Learning and Individual Differences 121 (2025) 102705

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Disentangling doctoral well-being support in progress-focused workshops: Combining qualitative and quantitative data in single-case learning analytics☆

Luis P. Prieto[a,*], Jelena Jovanovic[b], Paula Odriozola-González[c], María Jesús Rodríguez-Triana[a,d], Henry Benjamín Díaz-Chavarría[a], Yannis Dimitriadis[a]

[a] GSIC-EMIC research group, Universidad de Valladolid, Spain
[b] Department of Software Engineering, University of Belgrade, Serbia
[c] Department of Education, Universidad de Cantabria, Spain
[d] School of Digital Technologies, Tallinn University, Estonia

ARTICLE INFO

Keywords:
Doctoral education
Learning analytics
Well-being
Idiographic methods
Mixed methods

ABSTRACT

Doctoral education (DE) suffers from widespread well-being issues. Recent evidence from short-term training actions shows potential to address them, but also large variability. Further, DE practitioners face challenges in understanding whether (and for whom) such interventions work, due to small sample sizes, short intervention durations, and the inherent uniqueness of each dissertation. This methodological paper proposes a novel, practice-oriented, and idiographic approach to such understanding, supported by learning analytics of quantitative and qualitative data. To illustrate this approach, we apply it to two datasets from six authentic doctoral workshops ($N = 105$ doctoral students), showcasing how it can provide individualized practice-oriented insights to doctoral students and help trainers better understand their interventions, while coping with typical limitations of data from doctoral training. These findings exemplify how the triangulation of simple, interpretable analytics models of mixed longitudinal data can improve students, practitioners', and researchers' understanding, redesign, and personalization of such training actions.

Educational relevance and implications statement: Collecting data about the context and process of a doctoral training action can help practitioners and students understand who benefits more (or less) from such training. The individualized analysis of such data, obtained with even very simple technologies, can also help students understand their processes and contexts, to better address progress and well-being issues. The use of student-authored short narratives (e.g., diaries), along with longitudinal quantitative data, plays an important role in these personalized analyses, and the promise of automated qualitative coding makes this approach increasingly feasible.

1. Introduction

Doctoral education (DE) suffers from high prevalence of emotional well-being issues (e.g., anxiety, depression, stress), estimated at about 10–25 % (Satinsky et al., 2021). Given recent estimates of the doctoral student population (about three million, see Taylor, 2021), these issues affect hundreds of thousands of doctoral students worldwide.

DE research has documented many factors as related to these problems (e.g., Sverdlik et al., 2018), including individual, interpersonal, and socio-economic factors. Despite this wealth of known factors, there is a lack of evidence-based interventions to address well-being issues in DE (Jackman et al., 2022; Mackie & Bates, 2019).

A key obstacle to solving these issues is the inherent uniqueness of each doctoral topic, process, and candidate. This makes it difficult to define meaningful cohorts. For instance, Van der Linden et al. (2018) suggest that contextual factors are crucial in understanding doctoral well-being phenomena. Others have noted similar problems of heterogeneity (in contexts and individual differences) in higher education in

☆ This article is part of a Special issue entitled: 'Person-specific analytics' published in Learning and Individual Differences.

* Corresponding author at: GSIC-EMIC research group, School of Telecommunications Engineering, Universidad de Valladolid, Paseo de Belén, 15-47011 Valladolid, Spain.
E-mail address: luispablo.prieto@uva.es (L.P. Prieto).

https://doi.org/10.1016/j.lindif.2025.102705
Received 15 June 2024; Received in revised form 1 March 2025; Accepted 16 April 2025
Available online 25 April 2025

general (Gašević et al., 2016; Kizilcec et al., 2020; Saqr, 2023) – but this is bound to be more acute in DE, where there is no cohort, no instructional design, and the learning process is, by definition, unique.

Collecting evidence about the effects of interventions to address this wicked problem is hard for both researchers and doctoral training practitioners (e.g., trainers or facilitators of doctoral seminars). Key problems in this regard are the small sample sizes, short durations, and highly heterogeneous groups of doctoral students attending these interventions, especially in authentic doctoral practice. Practitioners face the additional hurdle of developing training actions/interventions that usually are not a mandatory part of the doctoral curriculum.

An example of training actions facing these challenges are recent small-scale studies on the design and evaluation of doctoral workshops (Prieto et al., 2022), focusing on the importance of making steady progress in the development of dissertation materials, as a key motivational factor for doctoral completion and well-being (De Clercq et al., 2021; Devos et al., 2017; Milicev et al., 2021). Although initial evaluation results were positive, the variability of outcomes was high and, from the study's limited sample, it was unclear *for whom* the interventions worked effectively (or not).

Learning analytics (LA) have been proposed as an evidence-based approach to support learning in other areas of higher education (Leitner et al., 2017). Using data science and machine learning techniques, LA could help develop evidence-based (and potentially scalable) interventions to help understand the aforementioned differences in effectiveness of doctoral interventions for doctoral well-being. This in turn could lead to personalization of such interventions for higher effectiveness.

The overall goal of this paper is to explore the potential of a specific LA approach we denominate 'single-case learning analytics', to help practitioners to understand technology-enhanced training interventions' impact on doctoral student well-being. Such an approach could both help practitioners improve intervention design and provide students with practice-oriented and evidence-based insights they could use in their learning practice. All this, under the constraints typical of DE practice (i.e., small sample sizes, high heterogeneity, etc.).

1.1. Well-being in doctoral education

Studies on doctoral student well-being suggest that this collective shows high prevalence of emotional well-being issues (e.g., anxiety, depression) (Guo et al., 2021; Levecque et al., 2017). DE research has also uncovered factors related to low well-being: organizational support, scholarships/funding, interpersonal factors (e.g., relationship with supervisor), personal circumstances, or individual characteristics (e.g., motivation, self-efficacy) and behaviors (e.g., sleep, exercise, time management) (Byrom et al., 2020; Jackman et al., 2022; Levecque et al., 2017; Mackie & Bates, 2019; Pyhältö et al., 2012; Schmidt & Hansson, 2018; Sverdlik et al., 2018).

From this wealth of research on doctoral well-being and persistence, we should note that most of the aforementioned factors are out of the control of individual students, while others, such as the students' motivational aspects, beliefs, and behaviors, seem more amenable to intervention. Indeed, there exists a strand of research highlighting the importance of motivational factors, such as the students' perception of progress or the appropriation of the thesis project, in the development of well-being issues and the overlapping problem of high dropout rates in doctoral degrees (De Clercq et al., 2021; Devos et al., 2017; Milicev et al., 2021).

Reviews of DE research have noted the dearth of evidence-based interventions to address the problems of doctoral well-being (Jackman et al., 2022; Mackie & Bates, 2019). This may be, in part, due to the doctorate's inherent uniqueness, which leads to a heterogeneity of processes and outcomes (Hish et al., 2020). This heterogeneity calls for different approaches to evaluating and designing interventions in DE in an evidence-based manner. While novel technologies like LA could help in gathering data about, and modeling, this heterogeneity, initial LA studies in this direction have obtained mixed results (Di Mitri et al., 2017).

1.2. Person-centric and idiographic learning analytics methods

The complexity of factors related to DE well-being, along with inherent contextual challenges of this area (e.g., heterogeneity), suggest that a classic cohort-based, variable-centered approach to researching interventions will be limited (cf. the promising but underpowered intervention results by Prieto et al., 2022; Barry et al., 2019). As an alternative, person-centered approaches (Howard & Hoffman, 2018) try to determine subgroups within the population, thus preserving part of its heterogeneity. However, such grouping or clustering approaches, e.g., topological data analysis (Godwin et al., 2021), or latent profile analysis (Oberski, 2016) normally require large sample sizes (>200–500, see Howard & Hoffman, 2018). Further, such clustering methods often fail to provide meaningful practical insights, as they divide subjects into coarse groups that are low/high in certain behaviors, relative to the cohort. This can indeed be problematic in highly heterogeneous settings like doctoral education where meaningful comparisons between learners/processes often cannot be drawn.

At an even finer degree of granularity, idiographic approaches try to understand an individual over time (i.e., single-person dynamics – which is what practitioners actually need to deal with), not positing the universality of those dynamics (Piccirillo & Rodebaugh, 2019; cf. Shaffer & Serlin, 2004). Different data analysis methods have been proposed to aid idiographic research. For instance, Gaussian Graphical Models (GGMs) (Epskamp et al., 2018) or Group Iterative Multiple Model Estimation (GIMME) (Gates & Molenaar, 2012) try to understand patterns and trends in an individual. One restriction these idiographic models share is that they normally require a long series of observations ($k = 60$–100) (Gates, 2024). Methods coming from complex systems research (the study of systems in which distinct parts interact in ways that produce patterns that can't be understood nor predicted by studying the parts independently) have also been proposed to study individuals as nonlinear dynamical systems using methods like, e.g., recurrence quantification analysis (Webber Jr & Zbilut, 2005). Yet, these complex system methods also require long sequences of observations to be reliable.

Aside from quantitative methods, we should remember that the in-depth study of individual cases has long been a staple of qualitative research, under the term "case study" (Stake, 2005). In this paradigm, multiple (often, unstructured) sources of data are collected and then analyzed using methods such as qualitative content or thematic analysis. Here again, individual cases are often analyzed in isolation and no generalizability of results is claimed, making these methods idiographic in nature. While this qualitative approach can be extremely enlightening (and is often the prelude to quantitative studies), data analysis is laborious and often fails to support in a timely fashion the students/practitioners from whom the data was gathered. In recent years, mixed approaches combining qualitative coding and statistical modeling have emerged, e.g., epistemic network analysis (Shaffer, 2017). These methods typically rely on human-based qualitative coding thus sharing the labor-intensity weakness of qualitative case studies). However, recent approaches to automated coding of unstructured data (Cai et al., 2019; Garg et al., 2024; Hou et al., 2024; Pishtari et al., 2022) are making it feasible to integrate qualitative analyses in learning analytics pipeline and provision of timely feedback.

Idiographic approaches are indeed starting to be integrated in the technological support of learning. For instance, Saqr (2024) used GGMs to compare group-level and within-person patterns of engagement behavior in online learning, concluding that the two differ substantially and that "generalizing group-level findings to individuals may not be warranted" (p. 9). Other researchers combined longitudinal quantitative measures and indicators extracted from qualitative coding of unstructured data (e.g., diaries) to model a single person's contextual influences and patterns in lifelong learning, in what they denominated "single-case learning analytics" (SCLA) (Prieto et al., 2021). This approach is the focus of the current paper and is defined and described in more detail in a separate section.

1.3. Research (and practice) gaps and purpose of the present study

Despite numerous calls for evidence-based educational practice (Beerkens, 2018; Pellegrini & Vivanet, 2021), the research-practice gap in higher education is well documented (McIntyre, 2005; Vanderlinde & Van Braak, 2010), and even more salient in DE, in part due to its contextual constraints. Doctoral trainings, normally promoted by the local doctoral school (or program), tend to be non-mandatory parts of the doctoral degree and, being highly specialized, are taught to small and highly heterogeneous student groups (e.g., different backgrounds and programs/disciplines), for a short duration (e.g., a few weeks, not a full semester). All these aspects hamper practitioners in evidence gathering, analysis, and data-informed decision-making about the training's design.

These contextual constraints and the uniqueness of the doctoral journey suggest that an idiographic approach to understand doctoral learning experiences may be warranted. However, most idiographic analysis approaches (like GGMs or RQA) require long time series of observations. Given the scarce use of learning management systems and other typical sources of fine-grained temporal data, such long time series data are rare in DE. What evidence-based methods could a doctoral trainer/practitioner (and students) then use, under the constraints of authentic DE practice? We suggest that a variant of the aforementioned "single-case learning analytics" (SCLA) approach could be well-suited for DE settings. Our overarching research question (RQ) thus is: *What can educational stakeholders learn through the application of an SCLA approach during DE interventions?*

The following section describes in detail the main methodological contribution of this paper: how the SCLA approach can be operationalized under the practical constraints of doctoral education. To illustrate the usefulness of this approach, we will then explore the research question above through an illustrative study set in an authentic DE setting: a series of workshops to support doctoral well-being.

2. Single-case learning analytics: operationalization for doctoral training

Heterogeneity is a key feature of doctoral students, not only because of their uniqueness as individuals and learners when approaching the doctorate (often, with disparate backgrounds even in the same lab or research group), but also in the topical focus of their studies (being unique by definition), their learning processes, and the point in that learning journey in which each student is when entering a doctoral training action. This wide heterogeneity poses difficulties when trying to model students and their learning processes throughout a doctoral training quantitatively, using typical (group-level) methods.

A similar kind of heterogeneity, found in modeling lifelong learning, led researchers to propose "single-case learning analytics" (Prieto et al., 2021), as: "the use of computational means to collect, analyze and report data about a single learning entity and its context over long periods of time, for purposes of understanding and optimizing such learning and the environment(s) in which it occurs, without necessarily comparing it with other learning processes". Following this definition, we can deduce that SCLA is idiographic in nature, and that it does not seek to generalize its results to the individual's whole population.

Several key methodological features can be derived when applying this flavor of learning analytics: a focus on longitudinal (i.e., time series) data; the tracking of multiple contextual variables to understand the situational uniqueness of each learner; the combination of quantitative/structured data and qualitative/unstructured data in an embedded mixed-methods process (Creswell, 2009); and enlisting the collaboration of students themselves in the definition of quantitative variables and the gathering of qualitative evidence (e.g., through observation, reflections, or diaries). A focus on helping stakeholders in practice (especially, each individual learner) would logically lead to developing idiographic models (i.e., using the data from a single learner for modeling), and can also influence the choice of models to favor more interpretable ones (e.g., linear models or decision trees, as opposed to "black box" models), since they will likely be shown to learners in some form or another.

When operationalizing these ideas in the context of doctoral trainings, we encounter additional challenges, as noted above: the comparatively short duration of trainings leads to shorter time series (and/or requires data sources of finer granularity, like logs or physiological data), and the lack of a centralized learning management system for the doctoral activities makes it harder to gather longitudinal data about learning-related activities. Further, the fact that these trainings are often not mandatory, may lead to missing data as students tend to skip parts of the courses.

These challenges in turn shape a number of decisions when implementing SCLA systems in DE. First and foremost, decisions related to the data collection, in particular, the gathering of longitudinal data (e.g., daily) of both quantitative indicators relevant to the learning phenomenon at hand (in the case of the widespread doctoral problems mentioned at the outset of this paper, data related to students' perception of progress or well-being) and unstructured data (e.g., narrative of reflective diary). These quantitative variables should vary meaningfully at the scale of the training and the data gathering (e.g., not be traits or stable beliefs). Then, a qualitative content analysis can be performed on the unstructured data to extract further contextual predictors at each time point (e.g., each day).

Both of these sets of predictors then serve as inputs for building a series of idiographic models (as interpretable statistical or machine learning models) for each individual student. We should note that these idiographic models are intended as exploratory tools rather than inferential ones, aimed to uncover potential underlying structures in each student's limited data. Thus, they should be understood as "hypothesis generating tools" (Ludwig & Mullainathan, 2024) for students to explore in their everyday work, not for hypothesis testing or generalization to a larger population. Even with this exploratory aim, we suggest triangulating multiple very different (e.g., linear and non-linear) models (cf. "method triangulation" Denzin, 2017; Fisher et al., 2019) as each model individually is bound to be a "weak predictor" due to the limited amount of data. Ideally, these models should be sequence-aware (as most idiographic analyses, e.g., GGMs, are), but the length restrictions of many doctoral trainings may impede the application of such models. In that case, simpler models that do not account for sequence in the temporal data (e.g., a simple regression tree) may still give (limited) information about "intraindividual variation, that is, variation within a person over time" (Beltz et al., 2016) – that is, how some days and their characteristics, as recorded by the mixed-data diaries, are different from others. The decision of which kinds of statistical/machine learning models to develop in this way should be primarily guided by their interpretability, given that SCLA here is practice-oriented, intended to be understood and put to use by non-experts (e.g., doctoral students or instructors). Further, it is critical in this idiographic model building process to pay special attention to the features that go into the models. This includes, for example, the use of models with some form of feature selection as the number of predictors (including both the quantitative and qualitative-extracted ones) may be comparable to or larger than the number of data/time points. By using theory-based quantitative predictors and variables selected by students themselves (as "experts in their own situation") via diary narratives and qualitative coding, we can also maximize the relevant information in each data point to enable meaningful exploratory analyses despite small sample sizes (Brooks & Ruengvirayudh, 2016).

Finally, from the triangulated interpretation of these exploratory models (e.g., which predictors seem related to the daily progress of a student), practical personalized insights may be extracted (e.g., behaviors, strategies or contextual factors for the student to investigate in their work in terms of emphasizing or avoiding them). Aside from this immediate usefulness of such idiographic models, the parameters and shape of the idiographic models (e.g., which variables seem like more stable and stronger are significant predictors for a certain student, and in which direction, cf. Fisher et al., 2019), could also be potentially useful for modeling at the group level (as we show in the illustrative empirical study below). This potential for idiographic models to inform nomothetic analyses is *not* the main aim of SCLA, but rather an interesting secondary added value. This operationalization of SCLA is summarized graphically in Fig. 1 (using the example of triangulating linear and tree regression models, see the case study in the following sections).

3. Illustrative empirical study: context and methods

Goal of the illustrative study. To understand the added value of applying SCLA for doctoral practitioners/students in DE interventions, we need to consider both an idiographic perspective (what can SCLA show about each individual student – which is relevant for students and for instructors advising those students individually) and a more nomothetic one (can insights from SCLA enhance group-level analyses of the DE interventions – relevant for instructors). It is worth noting that even this latter perspective is aiming for insights and feedback that apply to the particular group of doctoral students, not generalizable inferences about the doctoral population (cf. Shaffer & Serlin, 2004). Following this rationale, we will illustrate the added value of SCLA through an empirical study of three research questions that combine nomothetic (analyzing group-level, average effects of the interventions) and idiographic elements (inspired by Bergman, 1998; see also Raufelder et al., 2013).

Illustrative study research questions. First, we take a student-centered and idiographic perspective to understand *what kind of actionable insights can students obtain from the application of SCLA?* (RQ1). Next, taking the instructors' point of view, we investigate the nomothetic (group-level) question of *what can instructors learn from applying SCLA, about the effects of their training actions?* (RQ2). As noted at the end of the previous section, this second question can be informed by the idiographic results of RQ1, and can itself be decomposed into sub-questions, such as: *were the workshops generally effective?* (RQ2.1) and *who benefited more/less from the workshops?* (RQ2.2). Finally, as LA researchers, we also pose the (nomothetic) question: *do (process-oriented) SCLA-based indicators provide added value over other (pre-test) quantitative indicators?* (RQ3). Of these questions, RQ1 is the most important, being learner-centred (Gašević et al., 2015) and idiographic in nature. RQ2.1, RQ2.2 and RQ3 rather aim to illustrate how SCLA can help other educational stakeholders and inform nomothetic analyses, thus representing a secondary added value of the approach.

Study design and datasets. We performed an embedded mixed-methods case study (Creswell, 2009) to answer these questions, based on two datasets gathered in authentic DE settings. The data come from two series of doctoral workshops to improve doctoral student well-being by better perceiving and making progress, which took place across two different universities in Spain and Estonia. While the two datasets come from workshops using the same learning activities and structure, they investigated well-being using different constructs coming from contrasting psychological approaches (see "Data collection and instruments" below), to increase the reliability of our findings. The first dataset (workshops W1-W3) had already been used in a prior publication focusing on the workshop's iterative design and initial (nomothetic) evaluation of average effects (what we labeled RQ2.1 above, see Prieto et al., 2022); here, however, we focus on the novel idiographic method to analyze such data, expanding the range of research questions and complementing the answers with a second, entirely new, dataset.

3.1. Context

Educational context. The doctoral workshops took place in two public mid-size universities in Estonia (2 workshops) and Spain (4 workshops), organized in collaboration with the doctoral schools of both universities. In Spain and Estonia doctoral schools are becoming aware of the widespread well-being challenges of their students and it is now commonplace for universities in these countries to offer at least one well-being-oriented course within their roster of transversal training offerings. The workshops, promoted within each university's doctoral school training catalogue, were open to students at any stage of their PhD, preferably in their second and later years of PhD. The workshops took place between 2020 and 2023, five of them in online format (mainly due to COVID-19 pandemic-related restrictions) and one face-to-face. Table 1 provides further contextual details.

Workshop intervention design and participants. Taking into account the intertwined relationship of "soft" aspects and skills (e.g., motivation, well-being, burnout) and "hard" skills (e.g., productivity, time management habits) that lead to better progress, the workshops adopted a dual, integrative approach, with a module on emotional well-being in the doctorate and strategies to foster it, and another module on productivity and time management habits; a third module focused on

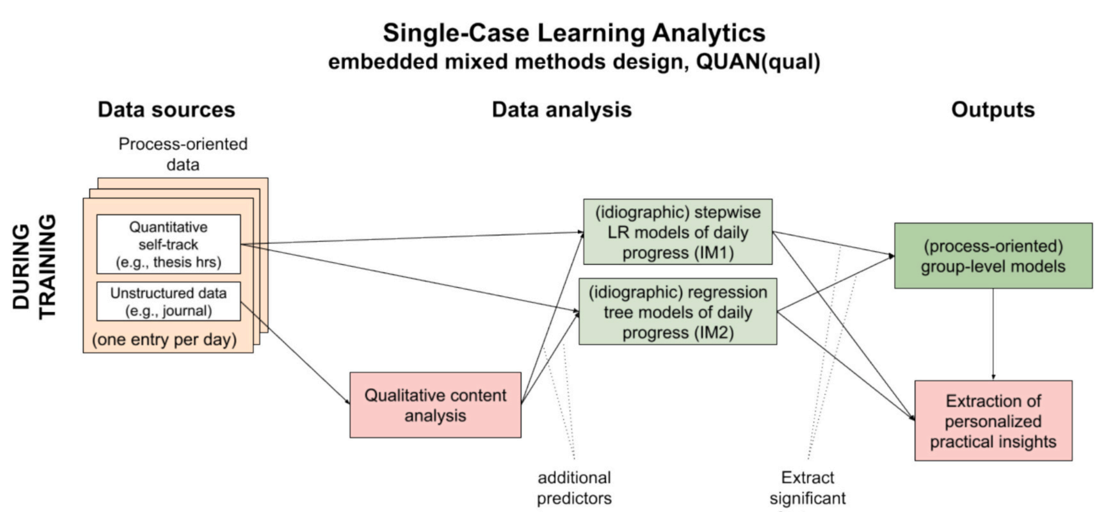

Fig. 1. Schematic operationalization of single-case learning analytics in a doctoral training.

Table 1
Workshop interventions included in the study and key contextual characteristics.

Dataset	Workshop	Semester	Country	Discipline	Format	Nr. students
	W1	Spring 2020	Spain	Health Sci.	Online	15
1	W2	Fall 2020	Spain	ALL	Online	20
	W3	Spring 2021	Estonia	ALL	Online	22
	W4	Spring 2022	Spain	ALL	Online	21
2	W5	Fall 2022	Estonia	ALL	Online	9
	W6	Spring 2023	Spain	ALL	Face-to-face	18

making and perceiving progress, as the key aspect linking "hard" and "soft" skills and leading to completion of the dissertation (De Clercq et al., 2021; Devos et al., 2017). Doctoral students were engaged in these topics through readings, reflection, discussions and targeted practices (e.g., journaling, see below). Each workshop spanned four 2-h synchronous sessions over four weeks, with students working on individual asynchronous activities in between the synchronous sessions. Further details about the workshops are available elsewhere (Prieto et al., 2022). The six workshops were attended by a total of $N = 105$ doctoral students (see Table 1). The doctoral schools awarded a certificate of attendance to participants, but the workshops were not part of the requirements to complete the doctoral degree. Please note how the workshops and the data exemplify the aforementioned constraints typical of DE training.

Intervention implementation details. The online workshops were held through video-conferencing, with communication with the participants handled via email. In both formats, students were asked to answer pre-post questionnaires, which were voluntary exercises within the workshops (aggregated results from them were used in the workshop activities), and thus not all participants chose to provide both pre- and post-questionnaire responses. To enable a journaling and self-tracking practice throughout the workshops (to promote a sense of progress, see Amabile & Kramer, 2011), students were offered a simple web-based questionnaire as quantitative self-tracker and journal (see below for details about the questions). All questionnaire responses (in the diaries and the pre-post data, described below) were anonymous, indexed to a student-chosen non-identifiable nickname, to foster more honest and meaningful responses (due to the sensitive nature of the workshops' topic). All participants in the workshops gave voluntary and informed consent to participate. The procedures for the data collection were approved by the CEITER project's Ethics Committee (i.e., the local Institutional Review Board – IRB, review cases #0002 and #0003).

3.2. Data collection and instruments

Data collection strategy. To account for the fact that doctoral progress, including issues of time management and productivity, is closely intertwined with well-being in the doctorate (Devos et al., 2017), we gathered data not only about doctoral students' well-being but also about progress-related aspects. The six workshops had similar data collection structures, but each workshop series/dataset used different well-being outcome measures and pre-workshop demographics/profiling questions.

3.2.1. Outcome measures (pre-post)

Well-being related outcome measures were: a) *positive psychological capital* (PsyCap, in dataset 1), a composite of hope, optimism, resilience and self-efficacy (Luthans et al., 2007) stemming from positive psychology, measured via the Compound PsyCap Scale (CPC-12) (Lorenz et al., 2016); and b) *general psychological distress symptoms*, i.e., of anxiety, depression and stress (dataset 2) – for brevity, referred to simply as distress symptoms below. These were measured via the Depression Anxiety Stress Scales (DASS-21, Lovibond & Lovibond, 1995), stemming from the classical view of well-being measured through the absence of negative symptoms. The shift in outcome measures from one dataset to the other was both to explore different perspectives on well-being, and due to our initial research efforts trying to measure benefits in a construct that is reportedly more prone to change (PsyCap, see Barry et al., 2019) and later attempting to understand the effects on a less modifiable well-being construct (symptoms), while maintaining the measurement effort approachable (given that it was a voluntary activity for participants).

3.2.2. Covariates (pre-test)

We also gathered covariates related to dropout and well-being in doctoral education literature, such as burnout (Cornér et al., 2017), experiential avoidance (Bond et al., 2011), or motivational aspects (perceived progress in dissertation materials, appropriation of one's thesis) (De Clercq et al., 2021; Devos et al., 2017). In dataset 2, we additionally gathered demographic data like year into the PhD, or full/part-time status (demographics were kept at a minimum to avoid inadvertently de-anonymizing the data). Table 2 provides a list of the pre-post variables used (see also Appendix B for the full description of the items that compose each of these variables/scales).

3.2.3. Process-oriented data (for the idiographic SCLA analyses)

A key aspect of the SCLA approach (and the workshops' experience) was to enlist the collaboration of students as data collectors and in selecting relevant variables, through the gathering of longitudinal qualitative data, to be later used for individualized (i.e., idiographic) modeling of each student. We used one of the workshop exercises (keeping a daily journal and self-tracking practice) as the main source of such process/experience data. A simple web-based questionnaire (with open and closed questions) asked students to answer, for each day, a small set of questions (see Table 2 and Appendix B), chosen to illustrate issues of time management, progress, self-care and exhaustion, in accordance to prior studies on doctoral well-being and dropout (De Clercq et al., 2021; Devos et al., 2017; Pyhältö et al., 2012). To encourage doctoral autonomy (core to the ethos of the doctorate, see Overall et al., 2011), students could also choose to follow journaling/self-tracking using other means. Yet, to incentivize the use of the provided data collection instrument (and later reflection by students on the data gathered), the workshops included an exercise in which students could explore their own data and the aggregate for the whole class, in an LA web-based dashboard (see Fig. 2).

3.3. Data analysis

3.3.1. SCLA analyses (pre-processing)

Qualitative analysis. We performed idiographic analyses of the process-oriented data described above, instantiating the SCLA process described in the "Single-Case Learning Analytics…" section. As a key feature of this embedded mixed methods approach (in which students themselves act as collaborators in defining relevant variables, via their noticing and reporting of events narratively), we used the unstructured (textual) data in the narrative journals as a source for personalized contextual and process predictors. We analyzed the diaries' narrative accounts using a deductive qualitative content analysis process, based on the Activity-Centered Analysis and Design (ACAD) framework (Goodyear et al., 2021). ACAD was originally developed to analyze emergent learning experiences in relation to a pedagogical design. Given

Table 2
Pre- and post-test variables as well as diary variables available in the different datasets.

	Dataset 1		Dataset 2	
	Construct	Instrument	Construct	Instrument
Outcome (pre-post)	Well-being * (PsyCap)	CPC-12 (Lorenz et al., 2016)	Well-being * (symptoms)	DASS-21 (Lovibond & Lovibond, 1995)
Covariates (pre-test)	Doctoral Burnout	8 items, taken from (Cornér et al., 2017)	Avoidance	AAQ-2 (Bond et al., 2011)
	Progress	1 item (Likert 1–5), inspired by (Devos et al., 2017)	Progress	3 items (Likert 1–5), inspired by (Devos et al., 2017)
	Dropout ideation	1 item (binary)	Appropriation	5 items (Likert 1–5), inspired by (Devos et al., 2017)
			Dropout ideation	5 items (binary), inspired by (Beck & Steer, 1993)
Demo-graphics (pre-test)			Year in the PhD	Scale (1−10)
			Discipline	Multi-choice (5 opts)
			Type of funding	Multi-choice (5 opts)
			Full-/Part-time status	Multi-choice (2 opts)
Process (diary)	Progress (that day)	1 item (Likert 1–7)	Progress (that day)	1 item (Likert 1–7)
	Time slept on the previous night	Continuous (nr. of hours)	Time slept on the previous night	Continuous (nr. of hours)
	Time spent on thesis-related tasks	Continuous (nr. of hours)	Time spent on thesis-related tasks	Continuous (nr. of hours)
	Time spent on non-thesis related tasks	Continuous (nr. of hours)	Time spent on non-thesis related tasks	Continuous (nr. of hours)
	Overall time spent working	Continuous (nr. of hours)	Overall time spent working	Continuous (nr. of hours)
	Narrative account of the day	Unstructured (open text)	Narrative account of the day	Unstructured (open text)

that, in the doctoral case, there is no explicit pedagogical design (students normally choose activities autonomously), the framework was used to map the nature of their *activities* (e.g., reading papers vs. doing a data analysis) and their social and physical context (i.e., *people* they were with, and *places* where the activity occurred). The qualitative analysis was performed by two members of the research team, who engaged in several iterations of coding small portions of the dataset and resolving disagreements until a common understanding of the codes was achieved. Then, each diary entry in the rest of the dataset was coded by one researcher. Inter-rater reliability was evaluated on 10 % of the dataset, achieving a substantial agreement (overall Cohen's K = 0.67; median single-code agreement K = 0.71), which was deemed sufficient for the purposes of this illustrative study. These codes, assigned to each diary entry (the unit of analysis), were structured into binary variables denoting whether a code (e.g., performing the activity of writing) was present or not that day. This gave a total of 32 binary code-related variables tied to each journal/self-track entry (3 for place-related codes, 9 for people-related codes, and 20 for activity-related codes).

Idiographic model building. This qualitative analysis process led to a qualitative-enhanced diary/self-tracking dataset which included both quantitative self-tracking variables (see Table 2) and additional variables coming from the qualitative analysis of the narrative entries. Then, to understand the individual patterns of progress and extract process/experience indicators, within the small sample size and limited length of time series (each student could provide at most 21 diary/self-track entries during the workshops), we triangulated two simple and highly-explainable (idiographic) statistical models of progress trying to predict the same proximal outcome (the satisfaction with progress on a given day, given the diary/self-track variables that day):

1. An idiographic stepwise linear regression model of each person's progress as a function of their other diary (i.e., self-tracking and journaling) variables (IM1). This process tried to balance model accuracy with simplicity (via the model's adjusted R-squared) – critical since each individual's sub-dataset often included more variables than data points. Linear regression was chosen over ordinal logistics regression due to the small sample size and easier interpretability (relevant for later use by instructors/students). These linear models were tested for multicollinearity using Variance Inflation Factors (VIF) for all predictor variables. Since the data forms a time series, these models were also tested for independence of residuals via the Durbin-Watson test.
2. An idiographic regression tree model of each person's progress as a function of the other self-tracking/journaling variables (IM2). The recursive partitioning algorithm automatically selects the variable on which a binary split most accurately predicts the target variable, and then iteratively builds a tree of such splits to predict the outcome (thus naturally performing variable selection). Given the small sample/time series size, we set a minimum of 4 observations in a node before split, a minimum of 2 observations in any terminal node, and a complexity parameter of 0.01. The trees were pruned by leave-one-out cross-validation to determine the optimal complexity parameter (CP). Each observation was iteratively held out as a validation set, with the model trained on the remaining n-1 observations. The CP value that minimized the cross-validated prediction error was selected, resulting in the final tree. Again, these models were also tested for independence of residuals via the Durbin-Watson test.

Idiographic model evaluation. Both kinds of idiographic models (IM1 and IM2 above) were evaluated using R-squared and root mean squared error (RMSE) in leave-one-out cross validation (LOOCV, chosen due to the small sample size of each idiographic model). Our idiographic models were developed only for those students who provided at least 5 diary entries. This (extremely small) lower bound was chosen to provide potentially interesting exploratory information to doctoral students as soon as possible (as an incentive for further data entry), but it also brings model reliability and stability issues (as the number of samples sometimes falls below heuristics such as having two samples per predictor, Austin & Steyerberg, 2015). This is noted explicitly as warnings in the idiographic models of Appendix C. While the triangulation of linear/nonlinear models and the use of theory-based and expert-selected variables somewhat mitigate these problems, these idiographic models need to be interpreted (and communicated to students/instructors) with extreme care, not as frequentist inference tools but rather as exploratory ones for mining hypotheses to be later tested by students in their daily practice.

3.3.2. Answering the study research questions

The data analysis to answer our three research questions remained similar across both datasets, albeit using different outcome variables and covariates (see Table 2) where appropriate.

Student-oriented (idiographic) data analysis (RQ1). To understand

Fig. 2. Example screenshots of the LA dashboard to analyze journal/self-track data, available to students in the workshops. *Note*. Temporal trend and weekly distribution of students' perception of progress (top), representation of coefficients and standard errors of a group-level linear regression of progress (middle), and idiographic regression tree of progress for one of the students, based on her self-track data (bottom).

what kind of actionable insights students can take from the application of SCLA, we extracted insights directly from the idiographic models of each student (IM1 and IM2, see "Pre-processing" above) and what could be their implications for student daily work practices. To understand the heterogeneity of these idiographic models, we also report descriptive statistics about the key features of these models (e.g., predictors with a p-value<0.05 in the IM1/linear models, first splits in the IM2/tree models).

Instructor-oriented (nomothetic) analysis of average intervention effects (RQ2.1). Then, to determine what instructors can learn from applying SCLA, about the effects of their training actions, we resorted to nomothetic/group-level analyses (informed by idiographic ones). We first looked at the general effectiveness of the workshops (RQ2.1). We did so through a paired t-test on the group-level difference in pre-post well-being values (i.e., to determine whether it is significantly different from zero), and calculated the effect size of the intervention via Cohen's d statistic. The Shapiro-Wilk normality test was used to confirm that the outcome variable of each dataset (PsyCap or DASS symptoms) was normally distributed. To understand whether there are clustering effects (and the corresponding group variance) in terms of workshop groups which would warrant a more complex multilevel modeling, we built linear mixed-effects models with workshop as the random (grouping) effect. However, very small intra-class correlations obtained for mixed effect models for both datasets (ICC = 0.02 and ICC = 0.01) showed that multilevel modeling was not needed. Thus, we include these models only in Appendix C.

Instructor-oriented (nomothetic, idiographic-informed) analysis of differential intervention effects (RQ2.2). Afterwards, to start disentangling who the interventions benefited more/less (RQ2.2) we used exploratory multivariate linear regression models to predict pre-post differences in well-being outcomes. Such model choice was due to the limited sample size and the models' relative simplicity and explainability (to enable eventual communication to instructors and students). A linear regression was chosen over an ordinal logistics regression due to the underlying assumption of continuous outcome variables (as they are calculated as averages/sums of Likert-scale items) (Norman, 2010), but also due to the comparatively small sample size. Several sets of variables were used as predictors in this linear regression (GM1): 1) the person's initial well-being score; 2) the other theory-backed and demographic pre-test covariates (see Table 2); and 3) indicators about the process/experience of students extracted from the idiographic modeling of the longitudinal (diary/self-track) data (see pre-processing step above, IM1 and IM2). Specifically, the latter group of predictors included:

1. Whether an indicator (e.g., time dedicated to the thesis) was a significant predictor of progress (taking the p-value<0.05 as a mere exploratory marker suggesting a potentially strong predictor, not in its original inferential sense), and the direction of effect, in the idiographic linear regression of progress (IM1) of a student. This led to a set of binary/bipolar variables (e.g., lm.Thesis = +1 if the time dedicated to the thesis was a potentially "significant" positive predictor of daily progress in that person's idiographic model, lm.Thesis = −1 if it was potentially "significant" negative, lm.Thesis = 0 otherwise).
2. A categorical variable indicating which variable was most important (i.e., the top-most split) in the individualized regression tree model (IM2) of a student, and the direction of relationship (positively or negatively related to progress). For instance, if the time dedicated to the thesis was the top-most split and larger values of it led to higher values of progress, the variable tree.firstsplit would be assigned the category "Thesis+".

Taking into account the large number of variables in the three sets of predictors above (compared to the limited sample size in each dataset), we performed a stepwise variable selection process to balance model accuracy with simplicity (via the model's adjusted R-squared). The result was our first group-level model of workshop benefits (GM1) indicating which factors (including idiographic ones) seem most related to benefiting more/less from the workshop. This linear model was tested for multicollinearity using Variance Inflation Factors (VIF) for all predictor variables. Instructor-oriented insights (e.g., about the design or student profiling of the workshop) were then extracted from the interpretation of this model.

Researcher-oriented (nomothetic, idiographic-informed) comparison of model performance (RQ3). Finally, we explored the researcher-oriented question of whether the process information extracted from the idiographic models (IM1 and IM2) would provide quantifiable added value over the traditional variable-centered approach (using just the pre-test covariates) in the (group-level) modeling of workshop benefits (RQ3). To understand what LA indicators better predicted the differential outcomes we developed three regression models: 1) A baseline group-level model of benefits (GM2) using the person's initial well-being score as predictor; 2) A variable-oriented group-level model (GM3) using initial well-being and other pre-test covariates (see Table 2); and 3) A more process-oriented group-level model (GM4) that used the indicators extracted from the idiographic modeling of the longitudinal (diary/self-track) data (IM1 and IM2). Again, the linear models were tested for multicollinearity using Variance Inflation Factors (VIF) for all predictor variables. These three models were compared using multiple performance metrics (root mean squared error, R-squared, adjusted R-squared, Akaike information criterion, and Bayesian information criterion), to understand if any one set of predictors showed consistent advantage.

Data analysis implementation details. The data analysis flow and relationships with the research questions are depicted in Fig. 3. All data analyses were performed using the R statistical software. We used the *leaps* package for stepwise linear regression modeling, *rpart* for regression trees, *performance* for model comparison and ICC calculations, *lme4* for mixed-effects linear models, and *rstatix* for Cohen's d and Shapiro-Wilk tests.

3.4. Dataset description

As noted above, two different datasets were used for this study, each one coming from three different doctoral workshops, and using different well-being outcome and pre-test variables (see Table 2). The full description of the dataset variables gathered is summarized in Table 3 below.

In dataset 1, a total of $N = 57$ doctoral students provided data, of which $N = 32$ completed both pre-post questionnaires. $N = 34$ students completed at least one journal entry (max = 19 entries, median = 8 entries, IQR = 9). $N = 26$ students provided 5+ entries and were thus used for building idiographic models of progress (median = 10.5 entries, IQR = 5). $N = 24$ students completed both pre-post and diary entries. The internal consistency reliability (Cronbach's alpha) was acceptable to good for PsyCap and questionable for burnout.

In dataset 2, a total of $N = 48$ doctoral students provided data, of which N = 32 completed both pre-post questionnaires. It is noteworthy that the average pre-test values for emotional distress symptoms (DASS) were considerably higher than those of non-clinical populations (Henry & Crawford, 2005). $N = 16$ students completed at least one journal entry (min = 2 entries, max = 20 entries, median = 13 entries, IQR = 3). The diaries of $N = 15$ students included 5+ entries and were used for idiographic modeling (median = 13 entries, IQR = 2.5). $N = 13$ students completed both pre-post and diary entries. The internal consistency reliability was excellent for distress symptoms (DASS) and avoidance, but it was questionable to acceptable for progress, appropriation and dropout ideation.

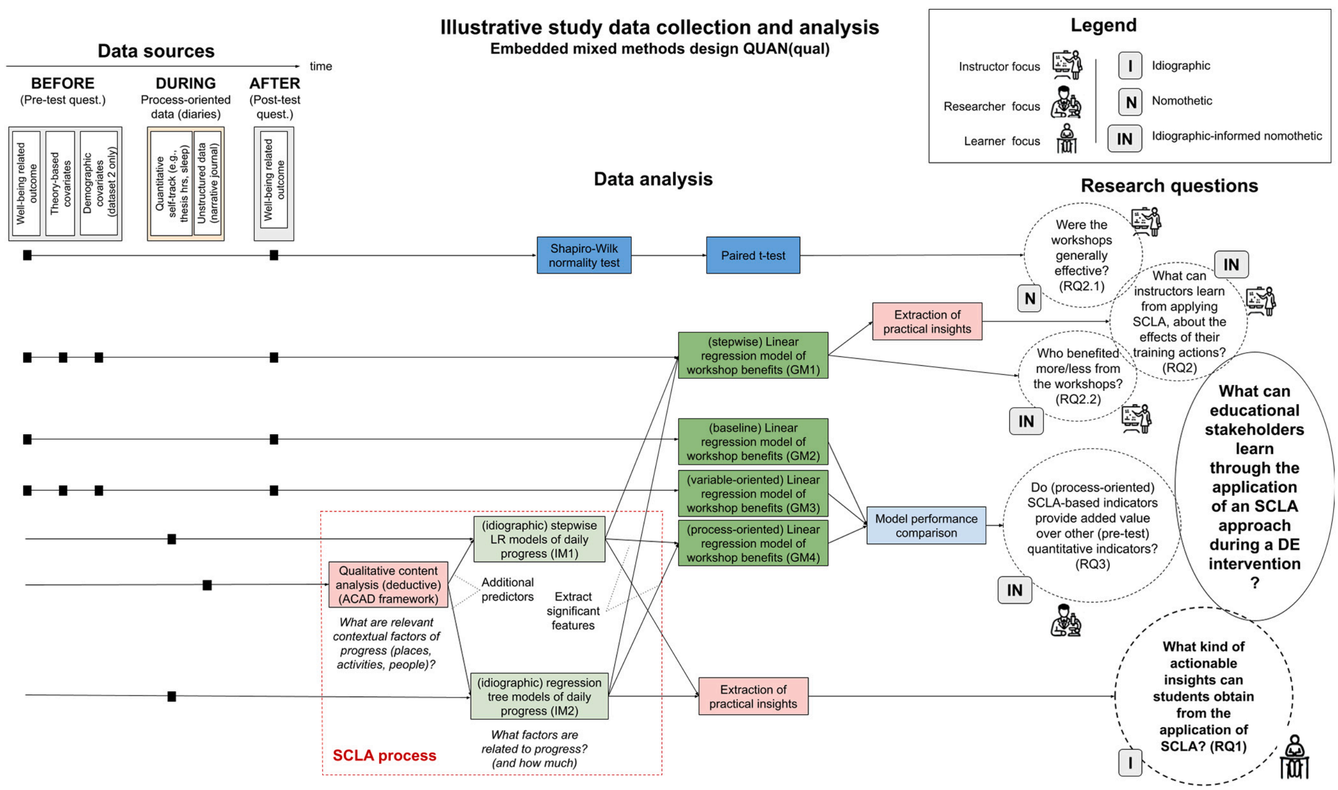

Fig. 3. Data analysis flow and relationship with the research questions of the study. *Note.* In blue, descriptive (light blue) and inferential (dark blue) statistics methods; in green, exploratory machine learning methods (light green for idiographic models, dark green for group-level ones); in red, manual processes.

Table 3
Description of the quantitative variables in the two datasets of the illustrative study.

variable	n	mean	sd	items	alpha	ci_lower	ci_upper
Dataset 1							
Pre-post variables							
PsyCap (pre)	47	4.44	0.62	12	0.79	0.72	0.87
PsyCap (post)	37	4.70	0.59	12	0.88	0.84	0.93
Progress	47	3.06	1.03	1	NA	NA	NA
Burnout	47	3.11	0.57	8	0.63	0.49	0.78
Dropout ideation	47	0.43	0.50	1	NA	NA	NA
Diary entries							
Progress	299	4.15	1.54	1	NA	NA	NA
Thesis (hrs)	299	3.55	2.58	1	NA	NA	NA
Work (hrs)	299	6.96	2.74	1	NA	NA	NA
Sleep (hrs)	298	7.31	1.07	1	NA	NA	NA
Dataset 2							
Pre-post variables							
DASS (pre)	46	17.20	14.56	21	0.95	0.93	0.97
DASS (post)	34	12.56	9.11	21	0.91	0.87	0.94
Avoidance	46	20.43	10.98	7	0.93	0.90	0.96
Progress	46	2.88	0.85	3	0.70	0.55	0.85
Appropriation	46	4.20	0.58	5	0.68	0.53	0.83
Dropout ideation	46	1.35	1.48	5	0.75	0.67	0.84
Diary entries							
Progress (that day)	214	4.20	1.85	1	NA	NA	NA
Thesis-related work (hrs)	201	3.87	2.96	1	NA	NA	NA
Total work (hrs)	214	7.92	3.09	1	NA	NA	NA
Sleep (hrs)	214	6.77	1.08	1	NA	NA	NA

While these sample sizes, internal consistencies, and missing data are sub-optimal for research purposes, they are quite typical of the voluntary, autonomy-oriented nature of doctoral-level training actions (i.e., what practitioners encounter in authentic DE settings), and thus considered appropriate for purposes of our illustrative study here.

4. Illustrative empirical study: results

4.1. RQ1: what kind of actionable insights can students take from the application of SCLA?

A total of 82 idiographic (linear and tree) models of progress (IM1, IM2) were built from each student's data with five or more diary entries in both datasets (n = 26 of each kind for dataset 1, n = 15 in dataset 2). Out of these idiographic models (see Appendix C), six did not pass the independence tests (Durbin-Watson test $p < 0.05$), and another two (linear) models showed signs of high collinearity (VIF > 10), and were thus discarded. Another model showed very high out-of-sample (cross-validation) RMSE (which implied large amounts of overfitting) and was discarded as well. The resulting set of $n = 73$ idiographic models were evaluated and examined for insights (see next).

To have a general understanding of the expected reliability and heterogeneity of these models, Fig. 4 below summarizes the performance metrics (out-of-sample R-squared and RMSE via leave-one-out cross-validation) of the models. The overall median performance was RMSE = 1.27 and R-squared = 0.20. Table 4 summarizes the structure of the linear and tree models, in terms of predictors with a p-value<0.05 and first tree splits (these will be later used as "process variables" in the group-level models of RQ2 and RQ3). We can see that in both datasets the distribution of strong model features is rather sparse and heterogeneous: the same predictors appear with different directions for different students, not being strong predictors for the majority of the students (e.g., mentions to organizing tasks in the diary). Still, some predictors (e.g., time spent working on the thesis on a day) are more often a strong predictor in the linear idiographic models (IM1) (in dataset 1, strongly positive for n = 8 students; in dataset 2, for one student). The first split of the idiographic tree models (IM2) is also heterogeneous, albeit the amount of time spent working on thesis materials seems to be the most common (for n = 13 models in dataset 1, $n = 11$ in dataset 2).

These diary-based exploratory idiographic models of progress could be a rich source of insights for individual practitioners/students. We could extract personalized contextual factors associated with each person's progress (itself linked to well-being and persistence in the doctorate).

For one of the participants in dataset 1 (P28, see Fig. 5, top-left and Table 5) the amount of time spent working on thesis materials, and mentions to learning, organizing, reading and writing activities all seemed to be strong and positive predictors of progress. The amount of sleep, however, was a negative predictor of progress. The tree model confirms the importance of time spent on thesis (especially, if >5 h that day) as a key predictor of progress while, e.g., the mentions to learning

Table 4
Summary of strong predictors in the linear models and first tree splits of the idiographic models.

Predictor (in linear model)	Dataset 1			Dataset 2		
	Strong, p < 0.05 (−)	$p >$ 0.05	Strong, p < 0.05 (+)	Strong, p < 0.05 (−)	p > 0.05	Strong, p < 0.05 (+)
Total work hrs.	3	16	7	0	14	1
Thesis-related hrs.	0	18	8	0	14	1
Sleep hrs.	3	18	5	0	15	0
Non-thesis work hrs.	0	26	0	1	14	0
Mentions of						
organizing activities	3	20	3	0	14	1
teaching activities	0	22	4	0	14	1
learning activities	1	21	4	0	14	1
reviewing (others')	3	22	1	0	14	1
data collection	2	20	4	0	15	0
reading activities	1	20	5	0	15	0
coworkers	2	24	0	0	15	0
data analysis	1	20	5	0	15	0
supervisors	3	22	1	0	15	0
email activities	1	23	2	0	15	0
meetings	1	24	1	0	15	0
writing activities	1	23	2	0	14	1
working at the lab	1	24	1	0	15	0
First tree split	Thesis hrs. (+):13 Total work hrs. (+): 5 Sleep hrs. (−): 2 Others (1 each): 6			Thesis hrs. (+): 11 Non-thesis work hrs. (−): 2 Others (1 each): 2		

Note. The numbers denote how many students have that predictor as strong or as first tree split.

activities sometimes lead to lower progress (depending on other parameters for that day, like time on thesis and open-text mentions to reviewing activities). In contrast, for another of the participants in dataset 1 (P31, see Fig. 5, top-right, and Table 5), the idiographic linear model suggested sleep, reading and working at the lab were associated with more progress; conversely, interactions with coworkers and learning seemed associated with less progress. The tree model for this

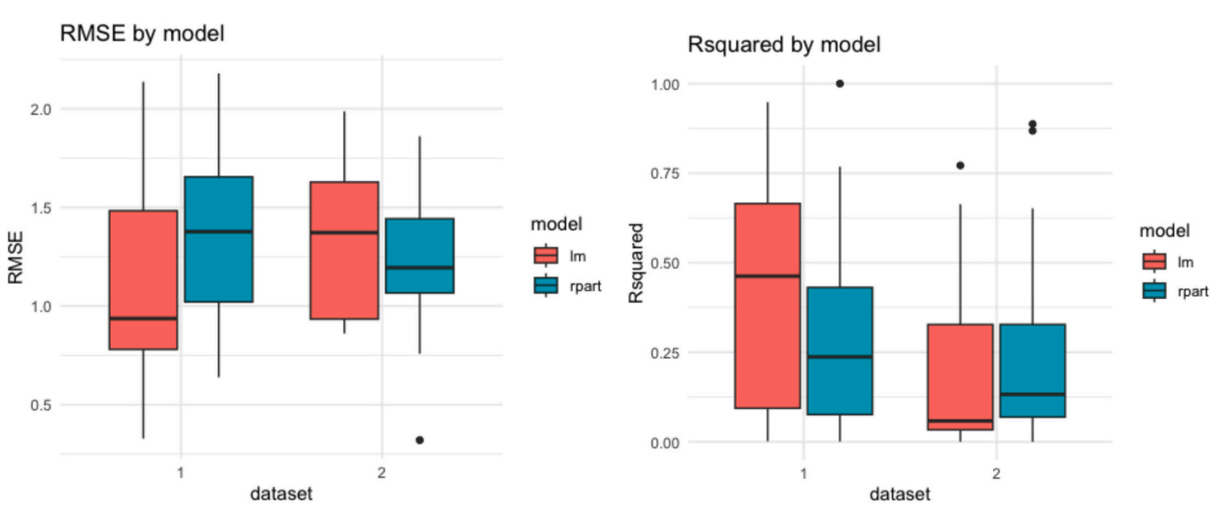

Fig. 4. Performance metrics of the idiographic models of doctoral student progress, measured by leave-one-out cross-validation.

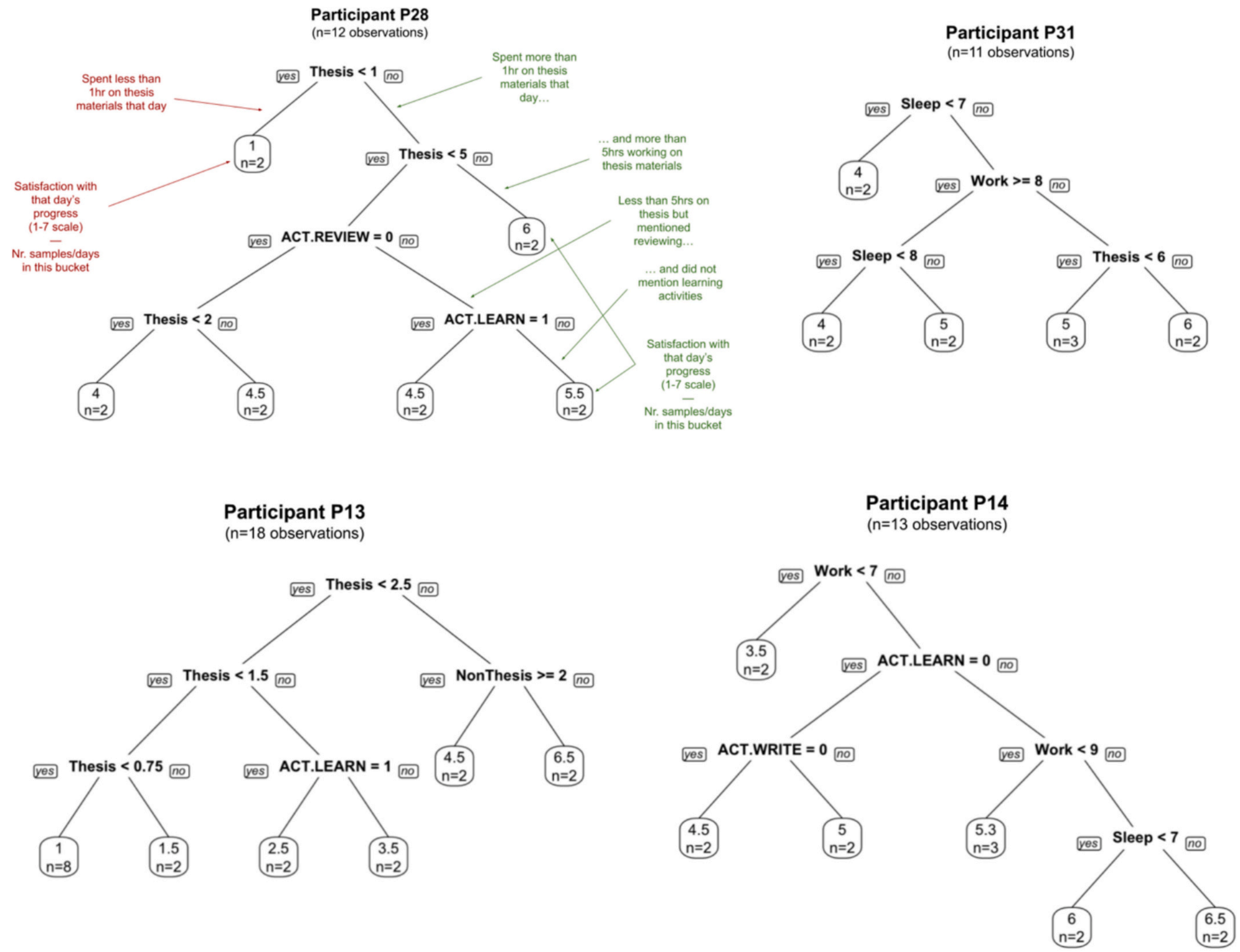

Fig. 5. Example idiographic regression tree models of progress. *Note.* Four doctoral students, two from dataset 1 (top, P28, P31) and another from dataset 2 (bottom, P13, P14).

person confirmed the role of sleep as important (as the top-most split) and suggested concrete thresholds for sleep (>7 h) and total work time (not >8 h) that seemed conducive to better progress. Personalized advice for these two students could focus on these factors or take the form of probing questions to understand deeper causes (why do you think sleep is negatively related to your progress? Do you have too much on your plate and only progress on the thesis at the cost of sleep? Are the learning activities mandatory? Is this a transitory phase in the doctorate? etc.).

From the idiographic, exploratory models of progress in dataset 2 we could also extract practice-oriented insights, by finding personalized contextual factors associated with each person's progress. For one participant in this dataset (P13), the linear regression (IM1, see Table 5) suggested that the time spent working on the thesis was a strong positive predictor, while mentions of reading activities seemed negatively related to progress (albeit non-significantly). The idiographic regression tree (IM2, see Fig. 5, bottom-left) for this person again pointed to the importance of the time spent working on the thesis (as the first and several other tree splits). It also suggested certain thresholds (two hours of non-thesis work tasks) as making a difference between mediocre and good progress in a day. Looking at another participant in this dataset (P14), the linear regression (IM1, see Table 5) suggested that overall time working was a key predictor of progress, but so was teaching, organizing, and writing activities. For this person, learning activities seemed positively associated with progress (opposite to the example P31 in dataset 1). The idiographic regression tree for this student (IM2, Fig. 5, bottom-right) supported the importance of overall work time as a positive predictor, providing certain thresholds that seemed to work best for this student. Mentions of learning activities seemed positively associated with this person's progress. Advice to these students could focus on doing more of the positive activities (e.g., learning activities, which appear in both the models of P14), but also probing questions about these models (why do you think you need 9+ hours of work to feel progress? Is it sustainable in the long-term? etc.). It is also worth noting that the evaluation of P31's linear model showed signs of overfitting (e.g., in terms of R-squared in the out-of-sample LOOCV) but the others generally seemed quite predictive of previously-unseen samples.

4.2. *RQ2: what can instructors learn from applying SCLA, about the effects of their training actions?*

4.2.1. *RQ2.1: Were the workshops generally effective? (RQ2.1)*

The baseline group-level effectiveness of the workshops to improve doctoral students' well-being was evaluated with classic (nomothetic) inferential statistics. In dataset 1 (which used psychological capital as the proxy for well-being), a Shapiro-Wilk test of normality ($W = 0.972$, $p = 0.54$) revealed that the outcome differences were normally distributed. A paired t-test of pre-post values suggested that the difference in

Table 5
Stepwise linear regression (idiographic) models of progress for four doctoral students.

	Dependent variable:			
	Progress (daily)			
Predictor/Participant ➔	P28	P31	P13	P14
Time spent in thesis-related activities	0.23 (0.07)**		1.13 (0.10)***	
Time slept the previous night	−3.54 (0.45)***	1.00 (0.16)***		
Overall time dedicated to work				0.32 (0.10)**
Mentions of learning activities that day	1.42 (0.26)***	−4.50 (0.54)***		1.21 (0.37)**
Mentions of organizing (materials, data) that day	1.03 (0.19)***			1.18 (0.38)**
Mentions of reading that day	1.00 (0.21)***	2.50 (0.42)***	−0.11 (0.53)	
Mentions to reviewing others' work that day				1.99 (0.71)**
Mentions of teaching activities that day				1.85 (0.39)***
Mentions to writing activities that day	1.72 (0.32)***			1.86 (0.37)***
Mentions of working at the lab		2.50 (0.54)***		
Mentions of coworkers that day		−1.50 (0.30)***		
(Intercept)	29.30 (3.48)***	−4.50 (1.36)**	0.92 (0.23)***	−0.09 (1.09)
Observations	12	11	18	13
(in-sample) R-squared	0.99	0.95	0.89	0.91
(out-sample/LOOCV) R-squared	0.63	0.02	0.77	0.47

Note. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Two students from dataset 1 (P28, P31) and another two from dataset 2 (P13, P14).

PsyCap (mean difference = +0.26) was statistically significant ($t = 4.96$, df = 31, $p < 0.001$). The effect size was $d = 0.88$ (considered "large"). However, 9 out of 32 participants saw zero or negative gains, and a similar proportion saw much larger gains. In dataset 2, where well-being was measured by the absence of emotional distress symptoms (using the DASS scale), again the Shapiro-Wilk test (W = 0.97, $p = 0.57$) and subsequent paired t-test of pre-post values of distress symptoms suggested that the difference (mean difference = −6.31) was statistically significant ($t = -3.15$, df = 31, $p = 0.004$). In this case the effect size was $d = 0.56$ (considered "moderate"). About a third of participants (11 out of 32) saw no change or increased symptoms and a similar proportion (12 out of 32) saw much larger benefits (e.g., twice the average). From the instructor's perspective, these results suggest that the workshops are a worthy pursuit. However, the aforementioned variability in outcomes from both datasets prompts the question of who benefited more/less from the interventions.

4.2.2. RQ2.2: who benefited more/less from the workshops?

A stepwise linear regression model using both pre-test variables and process-oriented variables (GM1) can be mined for insights about this question. In dataset 1 (see Table 6, left) we see that only one of the pre-test variables (initial self-perception of progress) was significantly related to larger workshop benefits (i.e., increase in PsyCap). We can also see quite a few of the process-oriented variables being significantly related to well-being benefits. For instance, the relationship of daily sleep and daily progress in the diaries (via the linear and tree idiographic models) appears in three of the model coefficients, in a consistent fashion: students that had a significant negative relationship between their daily sleep and progress, benefited more from the workshops. Certain kinds of doctoral processes/experiences, e.g., those in which a negative relationship was found between activities like email and progress, benefited more from the workshops. As another example, students who saw a positive relationship between their progress and certain work settings (e.g., working at the lab) also benefited more from the workshops. These patterns may point to the doctoral students finding out through the diary exercise certain ways of working that foster/hamper progress (which in turn could lead to exiting the workshops with higher psychological capital).

Table 6
Stepwise linear regression (nomothetic) models of differential well-being outcomes from the workshops, using pre-test and process/idiographic model predictors.

	Dependent variable:	
	Difference in PsyCap	Difference in DASS (symptoms)
Predictor	Dataset 1	Dataset 2
Perceived progress (pre-test)	0.14 (0.04)**	
DASS symptoms (pre-test)		−0.33 (0.14)*
Strong predictor (idiographic linear model, from diaries):		
Time spent on thesis tasks	−0.51 (0.12)**	
Time slept previous night	−0.60 (0.09)***	
Mentions in the diaries to …		
writing activities	−0.61 (0.10)***	
organizing activities	0.33 (0.09)**	
reading activities	−0.38 (0.10)**	
email activities	−0.87 (0.13)***	
working at the lab	0.85 (0.22)**	
First tree split (idiographic tree, from diaries):		
Mentions organizing activities (+)	−1.11 (0.20)***	
Mentions review activities (+)	0.13 (0.19)	
Time slept previous night (−)	0.09 (0.18)	
Time slept previous night (+)	−0.73 (0.27)*	
Time spent on non-thesis work (−)		2.67 (9.69)
Time spent on thesis work (+)	0.01 (0.14)	−0.59 (7.31)
Total time worked (−)	−0.03 (0.18)	
Total time worked (+)	0.14 (0.15)	−0.07 (9.84)
(Intercept)	0.14 (0.14)	1.94 (6.97)
Observations	20	13
R-squared	0.97	0.44
Adjusted R-squared	0.87	0.16

Note: *p < 0.1; **p < 0.05; ***p < 0.01.

In dataset 2 (see Table 6, right) we find a much simpler model. In this case, the only significant predictor of differential benefits is the initial level of symptoms: the worse their initial well-being, the more a student would benefit from the workshops (please note that the polarity of the DASS, a measure of symptoms, is reversed, so more negative scores are better). From the process-oriented variables we only find the ones coming from the idiographic regression trees, which show (non-significant) results consistent with those of dataset 1: students with certain patterns of experience (e.g., those for which the time dedicated to the thesis is an important predictor of progress) benefited more; those with other patterns (e.g., those for which time on non-thesis work activities is an important negative predictor of progress – typical of part-time doctoral students) benefited less.

What does the interpretation of these nomothetic (but idiographic-informed) models imply for the workshop instructors? The results suggest that the workshops will benefit more those students that already have emotional well-being symptoms at the outset but may not benefit as much those students that come feeling "blocked" (and hence additional remedial actions or activities may be needed for them). Aside from the potential for personalized advice coming from the idiographic variables (similar to what was presented in RQ1 above), the patterns of experience/process uncovered by these models could be used by instructors in advertising the course to attract students more likely to benefit (e.g., pointing to experiences of progressing at the cost of sufficient sleep, email interfering in one's progress), and to help emphasize to students the benefits of doing the diary exercise and uncovering such work patterns to improve well-being. It also may spur the development of specialized or new training actions (such as those for part-time doctoral students, which now may be benefitting less from the workshops). It could also prompt the instructors to increase the emphasis on the time management component of the workshops (to focus more on helping part-time doctoral students likely to struggle with these issues, and those finding out certain unhelpful patterns such as the issue with email being negatively related to progress).

4.3. RQ3: do SCLA-based indicators provide added value over other indicators?

As learning analytics researchers, we may wonder whether the extra effort and complexity of setting up a SCLA-based analysis provides value over more typical data gathering options (e.g., demographic or theory-based variables measured pre-training) when trying to understand the differences in outcomes from doctoral interventions. To understand the relative added value of these different kinds of variables, we compared the performance of three linear regression models of workshop benefits that used a) a baseline model using only the initial well-being as predictor (GM2); b) the demographic and theory-based variables measured at pre-test (GM3); and c) the process-oriented variables extracted from idiographic models of progress (see RQ2) (GM4). In dataset 1, using the subset of students ($N = 20$) that had both pre-post and 5+ entries of diary data (to enable meaningful model comparison), we found that that the process-oriented variables (and the resulting model) were the best-performing, according to the five different performance metrics tested (Table 7, top). In dataset 2, however, comparing the performance of the three models built with these different variable sets (GM2, GM3, GM4) using the subset of students that had both pre-post and 5+ entries of diary data ($N = 13$), we can see that no model consistently outperformed the others, with the baseline model using just the initial symptoms as predictor performing quite strongly (as noted in RQ2 above, the best predictor of workshop benefits in this dataset is entering the workshop with more emotional distress symptoms). Further details about each of these models built on different variable sets can be found in Appendix C.

5. Discussion

The present study explored the potential of single-case learning analytics to help understand the differential impact and potential personalization of doctoral education interventions. This approach was examined using datasets gathered in the context of doctoral workshops to address emotional well-being, under constraints typical of authentic

Table 7
Performance comparison for different kinds of regression models of workshop benefits.

Performance metric	Baseline model (GM2) (Initial well-being)	Pre-test variables model (GM3) (Initial well-being + pre-test variables)	Process variables (stepwise) (GM4) (Initial well-being + idiographic variables)
Dataset 1			
R-squared	0.059	0.083	**0.932**
Adjusted R-squared	0.016	−0.11	**0.74**
RMSE*	0.294	0.290	**0.083**
AIC*	15.4	20.8	**−10.7**
BIC*	18.9	27.8	**5.2**
Dataset 2			
R-squared	0.426	**0.500**	0.426
Adjusted R-squared	**0.374**	0.001	0.312
RMSE*	5.441	**5.078**	5.441
AIC*	**86.9**	95.1	88.9
BIC*	**88.6**	99.7	91.2

Note. Evaluation on the subsample of students with both pre-, post-, and diary data. * = Performance metric is reversed (lower is better). In **bold**, best performance.

DE practice. We posit that SCLA (as laid out in the "Single-case Learning Analytics" section) can help model doctoral student heterogeneity in a manner aligned with idiographic methods criteria, such as the search for patterns with commonalities across individuals, but not identical across individuals; or the observation of "typical patterns" (Bergman & Wångby, 2014).

From an idiographic, learner-centred perspective, we extracted practice-oriented advice and feedback from the triangulation of simple, interpretable idiographic models of the students' daily experience through a well-being related construct (one's satisfaction with daily progress) (RQ1). This individualized modeling of longitudinal data about each person's context/experience (enhanced by qualitative analysis of unstructured diary data) provided exploratory hints and hypotheses about contextual elements that could be leveraged for increased progress (and, hopefully, well-being – cf. Milicev et al., 2021). Such insights could be used for normative personalized feedback (e.g., do more of X) or probing reflective questions (why do you think X seems related to higher/lower progress?) which can be even more productive, considering that our simple models track correlations, not necessarily causes of progress. It is still unclear what form should such (automated) feedback take: our choice of simple models (linear regressions, trees) enables comparatively simple visual representations – although many students and practitioners may find textual explanations and tips more meaningful. Thankfully, such transformation of models into (personalized) feedback text at scale is becoming increasingly feasible with the advent of large language models (Stamper et al., 2024; Yan et al., 2024).

As a secondary added value of the SCLA approach, from a practitioner (i.e., doctoral trainer) perspective (RQ2), our illustrative empirical study first used the classic (nomothetic) pre-post approach to determine that workshops were generally beneficial for students' well-being (RQ2.1), confirming prior results (cf. Prieto et al., 2022). While this prompts instructors to continue these interventions, our descriptive look at the high outcome heterogeneity (e.g., proportion of students not benefitting, or doing so even more greatly) mirrored those of prior research in higher education (Kizilcec et al., 2020; Saqr, 2023). Regarding the thorny issue of who benefited more/less from the workshops (RQ2.2), our exploratory (stepwise) regression models using different kinds of data (pre-test covariates and idiographic process-oriented ones derived from the student diaries) led to insights about student profiles that may benefit more (those with higher distress symptoms) or less (e.g., those feeling "blocked" or having other non-thesis work tasks interfere with their daily progress). While such inferences are not the main purpose of SCLA, they could, with additional support from other doctoral education studies (e.g., the roles of progress, appropriation in doctoral dropouts, see De Clercq et al., 2021) inform profiling and recruiting strategies/criteria on the ground (e.g., questionnaires for triage during workshop registration). These models also uncovered typical processes (or features of processes, like the strong influence of time working on the thesis on progress) worth paying attention to. They also served to characterize student heterogeneity and discover atypical (but relevant) processes/features important for certain individuals (cf. Bergman & Wångby, 2014) (e.g., mentions to working at the lab as positively associated with progress, relevant after the COVID-19 pandemic that has prompted many to work from home more frequently). The sparsity of the features extracted from these diary models are a further sign of the high heterogeneity of students and experiences and suggests the need for future improvements in extracting features and modeling such short-length, small-sample time series.

Finally, as another secondary added value of the SCLA approach, the comparison of models of differential workshop outcomes based on different kinds of variables (pre-test vs. process-oriented from the idiographic models above) (RQ3) indicated that SCLA-based features can be helpful in nomothetic analyses of well-being outcomes (at least, from a positive psychology perspective – as demonstrated by the PsyCap models), by unearthing individual experience and process patterns. The failure to outperform other variables in the more classically-oriented modeling of well-being as the absence of distress symptoms (maybe due to the small number of students with enough diary entries to extract SCLA variables), however, suggests that more studies and evidence are needed to confirm these initial results.

We thus conclude that an SCLA approach, even when implemented using very simple technologies (like a web-based diary/questionnaire) and analysis methods (linear and tree regressions), has the potential to help DE researchers and practitioners understand *for whom* certain interventions worked, develop their interventions further, and aid students in leading more productive doctoral processes.

5.1. Implications for practice

The key value for learners of applying SCLA are the findings of its idiographic component, which suggest individual-specific focus points such as time spent on thesis tasks or research activities often associated with progress. The modeling of data about student experience and process from a relatively short period of time, albeit limited by the small amount of data, helped detect particular activities or contextual features associated with progress for each specific student (sometimes, countering or unrelated to the typical/group-level patterns) as exploratory hypotheses for further self-investigation. Indeed, the importance of these initial insights lies in their potential as motivators for longer, more consistent learner self-tracking (cf. tracking fatigue, Choe et al., 2014), rather than in their inferential power. Such longer time series are essential to expand the idiographic models in SCLA to more powerful sequence-aware methods (e.g., GGMs). In this regard, it is worth pointing out that the usefulness of SCLA as an intensive analytics method is not limited to doctoral training actions, and it could be more widely implemented in doctoral education, e.g., by individual students applying the method on an ongoing basis throughout (all or part of) their dissertation, to monitor and reflect on their progress and well-being. Indeed, pilots of such use beyond training actions are currently underway.

Another value of SCLA for DE practitioners stems from how it can inform nomothetic analyses. In our case study, this is exemplified by the suggestion of specific re-designs of the interventions: to introduce screening of relevant variables, triaging/prioritizing of certain student profiles, or the increase in emphasis on the time management component of the workshops, are just a few examples drawn from the triangulation of simple models in our study.

Technology designers can also derive insight from our exploratory mixed-methods case study. Regarding the practical feasibility of implementing this SCLA approach with current technology, the main hurdle is the qualitative analysis of unstructured diary text (performed manually for this paper). Recent advances in natural language processing (e.g., large language models) make this problem increasingly tractable, and there is evidence that such models are starting to perform such coding tasks (especially deductive coding about concrete aspects, such as our ACAD-based codes) with acceptable accuracy (Garg et al., 2024; Hou et al., 2024; Pishtari et al., 2022). Considering the short timeframe of this workshop and other doctoral training, gathering data once a day about the doctoral process may be insufficient to use more advanced idiographic methods (e.g., GGMs, GIMME, etc.). Thus, other (unobtrusive) strategies with higher data frequency (sensors, ecological momentary assessment), and/or the tracking of doctoral experience in the longer-term after the workshop ends, should also be part of future work in this area. To implement this, a human-centered approach to LA design seems to be a promising way of achieving the scalable "contextual personalization" that SCLA aims for (Prieto et al., 2023; Wiley et al., 2024).

5.2. Limitations

Of the SCLA method as operationalized here. Maybe the main limitation of the SCLA operationalization presented in this paper is the fact that the simple analysis methods proposed are not sequence-aware (i.e., specific for time series data). This modeling choice was motivated by very real limitations of the authentic setting of doctoral training (often leading to small sample sizes and short time series which made temporal idiographic analyses unfeasible). While our modeling choice limits the idiographic insights we can extract from the data and faces threats in terms of the independence of samples, we tried to compensate through the triangulation of linear and non-linear models (inspired by the work of Fisher et al., 2019), so that the triangulation of predictors from our exploratory models can still provide initial ideas of "variation within a learner over time" (Beltz et al., 2016). The high number of predictors that the triangulation of quantitative and qualitative data fosters (and the issues of multicollinearity they may bring with them) further compound these threats. Different statistical methods like Least Absolute Shrinkage and Selection Operator (LASSO), Partial Least Squares (PLS) regression or Bayesian methods could be applied to ameliorate some of these threats, albeit their interpretation by non-experts is less straightforward. Another limitation is that our idiographic models aimed to predict a proximal (indeed, cotemporaneous) outcome – more distal outcomes (like long-term progress or well-being) would be even more interesting to predict. Yet, the statistical models developed for this study only scratched the surface of what is possible with an SCLA approach, since the data source modalities were limited to (low-frequency) diaries and self-tracking. Data gathering using other techniques and modalities (e.g., ecological momentary assessment, physiological sensors – i.e., multimodal learning analytics) could help address some of the data size limitations of our current approach – and could unlock more advanced idiographic and idiothetic modeling of doctoral student processes, such as GGMs, GIMME, etc. (e.g., Saqr & Lopez-Pernas, 2021). Whichever analysis method we select, however, we should pay special attention to the interpretability of such models, as they are intended for use by practitioners, be them doctoral instructors or doctoral students themselves. A more thorough validation of the LA methods used is beyond the scope of the present paper, but such method replication in different doctoral education settings (also targeting well-being, or maybe different constructs) should be performed, to demonstrate that SCLA also provides value there.

Of the illustrative empirical study. Most of the limitations of the illustrative study are related, one way or another, to the small sample size of the interventions, which we already described as typical of the DE setting. This should make students, instructors, and LA technology designers wary of deriving too strong practical implications from the idiographic models of progress: they are rather meant as exploratory, hypothesis-generating tools (Ludwig & Mullainathan, 2024) to elicit potentially interesting everyday work factors that can be further investigated in a more sustained data gathering after the workshops. Larger studies in different countries and different settings are indeed needed to validate the particular findings of our illustrative empirical study – yet, our samples (and the sub-optimal internal consistency of certain variables) were typical of authentic DE settings for this kind of training. We hope the triangulation of "multiple simple models" (cf. Tsipras et al., 2018) and the application of more powerful temporal idiographic models when longer sequences of data are available, helps address the issue of the instability of the models developed with such small samples – not with the goal of demonstrating that these results and trends are true for the doctoral student population, but rather to reliably characterize the sample we *did* have (Shaffer & Serlin, 2004), which is more important for practitioners.

Another important methodological limitation of our study (which is likely to appear for practitioners reproducing this approach in DE) is the issue of missing data, due to the optionality of the workshop activities and anonymous data gathering. While this is certainly sub-optimal from a research standpoint, we aimed to reproduce the kinds of choices that DE practitioners often make in authentic settings, in particular to foster autonomy and honesty about sensitive topics such as well-being. In the future, instructors need to find ways to motivate students to fill in such diaries (or whichever process data is considered relevant), maybe using our data and insights as an example. Indeed, tracking for extended periods of time (or even the whole doctoral process) through journaling may not be feasible (or desirable) for everyone. To deal with tracking fatigue and the issue of missing data, SCLA-based research designs could use journaling as an occasional complement to other less effortful data sources (different logs and sensors, i.e. multimodal learning analytics), use analytics techniques robust to missing data (e.g., State Space Models), use different journaling formats (e.g., voice memos), and provide added value to students (e.g., through personalized analytics that idiographic models enable) from the outset of such longitudinal efforts.

Other limitations rather relate to the constructs and instruments used in our research design. The subjective perception of progress is a powerful predictor of emotional well-being (De Clercq et al., 2021; Devos et al., 2017; Milicev et al., 2021) but may not be fully accurate, and may prompt the use of data from other sources/stakeholders in the future (e.g., doctoral supervisors). Albeit several of the questionnaires used in the study were not validated (which limits the validity and generalizability of the study's findings), they still illustrate the use and value of the SCLA approach and mirror the kind of ad-hoc instruments often used by practitioners.

The fact that many of our models (e.g., in RQ2), even when kept simple via stepwise processes, showed limited performance or inconclusive predictors, is also a limitation (but is likely to occur in authentic DE settings). This is related to the duration and sample size limitations noted above, but in part also stems from our adherence to a frequentist statistics approach (i.e., the use of *p*-values, etc.). Alternative versions of SCLA using a Bayesian paradigm could somewhat sidestep this limitation, if ways are found to make the interpretation of resulting insights approachable to students and practitioners.

6. Conclusion

We are currently facing a mental health crisis in doctoral education (Evans et al., 2018) – and the practice and profession of research keeps changing (see, e.g., Gray, 2024). We thus need interventions that can help doctoral students develop the socio-emotional skills required not only to face the uncertainties of the PhD, but also of the scientific profession afterwards (and which help students with low well-being right now). All this, taking into account that each student and each thesis is unique.

This paper has proposed a mixed-methods idiographic LA approach using learning analytics to understand socio-emotional learning processes and the effects of such interventions within authentic DE practice constraints (using a simple web-based questionnaire data gathering strategy replicable by practitioners). Our initial evaluations suggest that the idiographic analysis of qualitative data for contextual/experience cues, combined with quantitative data about theory-grounded constructs, could provide practical insights for students about their doctoral processes. As a secondary contribution, we showcase that integrating idiographic indicators into nomothetic analyses can help instructors understand average and differential intervention effects in authentic doctoral education settings. Such insights include intervention re-design ideas, personalized factors for student feedback, and typical/atypical patterns that seem associated with higher/lower benefits. We hope these multiple added values of the SCLA approach helps shift idiographic analyses from its current niche status toward mainstream integration with other (often, nomothetic) learning analytics approaches.

Our future work will expand the variety and depth of data sources, analytics techniques, and temporal scope of SCLA application beyond these short interventions, toward a longer-term support throughout the doctoral process (indeed, initial pilots of longer-term usage out of doctoral training are already underway). To realize its aim as a practice-oriented analytical approach, SCLA will require further development about how its results actually inform stakeholders (i.e., how to present the triangulation of models to trainers/students) so that they can act upon them (Jung & Wise, 2024). We will also need evaluation studies of SCLA with end users in authentic DE settings.

We hope this kind of technology will help us (and the new generations of scientists) navigate the turbulent waters ahead. This kind of approach could also be useful for other areas in education where heterogeneity is high and sample sizes are small (e.g., special education or teacher professional development). And we hope this kind of study will kindle a wider conversation about how we can make the results of idiographic methods of analysis of learning experiences better available to stakeholders on the ground.

CRediT authorship contribution statement

Luis P. Prieto: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jelena Jovanovic:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Paula Odriozola-González:** Writing – review & editing, Validation, Resources, Methodology, Investigation, Conceptualization. **María Jesús Rodríguez-Triana:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization. **Henry Benjamín Díaz-Chavarría:** Writing – review & editing, Validation, Investigation, Formal analysis, Data curation. **Yannis Dimitriadis:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Ethics statement

All participants in the workshops gave voluntary and informed consent to participate. The procedures for the data collection were approved by the CEITER project's Ethics Committee (i.e., the local Institutional Review Board – IRB, review cases #0002 and #0003).

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

The present work has been supported by grants PID2020-112584RB-C32 and PID2023-146692OB-C32, and grants RYC2021-032273-I and RYC2022-037806-I, all financed by MCIN/AEI/10.13039/501100011033. These grants have been co-funded by the European Union's ERDF, NextGenerationEU/PRTR, and ESF+. The present work has also been supported by the Regional Government of Castile and Leon, under project grant VA176P23.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lindif.2025.102705.

References

Amabile, T., & Kramer, S. (2011). *The progress principle: Using small wins to ignite joy, engagement, and creativity at work*. Harvard Business Press.

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology, 68*(6), 627–636. https://doi.org/10.1016/j.jclinepi.2014.12.014

Barry, K. M., Woods, M., Martin, A., Stirling, C., & Warnecke, E. (2019). A randomized controlled trial of the effects of mindfulness practice on doctoral candidate psychological status. *Journal of American College Health, 67*(4), 299–307.

Beck, A. T., & Steer, R. A. (1993). *Beck scale for suscide ideation: Manual*. Psychological Corporation.

Beerkens, M. (2018). Evidence-based policy and higher education quality assurance: Progress, pitfalls and promise. *European Journal of Higher Education, 8*(3), 272–287. https://doi.org/10.1080/21568235.2018.1475248

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

Bergman, L. R. (1998). A pattern-oriented approach to studying individual development: Snapshots and processes. In *Methods and models for studying the individual* (pp. 83–122). Sage Publications, Inc.

Bergman, L. R., & Wångby, M. (2014). The person-oriented approach: A short theoretical and practical guide. Eesti Haridusteaduste Ajakiri. Estonian. *Journal of Education, 2* (1), 29–49. https://doi.org/10.12697/eha.2014.2.1.02b

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., … Zettle, R. D. (2011). Preliminary psychometric properties of the acceptance and action questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688.

Brooks, G., & Ruengvirayudh, P. (2016). Best-subset selection criteria for multiple linear regression. *General Linear Model Journal, 42*(2), 14–25.

Byrom, N. C., Dinu, L., Kirkman, A., & Hughes, G. (2020). Predicting stress and mental wellbeing among doctoral researchers. *Journal of Mental Health, 31*(6), 783–791.

Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D. W., Hu, X., & Graesser, A. C. (2019). nCoder+: A semantic tool for improving recall of nCoder coding. In *International Conference on Quantitative Ethnography* (pp. 41–54).

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014). Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1143–1152). https://doi.org/10.1145/2556288.2557372

Cornér, S., Löfström, E., Pyhältö, K., & others. (2017). The relationship between doctoral students' perceptions of supervision and burnout. *International Journal of Doctoral Studies, 12*, 091–106.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

De Clercq, M., Frenay, M., Azzi, A., Klein, O., & Galand, B. (2021). All you need is self-determination: Investigation of PhD students' motivation profiles and their impact on the doctoral completion process. *International Journal of Doctoral Studies, 16*, 189–209. https://doi.org/10.28945/4702

Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods* (1st ed.). Routledge. https://doi.org/10.4324/9781315134543

Devos, C., Boudrenghien, G., Van der Linden, N., Azzi, A., Frenay, M., Galand, B., & Klein, O. (2017). Doctoral students' experiences leading to completion or attrition: A matter of sense, progress and distress. *European Journal of Psychology of Education, 32* (1), 61–77.

Di Mitri, D., Scheffel, M., Drachsler, H., Börner, D., Ternier, S., & Specht, M. (2017). Learning pulse: A machine learning approach for predicting performance in self-regulated learning using multimodal data. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 188–197).

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

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*(3), 282–284.

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research, 20*(177), 1–81.

Garg, R., Han, J., Cheng, Y., Fang, Z., & Swiecki, Z. (2024). Automated discourse analysis via generative artificial intelligence. In *Proceedings of the 14th learning analytics and knowledge conference* (pp. 814–820). https://doi.org/10.1145/3636555.3636879

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*, 68–84. https://doi.org/10.1016/j.iheduc.2015.10.002

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71.

Gates, K. M. (2024). Group iterative multiple model estimation (GIMME). https://cran.r-project.org/web/packages/gimme/vignettes/gimme_vignette.html.

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage, 63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026

Godwin, A., Benedict, B., Rohde, J., Thielmeyer, A., Perkins, H., Major, J., … Chen, Z. (2021). New epistemological perspectives on quantitative methods: An example using topological data analysis. *Studies in Engineering Education, 2*(1), 16. https://doi.org/10.21061/see.18

Goodyear, P., Carvalho, L., & Yeoman, P. (2021). Activity-centred analysis and design (ACAD): Core purposes, distinctive qualities and current developments. *Educational Technology Research and Development, 69*(2), 445–464. https://doi.org/10.1007/s11423-020-09926-7

Gray, A. (2024). *ChatGPT "contamination": Estimating the prevalence of LLMs in the scholarly literature (version 1)*. arXiv. https://doi.org/10.48550/ARXIV.2403.16887

Guo, L., Fan, H., Xu, Z., Li, J., Chen, T., Zhang, Z., & Yang, K. (2021). Prevalence and changes in depressive symptoms among postgraduate students: A systematic review and meta-analysis from 1980 to 2020. *Stress and Health, 37*(5), 835–847. https://doi.org/10.1002/smi.3045

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*(2), 227–239. https://doi.org/10.1348/014466505X29657

Hish, A. J., Nagy, G. A., Fang, C. M., Kelley, L., Nicchitta, C. V., Dzirasa, K., & Rosenthal, M. Z. (2020). Acceptability and perceived effectiveness of approaches to support biomedical doctoral student wellness: One size doesn't fit all. *International Journal of Doctoral Studies, 15*, 653–684. https://doi.org/10.28945/4669

Hou, C., Zhu, G., Zheng, J., Zhang, L., Huang, X., Zhong, T., … Ker, C. L. (2024). Prompt-based and fine-tuned GPT models for context-dependent and -independent deductive coding in social annotation. In *Proceedings of the 14th learning analytics and knowledge conference* (pp. 518–528). https://doi.org/10.1145/3636555.3636910

Howard, M. C., & Hoffman, M. E. (2018). Variable-centered, person-centered, and person-specific approaches: Where theory meets the method. *Organizational Research Methods, 21*(4), 846–876. https://doi.org/10.1177/1094428117744021

Jackman, P. C., Jacobs, L., Hawkins, R. M., & Sisson, K. (2022). Mental health and psychological wellbeing in the early stages of doctoral study: A systematic review. *European Journal of Higher Education, 12*(3), 293–313. https://doi.org/10.1080/21568235.2021.1939752

Jung, Y., & Wise, A. F. (2024). Probing Actionability in learning analytics: The role of routines, timing, and pathways. In *Proceedings of the 14th learning analytics and knowledge conference* (pp. 871–877). https://doi.org/10.1145/3636555.3636914

Kizilcec, R. F., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., … Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences, 117*(26), 14900–14905. https://doi.org/10.1073/pnas.1921417117

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—A literature review. In A. Peña-Ayala (Ed.)*, Vol. 94*. *Learning analytics: Fundaments, applications, and trends* (pp. 1–23). Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_1.

Levecque, K., Anseel, F., Beuckelaer, A. D., Heyden, J. V.d., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868–879. https://doi.org/10.1016/j.respol.2017.02.008

Lorenz, T., Beer, C., Pütz, J., & Heinitz, K. (2016). Measuring psychological capital: Construction and validation of the compound PsyCap scale (CPC-12). *PLoS One, 11* (4).

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*(3), 335–343.

Ludwig, J., & Mullainathan, S. (2024). Machine learning as a tool for hypothesis generation. *The Quarterly Journal of Economics, 139*(2), 751–827. https://doi.org/10.1093/qje/qjad055

Luthans, F., Avolio, B. J., Avey, J. B., & Norman, S. M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psychology, 60*(3), 541–572.

Mackie, S. A., & Bates, G. W. (2019). Contribution of the doctoral education environment to PhD candidates' mental health problems: A scoping review. *Higher Education Research & Development, 38*(3), 565–578.

McIntyre, D. (2005). Bridging the gap between research and practice. *Cambridge Journal of Education, 35*(3), 357–382. https://doi.org/10.1080/03057640500319065

Milicev, J., McCann, M., Simpson, S. A., Biello, S. M., & Gardani, M. (2021). Evaluating mental health and wellbeing of postgraduate researchers: Prevalence and contributing factors. *Current Psychology, 42*, 12267–12280. https://doi.org/10.1007/s12144-021-02309-y

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson, & M. Kaptein (Eds.), *Modern statistical methods for HCI* (pp. 275–287). Springer International Publishing. https://doi.org/10.1007/978-3-319-26633-6_12.

Overall, N. C., Deane, K. L., & Peterson, E. R. (2011). Promoting doctoral students' research self-efficacy: Combining academic guidance with autonomy support. *Higher Education Research & Development, 30*(6), 791–805. https://doi.org/10.1080/07294360.2010.535508

Pellegrini, M., & Vivanet, G. (2021). Evidence-based policies in education: Initiatives and challenges in Europe. *ECNU Review of Education, 4*(1), 25–45. https://doi.org/10.1177/2096531120924670

Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review, 71*, 90–100. https://doi.org/10.1016/j.cpr.2019.01.002

Pishtari, G., Prieto, L. P., Rodríguez-Triana, M. J., & Martinez-Maldonado, R. (2022). Design analytics for mobile learning: Scaling up the classification of learning designs based on cognitive and contextual elements. *Journal of Learning Analytics, 9*(2), 236–252. https://doi.org/10.18608/jla.2022.7551

Prieto, L. P., Rodríguez-Triana, M. J., Ley, T., & Eagan, B. (2021). In A. R. Ruis, & S. B. Lee (Eds.), *Advances in Quantitative Ethnography, 1312* pp. 202–217). Springer International Publishing. https://doi.org/10.1007/978-3-030-67788-6_14.

Prieto, L. P., Odriozola-González, P., Rodríguez-Triana, M. J., Dimitriadis, Y., & Ley, T. (2022). Progress-Oriented Workshops for Doctoral Well-being: Evidence From a Two-Country Design-Based Research. *International Journal of Doctoral Studies, 17*, 039–066. https://doi.org/10.28945/4898

Prieto, L. P., Pishtari, G., Dimitriadis, Y., Rodríguez-Triana, M. J., Ley, T., & Odriozola-González, P. (2023). Single-case learning analytics: Feasibility of a human-centered analytics approach to support doctoral education. *JUCS - Journal of Universal Computer Science, 29*(9), 1033–1068. https://doi.org/10.3897/jucs.94067

Pyhältö, K., Toom, A., Stubb, J., & Lonka, K. (2012). Challenges of becoming a scholar: A study of doctoral students' problems and well-being. In *ISrn Education, 2012*.

Raufelder, D., Jagenow, D., Hoferichter, F., & Drury, K. M. (2013). The person-oriented approach in the field of educational psychology. *Problems of Psychology in the 21st Century, 5*(1), 79–88.

Saqr, M. (2023). Modelling within-person idiographic variance could help explain and individualize learning. *British Journal of Educational Technology, 54*(5), 1077–1094. https://doi.org/10.1111/bjet.13309

Saqr, M. (2024). Group-level analysis of engagement poorly reflects individual students' processes: Why we need idiographic learning analytics. *Computers in Human Behavior, 150*, Article 107991. https://doi.org/10.1016/j.chb.2023.107991

Saqr, M., & Lopez-Pernas, S. (2021). Idiographic learning analytics: A definition and a case study. In *2021 international conference on advanced learning technologies (ICALT)* (pp. 163–165). https://doi.org/10.1109/ICALT52272.2021.00056

Satinsky, E. N., Kimura, T., Kiang, M. V., Abebe, R., Cunningham, S., Lee, H., … Tsai, A. C. (2021). Systematic review and meta-analysis of depression, anxiety, and suicidal ideation among Ph.D. students. *Scientific Reports, 11*(1), Article 14370. https://doi.org/10.1038/s41598-021-93687-7

Schmidt, M., & Hansson, E. (2018). Doctoral students' well-being: A literature review. *International Journal of Qualitative Studies on Health and Well-Being, 13*(1), Article 1508171. https://doi.org/10.1080/17482631.2018.1508171

Shaffer, D. W. (2017). Quantitative ethnography. Lulu.com.

Shaffer, D. W., & Serlin, R. C. (2004). What good are statistics that don't generalize? *Educational Researcher, 33*(9), 14–25.

Stake, R. (2005). Qualitative case studies. In N. Denzin, & Y. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 443–466). Sage.

Stamper, J., Xiao, R., & Hou, X. (2024). *Enhancing LLM-based feedback: Insights from intelligent tutoring systems and the learning sciences (version 2)*. arXiv. https://doi.org/10.48550/ARXIV.2405.04645

Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of the factors influencing doctoral students' completion, achievement, and well-being. *International Journal of Doctoral Studies, 13*, 361–388. https://doi.org/10.28945/4113

Taylor, S. (2021). Towards describing the global doctoral landscape. UK Council for Graduate Education. http://www.ukcge.ac.uk/article/towards-global-doctoral-landscape-475.aspx.

Tsipras, D., Santurkar, S., Engstrom, L., Turner, A., & Madry, A. (2018). *Robustness may be at odds with accuracy (version 5)*. arXiv. https://doi.org/10.48550/ARXIV.1805.12152

Van der Linden, N., Devos, C., Boudrenghien, G., Frenay, M., Azzi, A., Klein, O., & Galand, B. (2018). Gaining insight into doctoral persistence: Development and validation of doctorate-related need support and need satisfaction short scales. *Learning and Individual Differences, 65*, 100–111. https://doi.org/10.1016/j.lindif.2018.03.008

Vanderlinde, R., & Van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal, 36*(2), 299–316. https://doi.org/10.1080/01411920902919257

Webber, C. L., Jr., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences, 94*(2005), 26–94.

Wiley, K., Dimitriadis, Y., & Linn, M. (2024). A human-centred learning analytics approach for developing contextually scalable K-12 teacher dashboards. *British Journal of Educational Technology, 55*(3), 845–885. https://doi.org/10.1111/bjet.13383

Yan, L., Martinez-Maldonado, R., & Gasevic, D. (2024). Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. In *Proceedings of the 14th learning analytics and knowledge conference* (pp. 101–111). https://doi.org/10.1145/3636555.3636856