell us by finding the important variables that we may use as targets for intervention.

**RQ1**. **On the group-level (RQ1a) and on the (individual-level RQ1b)**: What is the topology of interplay between engagement indicators, and how do they explain the outcome (final grade) or the lack thereof (disengagement)?

**RQ2**. To what extent does group-level network compare-to, provide or reflect the within-person engagement?

**RQ3**. What are the variables that are most central, i.e., drive the positive and strong connections between other variables, and to what extent can a change in a variable change other variables i.e., intervention?

## 3. Methods

### 3.1. Context

The study involved students in a blended program that is based on problem-based learning (PBL) which follows the "seven jump" method (Wood, 2003). The typical PBL consists of a group of students and a teacher who facilitates the PBL session. The group meets physically at the beginning of each week to discuss a "problem" (Wood, 2003). The problem is a scenario of a clinical case designed to cover the learning objectives of the week. For example, if the learning objective of the week is to study blood clotting, the problem would be about, e.g., a case that has a bleeding disorder with other medical complaints. The lectures, practical sessions, and seminars follow the same theme. During the beginning of the week, students read the problem, identify terms, discuss possible solutions and co-construct the learning objectives.

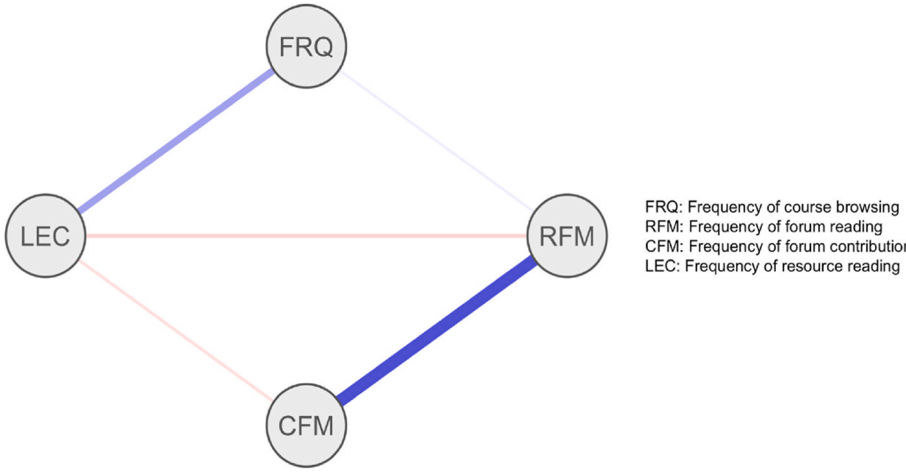


**Fig. 2.** An example of a psychological network of four variables. Blue edges are positive partial correlations and red edges are negative partial correlations.

Students post the learning objectives to the LMS and start the discussions online and continue online all over the week. On Friday, students meet face-to-face again and conclude what they have learned and reflect on the group performance. The PBL sessions represent the core of the program and therefore, engagement with online PBL is important or can be thought of as “required”. Moodle LMS is used to deliver all lectures, course announcements, updates, schedules, and course booklet. The grading scheme of the course is 80% exams, 20% mid-term, and on-going course assessment (not LMS).

The study involved the first six courses of the program which had the largest complete subset of data available at the time of data collection. The included courses were: 1) *Education* course, which teaches principles of PBL, assessment methods, and how students can approach learning tasks (practical, seminars, skills sessions); 2) *Principles of diseases* course, which teaches disease basics including physiology, pathology and biochemistry; 3) *Growth and development* course, which teaches anatomy, embryology, physiology and pathology of growth and development; 4) *Head and neck* course which teaches anatomy and pathology of the head, neck and face regions; 5) *Cell structure and function* course, which teaches the principles of histology, physiology and pathology at the cellular level, and 6) *Body systems* course, which teaches the physiology and pathology of major organs and body systems. All the courses follow the PBL, have an average six weeks of active studies, and follow the same assessment methods. Since courses have multiple subjects (e.g., physiology, histology, and pathology in every course), they are taught by several teachers, e.g., anatomy teachers teach the anatomy section, physiology teachers teach the physiology lectures and so on. The average number of teachers in each course was around 15 and, therefore, it is reasonable to say that courses are not influenced by any single teacher.

### 3.2. Collected data

The data included in this study were derived from the Moodle LMS logs. The logs were pre-processed so that teacher-related events, non-learning events (e.g., chats, profile and account related events were deleted). Some activities were rather rare and inconsistent among courses and therefore were excluded (e.g., Moodle workshop, choice, and polls). We consolidated granular logs so that closely related activities were combined, so activities related to contributing, editing, and writing a forum post were all categorized as “contributing to the forums” as detailed in the next section.

The choice of variables followed the literature on students’ indicators of online engagement (Henrie, Halverson, & Graham, 2015; Kassab, El-Sayed, & Hamdy, 2022; Saqr et al., 2023), context, and course design (Saqr, Jovanovic, Viberg, & Gašević, 2022; Saqr & Lopez-Pernas, 2021). Three types of indicators were collected that represent online engagement which are mostly behavioral and less so cognitive engagement: 1) frequencies of activities, e.g., interacting, participation and reading online material, 2) the time spent online learning or accessing these resources (as an indication of intensity of engagement and invested effort), and 3) regularity of access reflecting regulation of effort and commitment to learning or the lack thereof (gaps in online learning) (Dvorak, Jia, College, & York, 2016; Jovanović, Saqr, Joksimović, & Gašević, 2021; Saqr et al., 2022).

#### 3.2.1. Frequency of activities

These indicators reflect students’ investment in course work, contribution to the collaborative process and access to learning resources.

1. **Frequency of Course Browsing (FRQ):** The number of times a student viewed the course main page, which displays course announcements and updates (e.g., new announcements by the teacher, new lectures, posts from peers, assignments, etc.) and acts as the gateway for all other resources (Riestra-González, Paule-Ruíz, & Ortin, 2021).
2. **Frequency of Forum Reading (RFM):** The frequency a student reads contributions in the PBL forums, which are explanations, discussions, conclusions, or suggested links to helpful learning resources. Notifications of new forum contributions are sent to students’ emails, and students who follow the links are kept updated and are more likely to check other resources (lectures) within the same session (Conijn, Snijders, Kleingeld, & Matzat, 2017).
3. **Frequency of contributing to the forums (CFM):** The frequency of posting, updating, or creating a forum thread which requires students to compose an argument, co-construct knowledge, or interact with colleagues. As such, CFM reflects students’ participatory cognitive engagement (Saqr, Viberg, & Vartiainen, 2020).
4. **Frequency of learning resource access (LEC):** The frequency of downloading or opening learning resources (Jovanović et al., 2021; Riestra-González et al., 2021).

#### 3.2.2. Time variables

5. **Session count (SC):** The count of sessions which reflects time investment in studying. A session is commonly defined as a single continuous period of online activity (Jovanović et al., 2021).
6. **Average session time (DUR):** The average time that passes between the first and the last learning actions in a session (Conijn et al., 2017).

#### 3.2.3. Regularity variables

7. **Regularity (REG):** Reflects students' consistency in accessing learning resources, and self-regulation and cognitive engagement. The calculation followed the method by Jovanović et al. (2021).
8. **Inactive days gap (GAP):** The median number of inactive days between two periods of activity; by "inactive" we mean any single day without recorded learning activity. The indicator reflects disengagement from the LMS, or possible offline activity.

### 3.3. Analysis

To answer the RQs, two networks (within and between-person) were estimated from the eight variables described previously (in section 3.2). The following steps were taken: First, the data were prepared, and the assumptions were checked. Then, the indicators of online engagement were used to build the networks, then the networks were estimated, plotted, and compared as described in the following subsections.

#### 3.3.1. Data preparation and assumptions check

First, the data was cleaned and prepared for the analysis. Second, the variables were tested to ensure that the correlation matrix is positive-definite i.e., the included variables are not a linear combination of each other (Epskamp & Fried, 2018). Third, the Goldbricker algorithm with Hittner method was used to check for redundantly correlated items. Goldbricker algorithm compares the variables with each other as well as their correlation patterns with variables in the dataset. The algorithm followed the methods recently described in detail by Heeren et al. (2021) and was performed with the *networktools* package (Jones, 2018). The algorithm suggested the removal of some redundant items, which we removed and subsequent iterations resulted in no further suggested redundant items. Fourth, to ensure variables follow a normal distribution, a gaussian transformation was performed using the *huge R* package (Zhao, Liu, Roeder, Lafferty, & Wasserman, 2012) following the methods described in (Epskamp & Fried, 2018). Fifth, both networks were checked for normality using the Shapiro-Wilk test (Jones, 2018). Sixth, the variance of the variables was checked to confirm that the variance is (roughly) equal (Jones, 2018).

#### 3.3.2. Constructing the networks

The methods of constructing the networks relied on the established state of the art as described by previous literature (Bell, Fairbrother, & Jones, 2019; Costantini et al., 2019; Curran & Bauer, 2011; Epskamp, Borsboom, & Fried, 2018) as follows:

**Preparing data for the within-person network:** A within-person network encodes the individual variability from a subject's own "personal average" or how a person varies from his/her own average (Epskamp, Waldorp, et al., 2018). This can be performed using repeated measure data by person-mean centering often referred to as *de-meaning* (Epskamp, Borsboom, & Fried, 2018; Saqr, 2023). *De-meaning* is done by subtracting the mean from each observation for each person. For demonstration, in Fig. 3, Vera's LMS clicks had a mean of 22 clicks across the three measurement points; the de-meaned value for the first measurement point is the mean subtracted from the value of the current value (16–22 = −6) indicating that she has 6 clicks less than her average in the first time-point and so on for each time point and person, and variable (Costantini et al., 2019).

**Preparing data for between-person network:** A between-person network encodes the average group-level behavior and was constructed by averaging variables across all measurement points. Since we have several measurement points for the same person, we compute the mean for all the person observations which represents the central tendency (average) of each variable, offering an arguably less biased estimate (Borsboom et al., 2021; Costantini et al., 2019). As shown in Fig. 2, the mean of Vera's clicks is 22, and that represents her average clicking.

#### 3.3.3. Network estimation

Two psychological networks (between and within-person networks) were estimated using the mean-centered (de-meaned) dataset for the within-person network and the meaned data for the between-person network using the latest recommended estimation methods (Borsboom et al., 2021; Epskamp, Borsboom, & Fried, 2018). The networks were estimated using partial correlation with regularization, where the variables are connected by an edge if they are correlated above and beyond their correlation with all other variables in the network (i.e., *ceteris paribus*). The absence of an edge between two variables indicates that the two variables are independent from each other after controlling for all other variables in the network (Borsboom et al., 2021). The regularization is performed by applying an extra penalty to the network model, and recent literature recommends the procedure for several reasons: 1) it helps eliminate spurious edges, and 2) it shrinks trivial edges to zero and thus helps eliminate type 1 error or "false positive" edges (Epskamp, Borsboom, & Fried, 2018). In doing so, the resulting network model is less complex, sparser, simpler to interpret (Costantini et al., 2019;

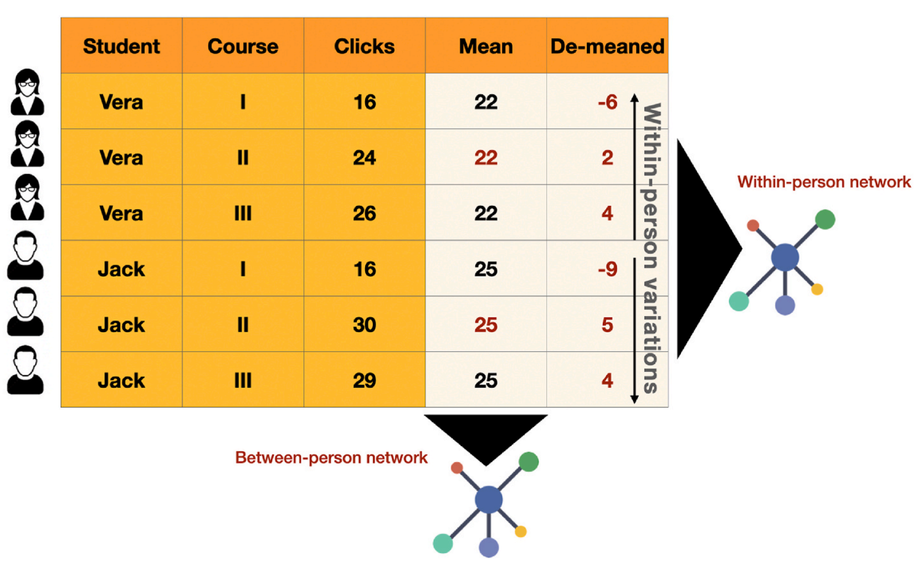


**Fig. 3.** Constructing networks using students' data. The between-student network is constructed of the meaned data (averaged) and the within-person network is constructed from the de-meaned data (mean subtracted from the values of each time point).

Epskamp & Fried, 2018). The network model was estimated using the package *bootnet,* which computes 100 models with various degrees of sparsity. The best model is selected based on the lowest Extended Bayesian Information Criterion value (EBIC) (Epskamp & Fried, 2018). For each of the two networks, another variant was estimated with the addition of the grade as a variable in the network to examine how controlling for the grade influences the interaction between the included variables.

### *3.3.4. Network inference*

Network inference measures were calculated to answer the RQ3 and find out which indicators of engagement are most central, drive the network of engagement indicators to connectivity, or can be explained by interactions with other indicators. We calculated two measures relevant to the research question: *Expected influence and predictability. Expected influence* is the sum of edge weights; a node with higher Expected influence is expected to drive the network positive connectivity (Borsboom et al., 2021). *Predictability* of a given node is a quantification of the extent to which the node connections predict or explain node variance. A node that has high predictability indicates that the network model and the connected variables appropriately explain it. Predictability has also been linked to controllability, i.e., the extent to which acting on the connections of the node would influence the node (Jonas M. B. Haslbeck & Waldorp, 2018). Predictability was computed for each node using the library *mgm* which uses all node connections as parameters for a regression model to calculate the proportion of explained variance of a node given its current connections as (R-square) and Root Mean Squared Error (RMSE). A node with high predictability is well explained by its connections, and vice versa (Jonas M. B. Haslbeck & Waldorp, 2018).

### *3.3.5. Evaluation of stability and rigor of the estimated networks*

To evaluate the stability and accuracy of the estimated networks, we followed the recommended procedure through the implementation of bootstrapping (Epskamp, Borsboom, & Fried, 2018). The results of bootstrapping, replicability, and reproducibility are covered in detail in the appendix.

## 4. Results

The study included 238 students enrolled in six courses with a total of 1428 data points. The average student visited the course main page (FRQ) $\mu = 63.9$, SD = 43.9, contributed to forums (CFM) $\mu = 64.8$, SD = 50.1, read forums (RFM) $\mu = 164.2$, SD = 117.3 and spent an average $\mu = 423.6$, SD = 172.6 min in each course. The descriptive data with mean, standard deviation, 25th percentile, and 75th percentile are displayed in Table 1.

**RQ1a.** between-person (group-level networks)

Mapping the between-persons using the aggregated data across all students represents the "*average engagement process*" and is performed here to establish the group-level insights. The between-person network (Fig. 4) shows strong correlations between *session count* (SC), *frequency of course browsing* (FRQ), *regularity* (REG) and *reading forums* (RFM) after controlling for all other variables in the network, indicating a strong conditional association of these activities on the group-level. Negative correlations between *session count* (SC) and *duration* (DUR), as well as a weak negative correlation between *session count (SC)* and *inactive days* (GAP) were also observed. This constellation indicates that the more sessions students had —on average—, the more they browsed, the more they read the forums, the less time they spent online and the more regular they were. As Fig. 4 shows, *session count* (SC) had the highest predictability $R^2 = 0.9$ and so were the three indicators that were tightly connected, e.g., *frequency course browsing* (FRQ) $R^2 = 0.79$ and *reading forums* (RFM) $R^2 = 0.85$. In other words, the number of sessions is well-explained by its connections and so are the connected variables. Expectedly, *reading forums* (RFM) was strongly correlated to *forum contributing* (CFM) and *duration (DUR)*. However, duration had $R^2 = 0.62$ and *forum contributing* (CFM) had $R^2 = 0.72$ indicating above average predictability, which was lower than most other variables; see Table 2 for details. To summarize, we see that– on the group-level behavioral engagement is self-amplifying and acts as a catalyst for cognitive engagement with PBL forums.

The constellation around the *inactive days* (GAP) is of particular interest since this is the behavior that we —as educators or researchers— seek to understand or optimize. *Inactive days* (GAP) was strongly and negatively correlated with *regularity (REG)* and weakly with *session count* (SC) and *forum contributing* (CFM). Nonetheless, the predictability of *inactive days* (GAP) was the least in the network $R^2 = 0.61$ indicating that it is the least controllable or responsive to possible intervention. In psychological networks, the absence of links is equally important to the presence thereof and we see that *inactive days* (GAP) is the least connected, indicating conditional independence from most variables, e.g., *forum contributing* (CFM) and *frequency course browsing* (FRQ) and so they poorly explain why a student is inactive for some days (disengaged). Put another way, poor online engagement reflects or explains disengagement.

The between-person network with grade was constructed so each connection in the network is conditioned on the grade level, i.e., independent of level of achievement. The correlations in the network in Fig. 4 (right side) were similar to the between-person network with small differences that ranged from 0 to 0.02. Only the connection between *forum contributing* (CFM) and *regularity (REG)* was present in the between-person network but absent from the grade between-person network, which means that regularity and forum posting are dependent on achievement level. In other words, students' approach to online learning was rather similar except for indicators of cognitive engagement.

The grades were strongly connected to *regularity* (REG), and *forum reading* (RFM) after controlling for all other variables. However, the predictability of the grade was the lowest of all variables $R^2 = 0.42$ indicating that while the variables in the network are fairly explained with a mean of $R^2 = 0.74$, there are other variables that contribute to the explanation of grades not included in the network i.e., online behavior can partially explain the variability of grades. The predictability of *inactive days* (GAP) did not improve after inclusion of the grades in the network indicating that inactivity is also poorly explained by online data only.

**RQ1a and RQ2.** within-person (personal-level networks)

The within-person networks in Fig. 5 show the individual behavior compared to their previous state. There was a strong correlation between *session count* (SC), *frequency of course browsing* (FRQ), *regularity* (REG) and *forum reading* (RFM) as well as a strong correlation between *forum reading* (RFM) and *forum contributing* (CFM), a negative correlation between *regularity* (REG) and *session count* (SC) as well as duration (DUR), indicating that such a constellation of behaviors does relatively generalize from between-person to individuals.

**Table 1**
The descriptive statistics of each student per course.

| | Mean | SD | 25% | 75% |
|---|---|---|---|---|
| REG | 73.07 | 11.40 | 66.48 | 81.10 |
| FRQ | 63.92 | 43.91 | 35.00 | 81.00 |
| RFM | 164.21 | 117.33 | 78.00 | 226.25 |
| CFM | 64.76 | 50.11 | 28.00 | 89.00 |
| LEC | 59.85 | 40.30 | 32.00 | 78.00 |
| SC | 55.46 | 31.79 | 33.00 | 71.00 |
| DUR | 423.58 | 172.60 | 302.78 | 512.68 |
| GAP | 0.81 | 2.17 | 0.00 | 2.04 |
| FG | 74.25 | 11.11 | 67.96 | 82.09 |

SD = standard deviation, 25% = 25th percentile, 75% = 75th percentile.

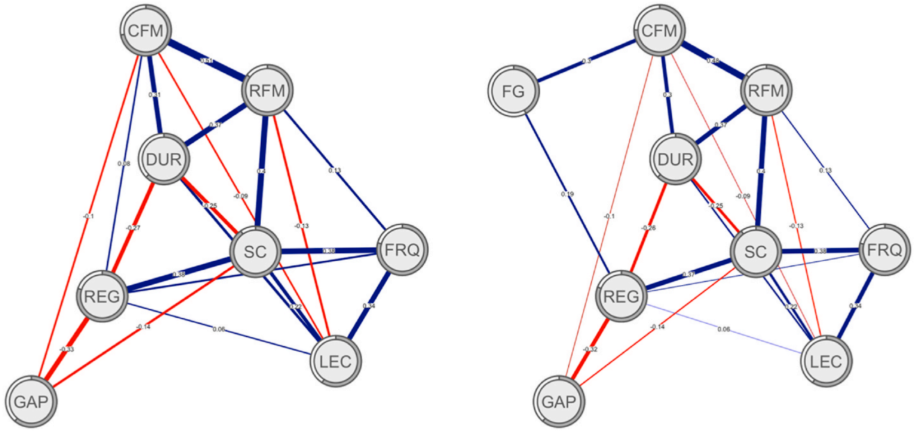


**Fig. 4.** Left: Between-person network. Right: Between-person with grades.

**Table 2**
Predictability $R^2$ and RMSE of the four networks.

| Variable | Between-person | | Between-person with grades | | Within-person | | Within-person with grades | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | $R^2$ | RMSE | $R^2$ | RMSE | $R^2$ | RMSE | $R^2$ |
| REG | 0.45 | 0.80 | 0.44 | 0.81 | 0.76 | 0.42 | 0.76 | 0.43 |
| FRQ | 0.45 | 0.79 | 0.45 | 0.79 | 0.60 | 0.64 | 0.60 | 0.64 |
| RFM | 0.38 | 0.85 | 0.38 | 0.85 | 0.58 | 0.66 | 0.58 | 0.66 |
| CFM | 0.53 | 0.72 | 0.50 | 0.75 | 0.71 | 0.49 | 0.71 | 0.49 |
| LEC | 0.64 | 0.59 | 0.64 | 0.59 | 0.84 | 0.29 | 0.84 | 0.29 |
| SC | 0.31 | 0.90 | 0.31 | 0.90 | 0.47 | 0.78 | 0.47 | 0.78 |
| GAP | 0.62 | 0.61 | 0.62 | 0.61 | 0.89 | 0.20 | 0.89 | 0.20 |
| DUR | 0.61 | 0.62 | 0.61 | 0.62 | 0.79 | 0.38 | 0.79 | 0.38 |
| FG | NA | NA | 0.76 | 0.42 | NA | NA | 0.98 | 0.05 |
| Mean | 0.50 | 0.74 | 0.52 | 0.71 | 0.71 | 0.48 | 0.74 | 0.44 |
| SD | 0.12 | 0.12 | 0.14 | 0.15 | 0.14 | 0.20 | 0.16 | 0.24 |

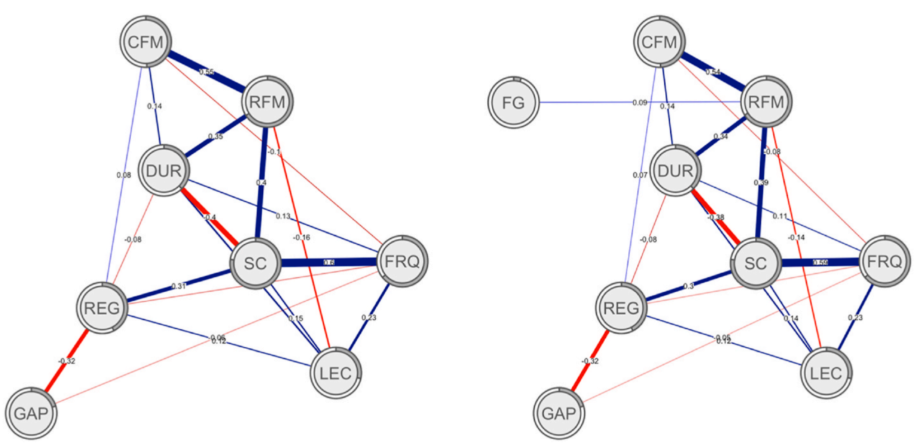


**Fig. 5.** Left, the within-person network. Right, the Final grade within-person network.

However, there were several differences regarding the strength of correlations, different connection constellations, correlations with reversed signs and differences in the *predictability* of variables as shown in Fig. 6. First, in the within-person network, the correlation between *session count* (SC) and *frequency of course browsing* (FRQ) was remarkably stronger; *session count* (SC) had relatively more negative correlation with *duration* (DUR) and was strongly correlated to *regularity* (REG), indicating a more consistent pattern of behavior at the person-level consisting of regular sessions, shorter durations, more frequent browsing, reading, and contributing to forums. In other words, behavioral engagement drives more engagement or is more consistent on the individual level. Second, the correlation between *inactive days* (GAP) and *forum reading* (RFM) does not exist in the within-person network. Third, *regularity* was **negatively** and weakly correlated with *frequency of course browsing* (FRQ) compared to a positive correlation in the between-person network, which shows a case of Simpson's paradox. Fourth, the predictability of all variables was lower than the between-person network (see next section for details). The grade was only connected to *forum reading* (RFM) after controlling for all other variables. The predictability of the grades (FG) was 0.05, which is the lowest of all the variables in the four networks, with a RMSE of = 0.98, the highest of all the variables.

**RQ3.** Inference, comparison between the two networks

The average predictability of the between-person network was 0.74,

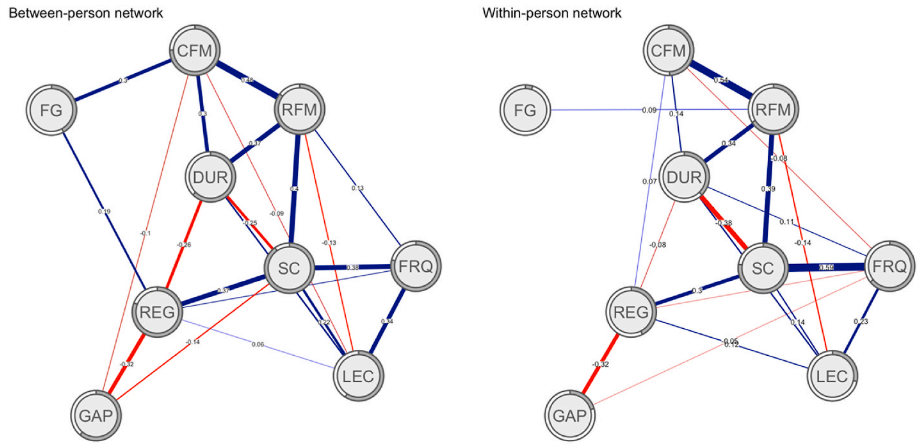


**Fig. 6.** Between-person network side-by-side with the within-person network.

which means that —on average— 74% of the variance of each node was explained by its connections. The average predictability of the within-person network was $R^2 = 0.48$ and the predictability of all the variables was less than their counterparts in the between-person network as shown in Fig. 7. The difference was least remarkable —though not trivial— in the frequency variables, e.g., *session count* (SC) and *frequency of course browsing* (FRQ). In both networks, *inactive days* (GAP) and grades were the indicators with the least predictability. Given the overall low predictability for within-person network, the disengagement, and grades, we need further research on what offers better account for the variation of individual processes.

The expected influence centrality was relatively higher in the between-person network in most variables, except for *session count* (SC) as shown in Fig. 7. In the between-person network, the *forum reading* (RFM) was the variable with most expected influence, followed by *session count* (SC), *frequency of course browsing* (FRQ) and *forum contributing* (CFM). *Inactive days* (GAP) was the variable with the least influence −0.47 indicating, the more the inactive days, the more we expect to see decrease in strength of correlations, see the appendix Table A1 for detailed numerical results.

## 5. Discussion

A central assumption of the scientific method is that inferences derived from group-level analysis align-with and generalize-to the individual-level (Fisher et al., 2018). Such assumptions are used as a base to design adaptive learning environments, to create intervention and to understand individual students' processes at large. Therefore, which methods can be used to build adaptive insights that generalizes to individual students is one of the main challenges of learning analytics nowadays (Sailer et al., 2023). This study was performed to put this assumption to the test and examine if and to what extent our analysis – of online engagement – generalizes from group to individual level. For that purpose, the study used the latest developments in psychological network methods to account for the interactions of different indicators.

On the group-level, the topology of engagement indicators showed a constellation of variables that were tightly connected —reflecting mostly behavioral engagement— including frequency of browsing, reading, and writing forum posts and shorter duration of sessions (controlling for all other variables). More importantly, the predictability of such indicators (how the connections of variables explain them) was high ($R^2 = 0.74$); i.e., such a constellation of behavioral engagement was fairly explainable by the between-person network. A similar constellation was obtained in the within-person network–with relatively lower correlations, indicating that such a pattern of tightly interconnected online behavioral engagement can be expected to generalize from group to individual-level. Furthermore, these findings offer evidence that engagement kindles more engagement both in quantity (behavioral

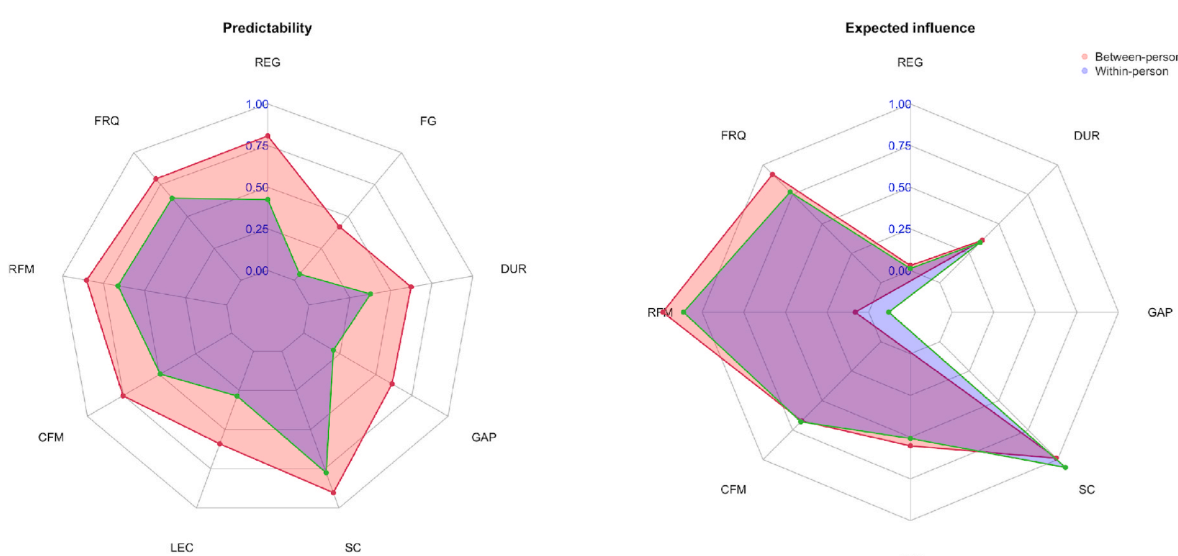


**Fig. 7.** Comparison between the two networks regarding the predictability (Left), and regarding expected influence (Right).

engagement) and quality, i.e., cognitive engagement.

However, the two variables that have potential for intervention and thus are important for educators (the grades, and days of online disengagement) were the least explainable by the group-level network. Similarly, in the within-person network, these variables had the least predictability (low $R^2$). Predictability reflects the upper bound of controllability or the extent of change that an intervention could bring (J. M. B. Haslbeck & Fried, 2017). Such results indicate that an intervention – using only online data– that targets improving online engagement may moderately result in an improvement in the overall average grade of a classroom and can hardly improve individual students' grades or disengagement. The results are also a strong indicator that online engagement poorly reflects, explains or accounts for disengagement.

The centrality measure of *expected influence* gave us a clue about the variables that are expected to positively influence online engagement. The variables with highest *expected influence* were frequency of sessions, browsing, contributing-to or reading the forums (behavioral engagement indicators). The finding that forum reading had the highest *expected influence* comes from the nudging effect of forum summary (which is sent via email) and also from the fact that reading the forums stimulates deeper engagement with the cognitive content of the PBL problems. That is, Moodle sends an email notifying students of new posts. Such notifications serve a dual function: first, alerts students to reply; second, students may explore other learning resources in the same session. In fact, we think that one of the most important take-home messages is how nudging forum reading positivity stimulates other elements of the tightly associated online behaviors. Understanding such insights was possible by studying the structure and tight interdependence of online behavior. Thus, nudging can possibly be applied to other activities, e.g., lectures, and expectedly, nudging lectures may result in engagement with forum reading and posting. There is vast evidence that has shown that nudging and prompts helps enhance behavior on a wide range of activities (Mertens, Herberz, Hahnel, & Brosch, 2022), and learning (Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015; Berthold, Nückles, & Renkl, 2007). Yet, caution should be exercised that such intervention may not be very effective in influencing individual students' grades according to within-person network low levels or predictability. In summary, behavioral engagement drives more behavioral and cognitive engagement and efforts that stimulate any type of engagement are expected to catalyze other types of engagement as well.

The second research question examined how the within-person network compares to the between-person network. Besides the aforementioned similarities, several differences have been revealed. First, there was a negative correlation between regularity and frequency (after controlling for all other variables), which is contrary to the group-level analysis in our study and other studies (Jovanović et al., 2021; Saqr, Fors, & Tedre, 2017). Such a reverse relationship is a case of Simpson's paradox i.e., a relationship on the group level exists in the opposite direction from the individual level. Simpson's paradox may occur when —inter alia — a phenomenon has a mechanism at the group-level different from the individual level (Kievit, Frankenhuis, Waldorp, & Borsboom, 2013; Mangalam & Kelty-Stephen, 2021; Tu, Gunnell, & Gilthorpe, 2008). On the individual-level, this can be explained by the fact that human behavior follows a bursty and irregular nature based on perceived priority (Barabási, 2005; Cencetti, Battiston, Lepri, & Karsai, 2021; Vázquez et al., 2006; Wu, Qi, Shi, Li, & Yan, 2022); some tasks are executed in time, while other tasks may be queued with long waiting times, resulting in an irregular rhythm (Barabási, 2005; Cencetti et al., 2021; Wu et al., 2022). Further support for the randomness and hence irregularity of human activity has been proven across several fields, e.g., browsing the internet, communicating with email or visiting the library (Cencetti et al., 2021; Vázquez et al., 2006). Time, regularity and scheduling are buzz words in self-regulated learning, yet, there seems to be a poor understanding of the time dynamics within the process of self-regulation on the individual level (Järvelä & Bannert, 2021; Reimann, 2019), and therefore, research may be needed to unravel the time dynamics of learning, especially with the prevalence of advanced sensors and monitors (Järvelä & Bannert, 2021).

Previous research —covered in section 2.4— has shown and emphasized the importance of within-person variability Malmberg et al. (2022b), addressed different time scales, and confirmed the differences between group-level and individual level inferences e.g., (Bakker et al., 2015; Collie, Malmberg, Martin, Sammons, & Morin, 2020; Dietrich, Viljaranta, Moeller, & Kracke, 2017; Vasalampi, Muotka, Malmberg, Aunola, & Lerkkanen, 2021). While our results emphasize the within-person aspect, it is hard to draw parallels with such previous research given the different time scale (course in our study), different data source (online behavior) and different methods (networks). Our methods have demonstrated the interplay between variables, the mutual influence and to what extent the variables explain each other; a gap that Järvelä and Bannert (2021) described as "what is still not clear, is when those actions take place, how they influence each other, and how they refer to learning performance".

## 6. Implications

The low average predictability of the within-person network, the inverse pattern of correlation between frequency and regularity, and the differences in network topology are indications of the poor congruence between the within-person and between-person networks, in particular, grades and inactivity (the variables that matter the most to educators). Therefore, generalizing group-level findings to individuals may not be warranted (Dirk & Schmiedek, 2016; Fisher et al., 2018; Winne, 2017). As Fisher et al. (2018) put it, using aggregate group-level inferences to draw conclusions about intervention at the individual level (student in our case) who is the basic unit of our analysis, is "far less accurate or valid than it may appear in the literature".

The ability to model within- and between-person dynamics opens the door for future opportunities to model the complex phenomena on the individual level. We see a potential that other constructs, e.g., self-regulation, motivation, or self-efficacy, can be studied using the psychological network methods; such methods could model the complexity of the temporal and multi-dimensional aspects of phenomena and help identify possible targets for intervention. Since psychological networks accommodate the different interactions within a dynamical complex system, "they form a natural bridge from data analysis to theory formation based on network science principles" (Borsboom et al., 2021). Furthermore, the analysis introduced in this study offers a method for "closing the loop" form analysis of data to proper adaptive insights that can better generalize to individual learners (Sailer et al., 2023).

Three levels of intervention are relevant to network methods: the variables (nodes), the connections (correlations), and external factors (factors not in the network). The centrality measure (expected influence) offered an idea about the important nodes as targets of intervention (e.g., driving behavioral engagement), possibly through nudging or prompting. The predictability offered an indication of the upper bound of expected change given the network structure and variables. An intervention using the existing variables could help improve achievement on the group level moderately, and very weakly on the individual level. The third aspect that was not studied are the variables outside the network, which are theoretically numerous. However, we have an estimate that these variables are very significant on the person-level; in other words, the large unexplained variance in grade and inactivity on the individual level indicate the need for research on how to understand person-level processes and how to positively intervene. Future research could attempt to understand the within-person learning processes, possibly involving other indicators (e.g., dispositions), use different timescales, or develop novel methods and data analysis methods.

## 7. Limitations

This study has limitations regarding the collected data and the analysis methods. The trace data is far from perfect, and has limitations regarding accuracy, breadth, and scope. For instance, trace data does not tell if a student is actively studying, multitasking, or just clicking and therefore, the time estimated from logs may suffer some inaccuracies. Trace data contains only clickable items, and therefore, does not contain data related to other non-clickable learning events, e.g., studying from an online book or a YouTube video. Researchers should be aware that they are modeling a part of the picture that is definitely incomplete and can be – sometimes – misleading. In fact, our data has clearly shown that when students have *gaps* in their engagement, modeling becomes less reliable. It is therefore advisable to expand the data repertoire about students and resort to more ecologically valid data sources e.g., Ecological Momentary Assessment (EMA). Other solutions include using smart devices that can capture passively sensed data (e.g., screen times, activities, eye movements, etc.). Psychological networks have known limitations regarding their estimation methods. First, we have borrowed methods and parameters developed for the study of psychological phenomena, which may need to be adjusted for educational research, however, there is not enough previous research to guide the optimal choices. Network methods use regularization techniques to ensure networks are sparse, while this method has been shown to return an interpretable network structure, it is most appropriate when the network is actually sparse. Several methods are currently tested that may enable possible alternatives for researchers (Borsboom et al., 2021). Modeling complex and highly dimensional data is challenging, whereas psychological networks employ several techniques to facilitate the interpretation of such data (e.g., regularization), novel methods of dimensionality reduction may be needed. Several researchers have suggested solutions such as clustering similar constructs together using traditional community detection methods or dimensionality reduction techniques such as principal component analysis (Wigman et al., 2015). Once several variables are grouped into constructs, it becomes easier to visualize a network with few constructs compared to another with a large number of nodes. In particular, the recently introduced Exploratory Graph Analysis (EGA) – though experimental– offers an intuitive solution that clusters lower-order dimensions together using the Louvain algorithm and then uses factor or network loadings to map them to higher-order dimensions with fewer nodes (Jiménez et al., 2023).

This study has just scratched the surface in a vast field that requires concerted efforts to map the within-person processes, understand idiographic behavior compared to group-level behavior and most importantly develop accurate models that can predict the need for help and support and ideally guide the steps or approach to such help (Bakker et al., 2015; Dirk & Schmiedek, 2016; Martin et al., 2015).

### Credit author statement

The article was conceptualized by the MS. Data curation, preparation, analysis and writing has been performed by MS.

### Ethical approval

The study received ethical approval ID:6065 from the University Ethical Board.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

The authors do not have permission to share data.

### Acknowledgments

The paper is co-funded by the Academy of Finland for the project TOPEILA, Decision Number 350560 which was received by the author.

The authors also acknowledge the help, support and valuable feedback on the manuscript by Sonsoles Lopez-pernas and Ramy Elmoazen.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chb.2023.107991.

### References

Bakker, A. B., Sanz Vergel, A. I., & Kuntze, J. (2015). Student engagement and performance: A weekly diary study on the role of openness. *Motivation and Emotion, 39*(1), 49–62. https://doi.org/10.1007/s11031-014-9422-5

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038

Barabási, A.-L. (2005). The origin of bursts and heavy tails in human dynamics. *Nature, 435*(7039), 207–211. https://doi.org/10.1038/nature03459

Beck, E. D., & Jackson, J. J. (2020). Consistency and change in idiographic personality: A longitudinal ESM network study. *Journal of Personality and Social Psychology, 118*(5), 1080–1100. https://doi.org/10.1037/pspp0000249

Beck, E. D., & Jackson, J. J. (2021). Within-person variability. In *The handbook of personality dynamics and processes* (pp. 75–100). Elsevier. https://doi.org/10.1016/b978-0-12-813995-0.00004-2.

Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality and Quantity, 53*(2), 1051–1074. https://doi.org/10.1007/s11135-018-0802-x

Beltz, A. M., Wright, A. G. C., Sprague, B. N., & Molenaar, P. C. M. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment, 23*(4), 447–458. https://doi.org/10.1177/1073191116648209

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., et al. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers, 1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Cencetti, G., Battiston, F., Lepri, B., & Karsai, M. (2021). Temporal properties of higher-order interactions in social networks. *Scientific Reports, 11*(1), 7028. https://doi.org/10.1038/s41598-021-86469-8

Collie, R. J., Malmberg, L.-E., Martin, A. J., Sammons, P., & Morin, A. J. S. (2020). A multilevel person-centered examination of teachers' workplace demands and resources: Links with work-related well-being. *Frontiers in Psychology, 11*, 626. https://doi.org/10.3389/fpsyg.2020.00626

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using moodle LMS. *IEEE Transactions on Learning Technologies, 10*(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. *Self Processes and Development, 268*, 43–77. https://psycnet.apa.org/fulltext/1991-97029-002.pdf.

Cook, C. R., Kilgus, S. P., & Burns, M. K. (2018). Advancing the science and practice of precision education to enhance student outcomes. *Journal of School Psychology, 66* (December 2017), 4–10. https://doi.org/10.1016/j.jsp.2017.11.004

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences, 136*, 68–78. https://doi.org/10.1016/j.paid.2017.06.011

Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*, 583–619. https://doi.org/10.1146/annurev.psych.093008.100356

Dietrich, J., Viljaranta, J., Moeller, J., & Kracke, B. (2017). Situational expectancies and task values: Associations with students' effort. *Learning and Instruction, 47*, 53–64. https://doi.org/10.1016/j.learninstruc.2016.10.009

Dirk, J., & Schmiedek, F. (2016). Fluctuations in elementary school children's working memory performance in the school context. *Journal of Educational Psychology, 108* (5), 722–739. https://doi.org/10.1037/edu0000076

Dvorak, T., Jia, M., College, Union, & York, New. (2016). Do the timeliness , regularity , and intensity of online work habits predict academic performance. *Journal of Learning Analytics, 3*(3), 318–330. http://www.learning-analytics.info/journals/index.php/JLA/article/view/4676.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods, 23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America, 115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Glaser, R. (1989). Learning theory and the study of instruction. *Annual Review of Psychology, 40*(1), 631–666. https://doi.org/10.1146/annurev.psych.40.1.631

Hamaker, E. L. (2012). Why researchers should think within-person: A paradigmatic rationale. *Handbook of Research Methods for Studying Daily Life*, 43–61. October.

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine, 47*(16), 2767–2776. https://doi.org/10.1017/s0033291717001258

Haslbeck, J. M. B., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior Research Methods, 50*(2), 853–861. https://doi.org/10.3758/s13428-017-0910-x

Heeren, A., Lannoy, S., Coussement, C., Hoebeke, Y., Verschuren, A., Blanchard, M. A., et al. (2021). A network approach to the five-facet model of mindfulness. *Scientific Reports, 11*(1), Article 15094. https://doi.org/10.1038/s41598-021-94151-2

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Järvelä, S., & Bannert, M. (2021). Temporal and adaptive processes of regulated learning - what can multimodal data tell? *Learning and Instruction, 72*, Article 101268. https://doi.org/10.1016/j.learninstruc.2019.101268

Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2023). Dimensionality assessment in bifactor structures with multiple general factors: A network psychometrics approach. *Psychological Methods*. https://doi.org/10.1037/met0000590

Jones, P. J. (2018). *Networktools: Tools for identifying important nodes in networks [R package*.

Jovanovic, J., Mirriahi, N., Gašević, D., Dawson, S., Pardo, A., Gasevic, D., et al. (2019). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers and Education, 134*(February), 156–168. https://doi.org/10.1016/j.compedu.2019.02.011

Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education, 172*(April), Article 104251. https://doi.org/10.1016/j.compedu.2021.104251

Kassab, S. E., El-Sayed, W., & Hamdy, H. (2022). Student engagement in undergraduate medical education: A scoping review. *Medical Education, 56*(7), 703–715. https://doi.org/10.1111/medu.14799

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology, 4*, 513. https://doi.org/10.3389/fpsyg.2013.00513

Malmberg, J., Saqr, M., Järvenoja, H., Haataja, E., Pijeira-Díaz, H. J., & Järvelä, S. (2022a). Modeling the complex interplay between monitoring events for regulated learning with psychological networks. In *The multimodal learning analytics handbook* (pp. 79–104). Springer International Publishing. https://doi.org/10.1007/978-3-031-08076-0_4.

Malmberg, J., Saqr, M., Järvenoja, H., & Järvelä, S. (2022b). How the monitoring events of individual students are associated with phases of regulation: A network analysis approach. *Journal of Learning Analytics, 9*(1), 77–92. https://learning-analytics.info/index.php/JLA/article/view/7429.

Mangalam, M., & Kelty-Stephen, D. G. (2021). Point estimates, Simpson's paradox, and nonergodicity in biological sciences. *Neuroscience & Biobehavioral Reviews, 125* (February), 98–107. https://doi.org/10.1016/j.neubiorev.2021.02.017

Martin, A. J., Papworth, B., Ginns, P., Malmberg, L. E., Collie, R. J., & Calvo, R. A. (2015). Real-time motivation and engagement during a month at school: Every moment of every day for every student matters. *Learning and Individual Differences, 38*, 26–35. https://doi.org/10.1016/j.lindif.2015.01.014

Mertens, S., Herberz, M., Hahnel, U. J. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences of the United States of America, 119*(1), Article e2107346118. https://doi.org/10.1073/pnas.2107346118

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*(2), 112–117. https://doi.org/10.1111/j.1467-8721.2009.01619.x

Reimann, P. (2019). Methodological progress in the study of self-regulated learning enables theory advancement. *Learning and Instruction, 72*, Article 101269. https://doi.org/10.1016/j.learninstruc.2019.101269

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–19). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_1.

Reschly, A. L., & Christenson, S. L. (2022). Jingle-jangle revisited: History and further evolution of the student engagement construct. In A. L. Reschly, & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 3–24). Springer International Publishing. https://doi.org/10.1007/978-3-031-07853-8_1.

Richters, J. E. (2021). Incredible utility: The lost causes and causal debris of psychological science. *Basic and Applied Social Psychology, 43*(6), 366–405. https://doi.org/10.1080/01973533.2021.1979003

Riestra-González, M., Paule-Ruíz, M. del P., & Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers and Education, 163*(July 2020). https://doi.org/10.1016/j.compedu.2020.104108

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. In *WIREs data mining and knowledge discovery* (Vol. 10). https://doi.org/10.1002/widm.1355. Issue 3.

Saqr, M. (2023). Modelling within-person idiographic variance could help explain and individualize learning. In *British journal of educational technology*. https://doi.org/10.1111/bjet.13309

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher, 39*(7), 757–767. https://doi.org/10.1080/0142159X.2017.1309376

Saqr, M., Jovanovic, J., Viberg, O., & Gašević, D. (2022). Is there order in the mess? A single paper meta-analysis approach to identification of predictors of success in learning analytics. *Studies in Higher Education, 1–22*. https://doi.org/10.1080/03075079.2022.2061450

Saqr, M., & López-Pernas, S. (2021). The longitudinal trajectories of online engagement over a full program. *Computers & Education, 175*, Article 104325. https://doi.org/10.1016/j.compedu.2021.104325

Saqr, M., & Lopez-Pernas, S. (2021). Idiographic learning analytics: A single student (N=1) approach using psychological networks. *Companion Proceedings 11th International Conference on Learning Analytics & Knowledge*, (LAK21), 397–404.

Saqr, M., López-Pernas, S., Helske, S., & Hrastinski, S. (2023). The longitudinal association between engagement and achievement varies by time, students' profiles, and achievement state: A full program study. *Computers & Education, 199*(104787), Article 104787. https://doi.org/10.1016/j.compedu.2023.104787

Saqr, M., Viberg, O., & Peteers, W. (2021). Using psychological networks to reveal the interplay between foreign language students' self-regulated learning tactics. *STELLA2020 Proceedings, 2828*, 12–23.

Saqr, M., Viberg, O., & Vartiainen, H. (2020). Capturing the participation and social dimensions of computer-supported collaborative learning through social network analysis: Which method and measures matter? *International Journal of Computer-Supported Collaborative Learning, 15*(2), 227–248. https://doi.org/10.1007/s11412-020-09322-6

Sailer, M., Ninaus, M., Huber, S. E., Bauer, E., & Greiff, S. (2023). Conceptualizing learning analytics from a holistic perspective. *Computers in Human Behavior (in press)*.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities, 48*(4), 369–390. https://doi.org/10.1177/0022219413504995

Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: A systematic literature review. *IEEE Access: Practical Innovations, Open Solutions, 10*, 72480–72503. https://doi.org/10.1109/access.2022.3188767

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

Skinner, E. A., & Raine, K. E. (2022). Unlocking the positive synergy between engagement and motivation. In A. L. Reschly, & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 25–56). Springer International Publishing. https://doi.org/10.1007/978-3-031-07853-8_2.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development, 88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. https://doi.org/10.3102/00346543045001089

Tinto, V. (2022). Exploring the character of student persistence in higher education: The impact of perception, motivation, and engagement. In A. L. Reschly, & S. L. Christenson (Eds.), *Handbook of research on student engagement* (pp. 357–379). Springer International Publishing. https://doi.org/10.1007/978-3-031-07853-8_17.

Trautwein, U., Schnyder, I., Niggli, A., Neumann, M., & Lüdtke, O. (2009). Chameleon effects in homework research: The homework–achievement association depends on the measures used and the level of analysis chosen. *Contemporary Educational Psychology, 34*(1), 77–88. https://doi.org/10.1016/j.cedpsych.2008.09.001

Tu, Y.-K., Gunnell, D., & Gilthorpe, M. S. (2008). Simpson's Paradox, Lord's Paradox, and Suppression Effects are the same phenomenon–the reversal paradox. *Emerging Themes in Epidemiology, 5*(1), 2. https://doi.org/10.1186/1742-7622-5-2

Urdan, T., & Kaplan, A. (2020). The origins, evolution, and future directions of achievement goal theory. *Contemporary Educational Psychology, 61*, Article 101862. https://doi.org/10.1016/j.cedpsych.2020.101862

Vasalampi, K., Muotka, J., Malmberg, L.-E., Aunola, K., & Lerkkanen, M.-K. (2021). Intra-individual dynamics of lesson-specific engagement: Lagged and cross-lagged effects from one lesson to the next. *British Journal of Educational Psychology, 91*(3), 997–1014. https://doi.org/10.1111/bjep.12404

Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K.-I., Kondor, I., & Barabási, A.-L. (2006). Modeling bursts and heavy tails in human dynamics. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 73*(3 Pt 2), Article 036127. https://doi.org/10.1103/PhysRevE.73.036127

Wang, M.-T., Degol, J. L., & Henry, D. A. (2019). An integrative development-in-sociocultural-context model for children's engagement in learning. *American Psychologist, 74*(9), 1086–1102. https://doi.org/10.1037/amp0000522

Webster, E. A., & Hadwin, A. F. (2015). Emotions and emotion regulation in undergraduate studying: Examining students' reports from a self-regulated learning perspective. *Educational Psychology, 35*(7), 794–818. https://doi.org/10.1080/01443410.2014.895292

Wigfield, A., & Eccles, J. S. (2020). Chapter five - 35 years of research on students' subjective task values and motivation: A look back and a look forward. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 7, pp. 161–198). Elsevier. https://doi.org/10.1016/bs.adms.2019.05.002.

Wigman, J. T. W., Van Os, J., Borsboom, D., Wardenaar, K. J., Epskamp, S., Klippel, A., et al. (2015). Exploring the underlying structure of mental disorders: Cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine, 45*(11), 2375–2387. https://www.cambridge.org/core/journals/psychological-medicine/article/exploring-the-underlying-structure-of-mental-disorders-crossdiagnostic-differences-and-similarities-from-a-network-perspective-using-both-a-topdown-and-a-bottomup-approach/2D4B0E386A06A15049CFD19F9E4E549B.

Winne, P. (2017). Leveraging big data to help each learner and accelerate learning science. *Teachers College Record, 119*(3), 1–24. https://www.tcrecord.org/books/pdf.asp?ContentID=21769.

Wood, D. F. (2003). Problem based learning ABC of learning and teaching in medicine: Topic collections ABC of learning and teaching in medicine Problem based learning. *BMJ, 326*(7384), 328–330. https://doi.org/10.1136/bmj.326.7384.328

Wu, L., Qi, J., Shi, N., Li, J., & Yan, Q. (2022). Revealing the relationship of topics popularity and bursty human activity patterns in social temporal networks. *Physica A, 588*(126568), Article 126568. https://doi.org/10.1016/j.physa.2021.126568

Zhao, T., Liu, H., Roeder, K., Lafferty, J., & Wasserman, L. (2012). The huge package for high-dimensional undirected Graph estimation in R. *Journal of Machine Learning Research: JMLR, 13*, 1059–1062. https://www.ncbi.nlm.nih.gov/pubmed/26834510.