JOURNAL OF LEARNING ANALYTICS

Volume 9(2), 202–217. https://doi.org/10.18608/jla.2022.7509

Insights of Instructors and Advisors into an Early Prediction Model for Non-Thriving Students

Arnon Hershkovitz[1], G. Alex Ambrose[2]

Abstract

In this qualitative study (N=6), we explored insights of first-year students' instructors and advisors into an early identification system aimed at detecting non-thriving students in the context of an all-campus first-year orientation course for undergraduates. Following the development of that prediction model in a bottom-up manner, using a plethora of available data, we focus on how its end-users could help us understand the underlying mechanisms that drive the identification of non-thriving students. As findings suggest, participants were appreciative overall of the prediction and its timing and came up with various behaviours that could explain non-thriving, mostly motivation and engagement. They suggested additional data that could predict non-thriving, including background information, academic engagement, and learning habits.

Notes for Practice

- Change terminology regarding early identification systems from survive to thrive, to reflect students' potential realization
- Combine quantitative and qualitative approaches for model construction to achieve reliable, authentic learning analytics
- Set institutional policies regarding a wide range of data types, to support effective data-driven decision-making

Keywords

Early warning system, early identification system, non-thriving, prediction model, instructor perceptions, advisor perceptions, data-driven decision-making

Submitted: 03/06/20 — **Accepted:** 21/03/21 — **Published:** 21/06/22

Corresponding author [1]Email: arnonhe@tauex.tau.ac.il Address: School of Education, Tel Aviv University, Israel 6997801. ORCID ID: https://orcid.org/0000-0003-1568-2238

[2]Email: gambrose@nd.edu Address: Kaneb Center for Teaching and Learning, Notre Dame Learning, University of Notre Dame, Notre Dame, IN 46556, USA.

1. Introduction

Early prediction of non-desired academic behaviours has been recognized as an important step in bringing at-risk students back on track. In higher education, such prediction models — frequently referred to as early warning systems (EWSs) — have been mostly developed to identify performance or dropout in the context of a course or an academic program (Liz-Domínguez et al., 2019; Na & Tasir, 2017). Such a prediction will allow early intervention to bring the at-risk students to a successful completion of their academic endeavour (Lizzio & Wilson, 2013; Pérez, 1998).

EWSs use various types of data as indicators for their predictions; among the most common are demographics, assignment submission, grades, attendance, and patterns of using learning management systems (Liz-Domínguez et al., 2019). This is usually done using a purely quantitative approach, feeding the models with as much data as available. Therefore, understanding what the best predictors are to identify at-risk students, how to build a prediction model, and when is the most optimal time to suggest a prediction are all part of a well-established field of inquiry, usually considered as part of the research in learning analytics or educational data mining (e.g., Howard et al., 2018; Hu et al., 2014; Waddington et al., 2016).

Since EWSs are aimed at having a positive impact on student academic progression, some intervention is usually required upon their prediction. That is, red flags raised by the EWSs should be visible to relevant stakeholders, be it students, instructors, advisors, program heads, etc. Hence, stakeholders' understanding and engagement with EWS-driven notifications, the interventions that follow such notifications, and the impact of these interventions are also the focus of an ongoing line of inquiry (Freeman & All, 2017; Klerkx et al., 2017; Lizzio & Wilson, 2013; Michaeli et al., 2020).

However, so far there has been a lack of research that connects these two sides of the EWS coin. That is, how education stakeholders interpret the very mechanism that drives the prediction model — a qualitative endeavour by its very nature — is still under-researched. This angle, taken in the current study, may help enhance such models and improve their use. Our main goal was to explore and unpack education stakeholders' insights into the underlying mechanisms that drive the prediction model with which they were engaged.

Importantly, we also had another goal — to revise the common language, mostly negative, used in this field. We desire to shift the discourse in a more positive direction and to be explicit about this. Therefore, instead of talking about early warning systems that raise red flags to enable interventions for preventing dropout or failure, we talk about early identification for non-thriving that allows for a boost to help students thrive and reach their top potential. In other words, non-thriving is not about failing, but rather about not fulfilling your potential to the fullest. By doing so, we operationalize thriving in a way that raises the bar from simply surviving a course to getting a specific high grade (B and above, as will be explained below).

This semantic twist may have far-reaching implications, since data — and the language used to speak about data — simultaneously constitutes practice and is constituted by it (Kommoju, 2019, p. iv), therefore changing the discourse is an important first step towards changing the way we look at and act upon student behaviour. Risk or non-thriving is a condition or situation; hence it should not be used to describe a person (Strauss, 2019). A similar approach was suggested by Ocumpaugh et al. (2017) in their redesign of data-driven reports to counsellors; for example, carelessness was relabelled as meticulousness, confusion as adequate help-seeking, and boredom as interest level. This shift from a negative framing to a positive framing may impact decision-making, hence may impact student success. Since most of the existing literature still uses terms like early warning systems, at-risk, intervention, and dropout, we will use them in the Literature Review section when referring to previous studies and will shift towards our suggested framing in the sections that follow. For meeting these dual goals, we defined the following research questions:

1) Given an existing data-driven early identification system for non-thriving, how do end-perceive its validity and usefulness?
2) Which further variables, on which the current model is not based, could help in predicting non-thriving?
3) What are the end-users' perceptions of the shift to focus on thriving?

The rest of the paper is organized as follows. In the next section, we review the most relevant, up-to-date literature in two aspects: early identification of non-thriving students, and education stakeholders' engagement with prediction models. Then, in Section 3, we fully describe the methodology underlying our study. In Section 4, we present the findings from our exploration, and while doing so we make sense of them in light of the literature. Finally, in Section 5 we discuss our findings and their implications to the use and research of EWSs.

2. Literature Review

2.1. Predicting Students At-Risk

Higher education has some far-reaching benefits for individuals and may benefit society at large. At an individual level, for example, people who achieve higher levels of education earn more money and are healthier than others. From a broader perspective, higher education graduates, compared to others, pay more taxes and contribute more to the functioning of their communities (Hout, 2012; Ma et al., 2019). Therefore, dropout from higher education has become a major point of interest, which then increased the interest in students who are at-risk for dropping out of their academic studies (Akçapınar et al., 2019; Freeman & All, 2017; Nik Nurul Hafzan et al., 2019; Valentine et al., 2011). Recently there have been substantial efforts to develop and study data-based prediction models that would alert education stakeholders to at-risk students; these are used within EWSs and are mostly studied under the umbrella of learning analytics and educational data mining (Ifenthaler & Yau, 2020; Lawson et al., 2016; Na & Tasir, 2017).

Three major issues are most relevant to such endeavours, namely, the definition of at-risk, the variables used for prediction, and the method by which the prediction is implemented. First, the target variable should be well defined. Different definitions of at-risk have been suggested. Non-thriving is one alternative that raises the bar from at-risk of failing a course or losing overall retention status to an assumption that students should not just survive a course but thrive and complete the course beyond a passing grade (Bartolini et al., 2020; Syed et al., 2019). One of the most common of these refers to the risk of dropout, either from a full program or from a single course. Such predictions attempt to identify which students might not complete the academic path they initiated and to support these students until successful completion (Biswas et al., 2019; Borrella et al., 2019; Cohen, 2017; Yukselturk et al., 2018). Another approach aims at predicting student achievement; that is, identifying an at-risk behaviour at a lower threshold. Here, the chosen focus influences the prediction — e.g., a single course versus the whole course of study — using various points in time, starting from early in the first semester until the final grades for the full program (García-Poole et al., 2019; Lammers et al., 2017; Lehmann, 2014;

Tsiakmaki et al., 2018). Of course, there are other approaches to defining at-risk — or its counter term, success — including student satisfaction, persistence, acquisition of skills and knowledge, and career success (cf. Alyahyan & Düştegör, 2020).

Second, the variables upon which the prediction of at-risk students is built should be determined. A plethora of factors have been demonstrated to be effective in predicting at-risk behaviour. A recent literature review of the factors predicting student academic success categorize such factors into five main groups (Alyahyan & Düştegör, 2020):

1) Demographics, including gender, age, and socio-economic status
2) Prior academic achievement in both university and pre-university studies
3) Environment, including the type of program, semester duration, and class type
4) E-learning activity, focusing on the online components of the learning
5) Psychological, for example, stress or anxiety, motivation, or self-regulation.

Another recent review presented similar findings, dividing the factors impacting dropout into nine dimensions, including study conditions, academic integration, social integration, personal efforts and motivations for studying, information and admission requirements, prior academic achievement in school, personal characteristics, sociodemographic background, and external conditions (Kehm et al., 2019).

Third, the mechanisms underlying the models or algorithms should be set. Data mining and machine learning have been widely used in recent years to build such prediction models (Agrusti et al., 2019; Manrique et al., 2019; Tamada et al., 2019). While data mining generally refers to merely the identification of hidden patterns in large datasets, machine learning-based algorithms are unique in the sense that they can improve automatically via experience. The latter achieve very high levels of accuracy while predicting dropout; a recent literature review had identified 19 machine learning-based models, all of which achieved an accuracy of 91% or higher (Tamada et al., 2019).

2.2. Education Stakeholders' Engagement with Prediction Models

Supporting education stakeholders' decision-making with data is a process consisting of several steps. We differentiate between two major components: overt and covert. The overt components are visible to the stakeholder, those by which they make their decisions. This includes the output of the data-based analyses, usually presented in a friendly, actionable manner, either by summary reports or by dashboards. On the other hand, the covert components are invisible, usually including the data analyzed and mechanisms for analyzing it. When studying education stakeholders' engagement with EWSs, it is most common to explore their engagement with the overt layer, referring to aspects like design issues or actions taken (e.g., Michaeli et al., 2020; Molenaar & Knoop-van Campen, 2019; van Leeuwen et al., 2019). However, the engagement of education stakeholders with the covert layer of such systems has been barely studied.

Teachers and instructors are among the most common users of education-related EWSs. Therefore, their understanding of the models underlying such systems should be emphasized. In a way, this point of view is similar to Oh and Oh's (2011) notion that the application of model-based views in science classrooms requires a clear understanding of the nature of models and modelling in science (p. 1110). Therefore, key stakeholders should be taken as an integral component of any learning analytics implementation (Tsai et al., 2018). This is most important considering the distrust teachers may develop towards data-driven implementations, attributed mostly to the subjective nature of numbers, the fear of power diminution, and approaches to design and implementation (Tsai et al., 2021); it was recently pointed out that such a tendency is particularly prominent when referring to preditions for students at-risk (Kollom et al., 2021). On the other hand, involving teachers in the early stages of learning analytics design, particularly engaging them with data collected on students, may improve their pedagogic and assessment practises, leading potentially to the development of more suitable learning analytics tools (Holstein et al., 2019; Vezzoli et al., 2020). Interestingly, one of the standards for educators, published by the International Society for Technology in Education (Crompton, 2017), refers to teachers as analysts who understand and use data to drive their instruction and promote their students.

Moreover, the ways by which stakeholders interpret the model may impact their use of it, and consequently its effectiveness (Herodotou et al., 2019; Sun et al., 2019), which is why it is imperative to study end-users' understanding of such models. Eventually, incorporating quantitative and qualitative approaches for model construction — that is, building on available data and stakeholders' insights, respectively — would help develop models driven by reliable, authentic data (Whitman, 2020).

3. Methods

3.1. Research Participants

Our participants included six higher education faculty and counsellors (two males, four females) from the University of Notre Dame, a medium-sized private top-20 institution (a total of 8,731 undergraduate students by Fall 2020) located in the Midwest U.S. (see Table 1). The overall student body is 53% male and 47% female.

All participants had been associated with the development and running of a First-Year Experience (FYE) course, or the Introduction to Engineering Course (Intro to Eg), which are described in the next section. The FYE course serves all first-year students (over 2000), and the Intro to Eg has a typical enrollment of 400 students. Of the six participants, two served only in an instructional role (course coordinator and TA) in the Intro to Engineering Course. The other four participants served in an academic advising role for Intro to Eg and in an advising or instructor role in the FYE course. Two out of four advisors were veterans with 20+ years of experience, one of the advisors was recently hired, and both the Intro to Eg instructor and TA had less than two years of teaching experience.

Table 1. Information About the Participants

Participant	Gender	Academic Role	Course
P1	Male	Course Coordinator & Instructor	Intro to Eg
P2	Male	Teaching Assistant	Intro to Eg
P3	Female	Advising Dean	Intro to Eg & FYE
P4	Female	Advising Dean, Course Coordinator & Instructor	FYE
P5	Female	Advisor & Instructor	Intro to Eg & FYE (respectively)
P6	Female	Advisor & Instructor	FYE

3.2. FYE and Intro to Eg and the Non-Thriving Prediction Models

The First-Year Experience (FYE) course, first presented in 2015, is mandatory for all first-year students and draws on 125+ instructors, each of whom leads a standardized section of no more than 19 students. Consisting of two one-credit-hour semester-long courses (Fall and Spring semesters), with a total of about 30 weeks, the FYE course helps students make a meaningful transition to college life by integrating their academic, co-curricular, and residential experiences. Based on their performance in the course assignments, students are awarded a letter grade, with passing grades being (from highest to lowest): A, A-, B+, B, B-, C+, C-, D.

As a mastery-based course, FYE is designed with the expectation that all students who put in the necessary effort should not only succeed but also be on a pathway to thrive. As a result, on average, 90% of the students get an A as their final grade, with a standard deviation of 1.4% every semester. This is why the following operationalization was chosen for non-thriving: A student was defined as non-thriving based on their grade at week 6 of the Fall semester — which is considered an early stage of the course — if it was B- or lower. Since this is the only class that all incoming first-year students take, the course was designed from the start to be optimized to collect meaningful assessment data for early detection of student engagement and retention. Therefore, various types of available digital traces — i.e., learning management system log activity data, weekly exported gradebook scores, and lecture video analytics — were used to develop an early non-thriving prediction model. The earliest and best predictor for non-thriving found by the model was missing two or more weekly preparation homework assignments. This model is in line with previous models in the sense that early grade points and early engagement cues were found to be predictive of the course final grade (Gray & Perkins, 2019; Williams et al., 2021), hence allowing for early prediction that is actionable and useful in assisting at-risk students (Balfanz & Byrnes, 2019; McMahon & Sembiante, 2020).

Our predictions were sent to the advising dean and course coordinator, and they passed it along to the course advisors and instructors. They, in turn, contacted the students identified and gave them early boosts; these boosts included a personalized learner action plan and a customized set of campus resources (e.g., tutoring, mental health services, etc.). Boosting helped 85% of these students to improve by at least a letter grade from midterm to finals. A full description of the development and implementation of this prediction model can be seen in Syed et al. (2019).

Evolving from this success in FYE, the Introduction to Engineering course adapted a similar model, based on historical data, to identify students who might be non-thriving at the end of the semester and boosted those students in an attempt to improve their performance in the course. This first-year engineering gateway course, serving over 400 students a year, found that a trigger of 80% or lower grade on the first three homework assignments combined (out of the maximum score on those assignments) was the most effective for that task. Furthermore, students who responded to the personalized action plan in their boost email performed better than those who did not respond and better than those who would have been boosted based on the same trigger in the 2017 and 2018 fall semesters (Bartolini et al., 2020).

Given that the university had a 98% retention rate from first to second year for Fall 2019 to Fall 2020, the goal of these two large gateway first-year courses was to use learning analytics to create a system that would go beyond identifying at-risk students who may benefit from an intervention. Rather than just helping students survive, this prediction model was designed to help all students thrive. With the risk of incorrectly labelling students as at-risk and executing interventions, early boosts allowed students an opportunity to improve their course grades, and with enough time for doing so.

3.3. Research Approach and Tools

This study took a qualitative approach. Data collection involved in-depth semi-structured interviews conducted by the two authors in November 2019 at the University of Notre Dame. Interviews lasted between 36 and 51 minutes, and were conducted either individually (P3, P4) or with two participants together (P1+P2, P5+P6). The interviews were audio-recorded and fully transcribed before analysis. The design of the interviews allowed us to deviate from the protocol, based on participants' narratives and insights, to fully capture their perceptions on the topic. The questions for the interview protocol are presented in Table 2.

Table 2. Interview Questions

#	Question
1	What is your role in the university, and how are you associated with the FYE prediction model?
2	In your opinion, who is a non-thriving student?
3	To what extent do you think that the historical data, upon which the model was built, could be generalized to future students?
4	Which actions did you take upon getting the model prediction of non-thriving students, and did it help them?
5	Do you think those students who were predicted as non-thriving understood the mechanism of the model?
6	Imagine that you could have access to any data about the students — which data would you wish for?
7	Do you believe thriving students should be notified too?

3.4. Data Analysis

Whereas the initial prediction model was done externally — using a top-down and purely quantitative approach — the interview data were qualitatively analyzed using the conventional (inductive) data analysis approach, that is, in a bottom-up manner, with the researchers immersing themselves in the data independently to allow new insights to emerge (Hsieh & Shannon, 2005). Before coding the data, the two authors had read the interview transcripts and discussed their understanding of them. Then, coding of participants' statements was conducted mostly by the first author, with frequent discussions between the two authors about the emerging codes, until full agreement was achieved. Organizing the coded statement into higher-level themes was done by the two authors jointly. Therefore, the resulting themes are a product of a thorough spiral process of analyzing the data, which allowed us to answer the research questions thoroughly.

4. Findings

4.1. Insights into the Prediction Mechanism (RQ1)

4.1.1. Is the model valid and well timed?

Overall, participants seemed satisfied with how the model had worked, suggesting its validity. As one participant put it, "in the years that we had [the model], it was pretty steady, […] a pretty good predictor" (P3). Another participant linked the model prediction to the actions taken upon its identification of non-thriving students, arguing for its validity, saying that "a model is acceptable if [predicted] information was given to advisors and instructors and they did something [and] fewer people [were] getting a B and below" (P2) — as indeed happened. Therefore, there was overall agreement that the model could be used as-is, because there weren't any significant changes to the course (P2), or — looked at a different way — "if the nature of the class changes, then all bets are off, […] then we won't know what it means. It could mean anything" (P3). Furthermore, our participants agreed that an early-as-possible prediction is a good practice:

> *If we categorize them into the, you know, at-risk students, we want to get to them right away because if we don't, then they get in trouble quickly, both academically and then emotionally also. (P5)*
> *At the end of the semester, it's too late to do anything. (P3)*
> *Why wait to midterm to find out if a kid is struggling? (P4)*

One of the interviewees associated early identification with the data-driven culture of U.S. high schools:

> *The reality in high school is, with whatever the software they use, to report things, [the students are] being graded on everything. They get graded on, you know, if they turned up for lunch on time, and the parents are also getting that. And*

> *so they're actually used to getting that formative quick feedback all of the time. And I think we don't do a good enough job of giving them this quick feedback. (P6)*

One of the main arguments supporting the notion of intervention as early as possible was that the action to be taken was rather simple, and could be carried out at any point. Even a single email notification, or a short conversation, may lead students to change their direction and eventually succeed in the course. As two of the participants put it, "I was stunned how much it helps […] I was really surprised. […] It's mostly about awareness" (P3); "when we made an intervention, the majority of the students self-corrected and finished the course successfully" (P4).

Early identification allowed our participants to implement a simple early boost, like a face-to-face meeting with identified students. During these meetings, they would make "sure we let the students know what resources are there [… as] they might not be aware of office hours" (P1) or "help them find programs that will be funded to help these students" (P5). Notably, these meetings were perceived as more effective than email notifications, as sometimes non-thriving may be associated — as will be detailed below — with difficulties in handling the new academic life; "so if you can't plan your life out and you can't submit homework on time, you're probably not very on top of your email, which means that you're not going to see this opportunity to come through" (P1).

4.1.2. What's in a grade?

Recall that the definition for non-thriving in the FYE course was based on a B- grade in week 6 of the course and that the best predictor for that was missing two or more weekly preparation homework assignments. Such a threshold for labelling non-thriving — or one similar to it — was overall agreed upon by our participants; they suggested that "anyone who gets less than a B in [the course] isn't thriving. But it could be that really we're looking at the C minuses" (P6), or "Non-thriving at the end of the semester, it was deemed as a B" (P1). One participant raised the need for a more comprehensive view on student achievement in other courses as well:

> *The obvious ones are the students whose […] initial grades are low and who are struggling, especially in more than one course. If you have students that are actually doing fine and there's one course giving them trouble, that's not non-thriving. (P3)*

So, upon agreeing on a grade-based identification, how do we set up the predicted values? In other words, how do we decide which grades are indicative of non-thriving? Should it be based on a quantitative, data-driven approach or a qualitative, observation-based approach? Recall that among our interviewees, two were involved with building a prediction mechanism for another course (not the FYE). They described how a combination of different mechanisms led them to that decision. One of them took a bottom-up approach and was plotting all the students, trying to figure out where [non-thriving] was (P2), while the other had gone the other way:

> *Let me look at the non-thriving. Are there any trends among the non-thriving? [...] Can we find any commonality between those? And then, is that different than the general population? (P1)*

Based on this joint approach, they concluded that students whose total homework performance at the six-week early semester cutoff was below the class average (B-) were predicted as non-thriving. For them too, grades are only a proxy for a broader problem, specifically, not using the available resources:

> *That typically means that you're not using the resources that are available to you to do well on it, which then would trigger that you would continue to not use those resources. (P1)*

Importantly, the notion of irregularity in handling assignments — the best predictor for non-thriving — was perceived as a proxy for a broader issue, like disengagement or learning habits:

> *We knew from advising […] that there's always the students who don't connect. So we'd started to think about at what points can we find out any student who's showing signs of disengagement. (P4)*
> *Sometimes what the grade means is you just didn't study the right way. […] Sometimes the grade means you don't have the background and we need to get you a tutor to catch you up. (P3)*

Indeed, another interviewee first mentioned that a non-thriving student is "Someone […] whose motivation is lacking," but immediately made it clear that this lack of motivation will be reflected in academic behaviour:

> *So they're not motivated to either be here […] and so they're not engaged. […] They're not going to class. They're not involved in activities outside the classroom. They may have a cohort of students, but it's very small and they're probably similar in their lack of motivation. […] They're not achieving everything that they can achieve, both in and out of the classroom because they don't want to be here. (P6)*

Other participants emphasized that for first-year students, non-thriving may be associated with other factors, such as mental health (e.g., homesickness or stress) and substance abuse. In that sense, the specific course chosen for prediction has served as a good testbed since it was designed to assist first-year students to adjust to college and is the only common course all students take during the first and second semester. The course is mostly about learning habits and adjustment and not so much about content. As one of the interviewees put it, comparing to a standard, content-based Chemistry course:

> *In Chemistry, if you went to a school that had Honors Chem[istry] and then you did AP Chem[istry] and then you did a second semester, […] of course you're going to do well in Chemistry. […] But you may not know how to study […] So what do those grades mean in Chemistry? You're not sure what they mean, it could just mean that you didn't have the background or whatever. But in this [FYE course], it's pretty much every student has an equal chance to do well in this class. […] And so it's a little pure […] about, in some sense, your understanding of or attitude towards doing the work. (P3)*

Another participant put it this way: "if he's [a student] not doing [weekly homework] for a one-credit class […] he's probably not doing it in Calculus or Chemistry" (P4).

4.2. What Else Can Help in Identifying Non-Thriving? (RQ2)

We asked our participants to think about more types of data that might be useful in predicting non-thriving students, encouraging them not to be limited by existing practices of data collection, even to imagine that any data they could think of may indeed be available. Taken together, participants mentioned a wide range of data, covering different aspects of the learning process, based on their experience in teaching or advising. We grouped these into several categories.

4.2.1. Background information

Some basic information about their students was explicitly mentioned by participants as data that would be helpful. Our participants referred to data about students with disabilities (P5), particularly in the context of any sort of need for extra test time (P2). Additionally, our participants were interested in data about students' general life habits; for example, are they sleeping? (P1), or how much time do they spend on social media? (P6). Another item of background data that participants wished to get was related to college familiarity, that is, whether they could adjust quickly, and how familiar they were with the Notre Dame system (P1). One of the participants elaborated on that as a cue to non-thriving:

> *If we meet with them and they're not sure of the questions that they're asking, or if the questions are pretty far out […] — like first-generation students, for example, may not know exactly what to expect — so those are also students to watch [out for]. (P5)*

4.2.2. Academic engagement

Under this category, participants regarded two aspects of academic engagement. First, and more prominently, they referred to achievement, wishing for data like "One Test Score, AP, credits, those sorts of things" (P1); first exam scores (P5); "SAT scores, Calculus […] or Physics in high school" (P5); "mid-semester grades and then their final grades for the first semester" (P5); homework grades (P3). Note that the variety of scores mentioned refer to different types of assessment tasks and to different points in time.

The second aspect related to data about their studies referred to the extent to which students were engaged with their academic assignments. An obvious cue would be "whether they were not completing the requirements of this class" (P6). However, even when completing a task, it may be useful to have a deeper look at how the students were engaged with it:

> *Some of them put a lot of work into [the final course portfolios, and] I don't know if we spend enough time […] looking at that level of work. Some kids really invest in it [with] interesting graphics and design work. Other kids totally perfunctory, other kids late, insufficient. So maybe a little bit more time on the longer assignments and that might be more text mining. (P4)*

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

4.2.3. Learning habits

Our participants considered the learning process as a broad endeavour and therefore were interested in data that reflected student academic activity not limited to achievement. This was referred to from three points of view. First, they were interested in data about student study habits, for example, "How organized is the person? Do they have a planner?" (P1); "Do you have a system for getting your homework done? Do you have a place in which to study? When do you start studying for the test? How did you study for the test?" (P3); "When are they preparing for class? How many hours they're putting into preparation" (P6).

Second, our participants suggested that data about attendance and engagement would be helpful. Indeed, these are key to success (Torenbeek et al., 2013). In that context, they wished to know "Are they coming to class every day? […] Are they there on time?" (P3); "or if they do come to class [and] they don't participate, [or if] they're not doing any of the homework or the class prompts" (P6). As one participant put it, "attendance […] is the single best indicator of who's going to have not just a low grade here or there, but the kids who do not graduate, the kids who are on probation. My biggest problems […] all stem from excessive absences" (P4). Another participant extended her view on engagement to the classroom dynamics:

> *I'd also like to have a way of capturing the classroom dynamic and seeing how students interact, […] who's at least listening, who's participating. (P4)*

The third aspect regarding students' learning habits regarded the use of resources. This is important since "giving the students resources doesn't necessarily prepare them [if] they don't know how to use the resources" (P1). So, participants wished to know more about the resourcefulness of their students:

> *Where do [they] get [their] help from? […] do they ever see the professor? Do they ever see the TA? […] Is it the kid next door? Is it tutoring? Are you not asking anybody for help? (P3)*
> *Do they have study groups? Are they going to tutoring. Are they going to office hours? (P1)*

One participant emphasized that this issue could be detected even before the beginning of the semester:

> *In the summertime when we prepare their schedules […] they have to give us input […] and if we have trouble hearing back from them […] that can be a good indication. (P5)*

One participant mentioned that some cues could be extracted from the online learning environment he was using, since "the video watch […] gives a little bit of insight into some of this stuff" (P1).

Relevant to the issue of learning habits were also participant mentions of students' nonacademic habits, especially whether they were athletes and what sports are they in; do they have a balance with the curriculum and co-curriculum? (P5).

4.3. Every Student Can Thrive (RQ3)

Referring to the change in terminology — from detecting dropout to identifying non-thriving — and considering their overall appreciation of the model and the various behaviours depicted, we asked participants to think about the notion of early notification on thriving. This idea stems directly from the assumption underlying this shift in terminology, according to which every student should be encouraged to thrive and fulfill their potential.

Notably, we got mixed reactions on this idea. Those who supported it, some of whom even practised it, explained that in some cases, students — mostly in their first year — need perspective about their achievement. As one participant put it, "sometimes they have all A's. It's like you'd think they would know […] but often they need to be told [that they are on track]" (P3). Even if students are aware of their achievement, "recognizing it may be helpful, as they […] live in their own little world" (P1). This could be easily done, as one participant explained her practice:

> *I would have a little slide and I'd say, you know, the average grade was this. So I kind of gave them a sense of, you know, as a class […] 87% of you got this. (P4)*

For those participants who appreciated the idea of notifying on thriving, there was a clear educational agenda underlying it, which is aligned with the very idea of changing the terminology. As one of them put it:

> *I view myself as somebody that wants to help the students, not just as the gatekeeper of knowledge. I want to make sure that we are supporting first-year students. […] I feel they can thrive [if we] give them the skills that they need to do that [… if] we remove all barriers for their succeeding. (P1)*

ISSN 1929-7750 (online).

Another participant mentioned that she was deliberately telling students, “as they leave [a meeting]: ‘You’re doing well, it’s going well, this is good,’ as to decrease their level of anxiety” (P3). Importantly to the very notion of thriving as fulfilling one’s potential, one participant emphasized that in some cases, “a student’s best effort is going to be a C, and if they’re connected on campus and got friends, they’re involved in groups — that student to me is still thriving […] that’s the most they can achieve” (P6).

Contrary to this approach, other participants did not support the idea of notifying thriving students, stating that these students may not benefit from it. If anything, it may be too much pressure: “They’re under a ton of pressure as it is” (P5). According to this point of view, the very notification may even harm students, as it could boost their confidence above a realistic level and they might “get overconfident and […] quit working as hard as they need to because they think, ‘Oh, well I’m doing great’” (P2). P2 was also worried about a scenario where he might say “a 90 or above is great and [the student may think that only] 95 is great, so they may actually [think …] ‘That’s not my level of standards.’”

5. Discussion

In this study, we qualitatively analyzed different stakeholders’ insights into early identification of non-thriving in the context of first-year university students. Overall, our participants appreciated the model’s effectiveness. Such appreciation is crucial, since models are only effective if they help their end-users make decisions based on their predictions (Kappen et al., 2018). They were also appreciative of the opportunity for early boosting, which was rather simple. Recent literature reviews of early warning systems have argued that for an intervention to be effective, it should better be tailored to student needs. However, many schools lack the resources for such interventions (Balfanz & Byrnes, 2019; McMahon & Sembiante, 2020). Therefore, a short meeting may serve as the first low-cost step towards assisting the detected student. However, this may not be sufficient in preventing the predicted outcomes (Plak et al., 2022).

Our qualitative analysis helped us understand what drives the irregularity in handling assignments, which was found by the model to be the best predictor for non-thriving. Mostly, it was about student motivation and engagement, which were indeed shown to be associated with non-thriving (Dames, 2019; Taylor & Harrison, 2018). Also, mental health and substance abuse — which our participants mentioned — may also be associated with non-thriving (Iorfa et al., 2019; Seppälä et al., 2020).

Interestingly, our participants did not agree on whether those students who are predicted to be thriving should be notified. In practice, early warning systems are mostly used for detecting at-risk students (Balfanz & Byrnes, 2019; McMahon & Sembiante, 2020). The idea of notifying thriving students, however, is well established in many educational systems, particularly in higher education institutions where student honour systems have been implemented for decades and have proven success in maintaining students’ quality of academic work (e.g., Seaver & Quarton, 1976). Interestingly, both thriving and non-thriving notifications may result in improved achievement following the notification (Wright, 2020). The disagreement between our participants echoes the complexity of feedback acceptance among students, which depends on the feedback itself, on the context in which it is given; and on individual capacity; it may also yield various emotional responses (Henderson et al., 2019; Lim et al., 2020).

5.1. Which Data Can Help in Predicting Non-Thriving?

Notably, our findings suggest that early identification of non-thriving should be based — as end-users suggested — on a plethora of data types that could target improvements to the next iteration of the first semester, first year, early non-thriving prediction model. The first category includes background information, and most of the factors mentioned by our participants may indeed be relevant to student success in higher education: student disabilities are inherently assumed, and empirically proven, to be associated with objective challenges in higher education (Richman, 2013); sleep habits were shown to be negatively associated with academic performance (Owens et al., 2017; Suardiaz-Muro et al., 2020); and preparation for college may indeed increase a student’s chances of succeeding (Cates & Schaefle, 2011; Woods-Weeks, 2017). In contrast, habits of social media use, mentioned by our participants, may not straightforwardly indicate a student’s chances of succeeding in higher education (Holmes et al., 2020; Mastrodicasa & Metellus, 2013).

A second category that predicts thriving is academic engagement. As studies have repeatedly shown, success in higher education may indeed be predicted by a wide range of grades, either from their early stages of academic studies or from prior studies (O’Neill et al., 2011; Paura & Arhipova, 2014; Rovira et al., 2017). Engagement may also be related to success; one possible explanation relies on the time invested in school work (Masui et al., 2014; Wagner et al., 2008), while another refers to the benefits of using elaboration strategies (Cebesoy & Akinoglu, 2012; Weinstein, 1982).

Finally, the third category is learning habits, which is again a relevant factor in succeeding in higher education (Lammers et al., 2001; Nakayama et al., 2017; Selvig et al., 2015). Among these, the use of resources, mentioned in our study, is associated with academic achievement (Selvig et al., 2015). Interestingly, regarding academic habits, previous studies have shown mixed results, with some showing that college athletics may contribute to academic success (James, 2013), while others show that sports hinders the academic experience of student-athletes (Stiles Hanlon, 2018).

It is interesting to compare the types of data we found relevant for instructors for early identification of non-thriving with the perceptions of end-users about the data relevant to prevent dropout. This may be inferred from the host of factors that teachers perceive as leading to dropout, including low academic achievement, lack of effort or motivation, irregular attendance, disciplinary issues, and lack of parental involvement (Doll, 2010; Kennedy, 2017; Knesting-Lund et al., 2013; Owen, 2009; Zabalou, 2021). In terms of data, the categories we found pertaining to non-thriving are similar to those associated with dropout, which strengthen the importance of the terminology shift we suggest here. Therefore, changing the terminology will not require changing the whole student support system, since it relies on similar grounds in both cases. Contrarily, it may help educators and administrators understand that this support system should treat the student as a whole person. As we have seen, some participants in this study explicitly support this notion. Moreover, it may emphasize that thriving, which is a student's inner state, is also dependent on students themselves, hence promoting it should be considered a dual effort by both students and institutions (Larsen et al., 2019). This echoes O'Shea et al.'s (2016) suggestion of changing the discourse regarding diversity in the student population; instead of talking about disadvantaged students, they suggest talking about inclusion, which also highlights that both parties — students and institutions — should take part in this process.

Still, how to realize and implement such prediction models based on the rich data suggested here is an open question. Some types of data may be relatively easy to collect — for example, academic achievement (via learning management systems), or background information (via institutional administrative databases) — but even this collection —specifically, the triangulation of data from multiple datasets — may present higher education policy-makers with important ethical and privacy considerations (Willis et al., 2016). On the other hand, even before considering ethical and privacy issues, other types of data may prove a challenge to detect automatically in the first place, like engagement, learning habits, motivation, or social media activity. While feasible in principle (Dang & Koedinger, 2020; Hutt et al., 2021; Molenaar et al., 2020), detection of such variables may be heavily dependent on many contextual factors — such as systems used or sociocultural background — and require extensive efforts in building valid, reliable prediction models. As a first step, such data could be collected in a self-report manner that still can improve prediction (Salehian Kia et al., 2021).

5.2. Institute-Level Policies and Practices

At the institutional level, our findings suggest that data-related policies and practices should refer to a wide range of data to support learners. Based on our participants' experience, student success may be associated with-, and therefore predicted by, a plethora of factors. Hence, collecting these measures and incorporating them within EWSs may be beneficial. Such policies should be defined and carried out as part of an institutional structure that supports the relevant stakeholders while developing appropriate practices across different contexts and employing ethical standards (Vuorikari & Castaño Muñoz, 2016; Scholes, 2016; Willis et al., 2016). Designing these policies and practices together with the different stakeholders, including students, is preferable in handling this process (Tsai et al., 2020).

Lastly, these policies should also refer to the actions to be taken once a prediction model triggers an early identification. In the case of the model studied here, even a minor intervention — such as notifying the student, or having a short talk with them — may be helpful (Syed et al., 2019). Indeed, self-awareness and some consultation — even, as simple as a check-in, check-out — have been shown to be effective as an intervention tool for at-risk students (Hawken et al., 2014; Lizzio & Wilson, 2013; Zhang et al., 2014). However, one should consider that different predicted measures may require other interventions. In any case, such intervention should be considered as an important component of the learning analytics cycle. This cycle begins with learners, who generate data, which is processed into metrics, which are then used to inform actionable interventions, which in turn impact learners (Clow, 2013). While doing so, it is important to resolve potential ethical and privacy-related dilemmas — which may be inherent to measuring student behaviour (Macfarlane, 2015) — hence the institutional policy should encourage transparency of the full learning analytics cycle and should listen to student voices (Ifenthaler & Schumacher, 2016; Jones, 2019; Lawson et al., 2016; Sun et al., 2019).

5.3. Conclusions and Implications

In this study, we explored the insights of higher education stakeholders into a non-thriving prediction model. Overall, the model was appreciated as helpful, and we were able to identify some of the factors that may explain its predictions, including background information, academic engagement, and learning habits. Our findings may help in designing better prediction models, and to inform institution-level engagement with students vis-à-vis the notion of non-thriving, compared with the current common notion of survival.

Indeed, since the launch of this initial research in FYE, not only has it continued and expanded into Intro to Engineering, as reported here, but it has also been replicated and implemented in the Intro to Chemistry course (Schalk et al., 2021), and currently three of the biggest gateway first-year STEM courses (Calculus, Economics, and Biology) are also working on adopting this early prediction model for non-thriving. Moreover, we are working on a cross-course dashboard for advisors and deans to monitor non-thriving students across multiple courses. Such a dashboard will further and better support the important mission of helping every student to fulfill their potential. Future avenues of research will include testing the implementation of that cross-course dashboard from various stakeholder perspectives. Also, we suggest testing our non-thriving early identification model in other institutions and different educational and cultural contexts.

Importantly, we believe that our suggested shift from focusing on survival to focusing on thriving has some important implications. First, it may reduce harm to those students who might have been labelled as failures upon being predicted as potential dropouts. Second, it may support the infusion of learning analytics in institutions where dropout is not a problem, hence efforts in such technologies are supposedly redundant. Changing the terminology will make learning analytics relevant to these institutions, as the ultimate goal is now not to prevent dropout, but rather to support every enrolled student. Lastly, the focus on thriving sets a new standard for the way institutions of higher education communicate the notion of success to their students. Therefore, we truly believe that adopting this change in terminology may positively impact higher education at large.

5.3.1. Limitations

Of course, this study is not without its limitations. First, this is a qualitative study limited by the number of participants. Furthermore, it was situated at a single university, which may be characterized by a specific culture of education, technology, and implementation of technology in the educational process. Our findings should be validated by similar studies in other institutions and countries. Second, it was focusing on a single EWS, aimed at predicting the non-thriving of first-year university students in the context of an orientation-like course. Therefore, the study should also be replicated with other EWSs. Additionally, even when considering this narrowed-down point of view, the sampled population does not necessarily represent the whole user population of the studied EWS. Despite these limitations, we feel that the current study can contribute to promoting better data-driven decision-making for students. However, our policy recommendations should not necessarily be considered as immediate action items, but rather as suggestions for furthering the study in this field.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

This work was supported by the Schlindwein Family Tel Aviv University–Notre Dame Research Collaboration Grant.

References

Agrusti, F., Bonavolontà, G., & Mezzini, M. (2019). University dropout prediction through educational data mining techniques: A systematic review. *Journal of E-Learning and Knowledge Society*, *15*(3), 161–182. https://doi.org/10.20368/1971-8829/1135017

Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education*, *16*(40), 1–20. https://doi.org/10.1186/s41239-019-0172-z

Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: Literature review and best practices. *International Journal of Educational Technology in Higher Education*, *17*(3), 1–21. https://doi.org/10.1186/s41239-020-0177-7

Balfanz, R., & Byrnes, V. (2019). Early warning indicators and intervention systems: State of the field. In J. A. Fredricks, A. L. Reschly & S. L. Christenson (Eds.), *Handbook of student engagement interventions: Working with disengaged students* (pp. 45–55). Elsevier Inc. https://doi.org/10.1016/B978-0-12-813413-9.00004-8

Bartolini, A. C., Running, C. L., Duan, X., & Alex Ambrose, G. (2020). Integrated closed-loop learning analytics scheme in a first-year engineering course. Paper presented at the 2020 ASEE Virtual Annual Conference. https://doi.org/10.18260/1-2--34836

Biswas, A. A., Majumder, A., Mia, M. J., Nowrin, I., & Ritu, N. A. (2019). Predicting the enrollment and dropout of students in the post-graduation degree using machine learning classifier. *International Journal of Innovative Technology and Exploring Engineering*, *8*(11), 3083–3088. https://doi.org/10.35940/ijitee.K2435.0981119

Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2019). Predict and intervene: Addressing the dropout problem in a MOOC-based program. In *Proceedings of the 6th ACM Conference on Learning @ Scale (L@S),* 24–25 June 2019, Chicago, IL, USA (Article 24, pp. 1–9). ACM Press. https://doi.org/10.1145/3330430.3333634

Cates, J. T., & Schaefle, S. E. (2011). The relationship between a college preparation program and at-risk students' college readiness. *Journal of Latinos and Education*, *10*(4), 320–334. https://doi.org/10.1080/15348431.2011.605683

Cebesoy, U. B., & Akinoglu, O. (2012). The effectiveness of elaboration and organizational strategies in science education on students' academic achievement, attitude and concept learning. *Energy Education Science and Technology Part B: Social and Educational Studies*, (Special Issue), 212-218.

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, *18*(6), 683–695. https://doi.org/10.1080/13562517.2013.827653

Cohen, A. (2017). Analysis of student activity in web-supported courses as a tool for predicting dropout. *Educational Technology Research and Development*, *65*, 1285–1304. https://doi.org/10.1007/s11423-017-9524-3

Crompton, H. (2017). *ISTE standards for educators: A guide for teachers and other professionals*. International Society for Technology in Education. https://www.iste.org/standards/iste-standards-for-teachers

Dames, S. (2019). The interplay of developmental factors that impact congruence and the ability to thrive among new graduate nurses: A qualitative study of the interplay as students transition to professional practice. *Nurse Education in Practice*, *36*, 47–53. https://doi.org/10.1016/j.nepr.2019.02.013

Dang, S. C., & Koedinger, K. R. (2020). The ebb and flow of student engagement: Measuring motivation through temporal pattern of self-regulation. In A. N. Rafferty, J. Whitehill, C. Romero, & V. Cavalli-Sforza (Eds.), *Proceedings of the 13th International Conference on Educational Data Mining (EDM2020)*, 10–13 July 2020, Online (pp. 61–68). International Educational Data Mining Society.

Doll, J. J. (2010, May). *Teachers' and administrators' perceptions of the antecedents of school dropout among English language learners at selected Texas schools* [Unpublished doctoral dissertation]. Texas A&M University. https://core.ac.uk/download/pdf/147140442.pdf

Freeman, J. C., & All, A. (2017). Academic support programs utilized for nursing students at risk of academic failure: A review of the literature. *Nursing Education Perspectives*, *38*(2), 69–74. https://doi.org/10.1097/01.NEP.0000000000000089

García-Poole, C., Byrne, S., & Rodrigo, M. J. (2019). Implementation factors that predict positive outcomes in a community-based intervention program for at-risk adolescents. *Psychosocial Intervention*, *28*(2), 57–65. https://doi.org/10.5093/pi2019a4

Gray, C. C., & Perkins, D. (2019). Utilizing early engagement and machine learning to predict student outcomes. *Computers and Education*, *131*, 22–32. https://doi.org/10.1016/j.compedu.2018.12.006

Hawken, L. S., Bundock, K., Kladis, K., O'Keeffe, B., & Barrett, C. A. (2014). Systematic review of the check-in, check-out intervention for students at risk for emotional and behavioral disorders. *Education and Treatment of Children*, *37*(4), 635–658. https://doi.org/10.1353/etc.2014.0030

Henderson, M., Ryan, T., & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment and Evaluation in Higher Education*, *44*(8), 1237–1252. https://doi.org/10.1080/02602938.2019.1599815

Herodotou, C., Rienties, B., Verdin, B., & Boroowa, A. (2019). Predictive learning analytics 'at scale': Guidelines to successful implementation in higher education. *Journal of Learning Analytics*, *6*(1), 85–95. https://doi.org/10.18608/jla.2019.61.5

Holmes, H., Lara, A. E., & Brown, G. S. (2020). Social media use in college-age youth: A comprehensive review and a call to action. *Current Psychopharmacology*, *9*(2), 128–143. https://doi.org/10.2174/2211556009999200408112951

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, *6*(2), 27–52. https://doi.org/10.18608/jla.2019.62.3

Hout, M. (2012). Social and economic returns to college education in the United States. *Annual Review of Sociology, 38*, 379–400. https://doi.org/10.1146/annurev.soc.012809.102503

Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*, *37*, 66–75. https://doi.org/10.1016/j.iheduc.2018.02.001

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hu, Y. H., Lo, C. L., & Shih, S. P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, *36*, 469–478. https://doi.org/10.1016/j.chb.2014.04.002

Hutt, S., Ocumpaugh, J., Andres, J. M. A. L., Bosch, N., Paquette, L., Biswas, G., & Baker, R. S. (2021). Investigating SMART models of self-regulation and their impact on learning. In S. I-Han, S. Shaghayegh, F. Bouchet & J.-J. Vie (Eds.), *Proceedings of the 14th International Conference on Educational Data Mining (EDM2021)*, 29 June–2 July 2021, Online (pp. 580–587). International Educational Data Mining Society. https://educationaldatamining.org/EDM2021/EDM2021Proceedings.pdf

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, *64*, 923–938. https://doi.org/10.1007/s11423-016-9477-y

Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, *68*, 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Iorfa, S., Ugwu, C., Ifeagwazi, C. M., & Chukwuorji, J. C. (2019). Substance use among youths: Roles of psychoticism, social alienation, thriving and religious commitment. *African Journal of Drug and Alcohol Studies*, *17*(2), 133–146.

James, M. L. (2013). *Understanding the role of athletics and resiliency in the persistence and success of African American males in a community college setting* [Unpublished doctoral dissertation]. Northern Illinois University.

Jones, K. M. L. (2019). Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education*, *16*(24) 1–22. https://doi.org/10.1186/s41239-019-0155-0

Kappen, T. H., van Klei, W. A., van Wolfswinkel, L., Kalkman, C. J., Vergouwe, Y., & Moons, K. G. M. (2018). Evaluating the impact of prediction models: Lessons learned, challenges, and recommendations. *Diagnostic and Prognostic Research*, *2*(11), 1–11. https://doi.org/10.1186/s41512-018-0033-6

Kehm, B. M., Larsen, M. R., & Sommersel, H. B. (2019). Student dropout from universities in Europe: A review of empirical literature. *Hungarian Educational Research Journal*, *9*(2), 147–164. https://doi.org/10.1556/063.9.2019.1.18

Kennedy, C. B. (2017). *An analysis of perceptions of dropout factors and interventions by middle school and high school teachers in a southeastern school district* [Unpublished doctoral dissertation]. Liberty University.

Klerkx, J., Verbert, K., & Duval, E. (2017). Learning analytics dashboards. In C. Lang, G. Siemens, A. Wise, & D. Gašević (Eds.), *Handbook of learning analytics* (pp. 143–150). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.012

Knesting-Lund, K., Reese, D., & Boody, R. (2013). Teachers' perceptions of high school dropout and their role in dropout prevention: An initial investigation. *Journal of Studies in Education*, *3*(4), 57–71. https://doi.org/10.5296/jse.v3i4.4281

Kollom, K., Tammets, K., Scheffel, M., Tsai, Y. S., Jivet, I., Muñoz-Merino, P. J., Moreno-Marcos, P. M., Whitelock-Wainwright, A., Calleja, A. R., Gašević, D., Kloos, C. D., Drachsler, H., & Ley, T. (2021). A four-country cross-case analysis of academic staff expectations about learning analytics in higher education. *The Internet and Higher Education*, *49*, 1–18. https://doi.org/10.1016/J.IHEDUC.2020.100788

Kommoju, N. (2019). *The measure of our commitment is our commitment to measurement: A critical discourse analysis of the language of data in the public discourse of U.S. Secretaries of education* [Unpublished master's thesis]. Georgetown University.

Lammers, W. J., Gillaspy, J. A., & Hancock, F. (2017). Predicting academic success with early, middle, and late semester assessment of student–instructor rapport. *Teaching of Psychology*, *44*(2), 145–149. https://doi.org/10.1177/0098628317692618

Lammers, W. J., Onwuegbuzie, A. J., & Slate, J. R. (2001). Academic success as a function of the gender, class, age, study habits, and employment of college students. *Research in the Schools*, *8*(2), 71–81.

Larsen, A., Horvath, D., & Bridge, C. (2019). 'Get ready': Improving the transition experience of a diverse first year cohort through building student agency. *Student Success*, *11*(2), 14–27. https://doi.org/10.5204/ssj.v11i3.1144

Lawson, C., Beer, C., Rossi, D., Moore, T., & Fleming, J. (2016). Identification of 'at risk' students using learning analytics: The ethical dilemmas of intervention strategies in a higher education institution. *Educational Technology Research and Development*, *64*, 957–968. https://doi.org/10.1007/s11423-016-9459-0

Lehmann, A. C. (2014). Using admission assessments to predict final grades in a college music program. *Journal of Research in Music Education*, *62*(3), 245–258. https://doi.org/10.1177/0022429414542654

Lim, L.-A., Dawson, S., Gašević, D., Joksimović, S., Pardo, A., Fudge, A., & Gentili, S. (2020). Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: An exploratory study of four courses. *Assessment and Evaluation in Higher Education*, *46*(3), 339–359. https://doi.org/10.1080/02602938.2020.1782831

Liz-Domínguez, M., Caeiro-Rodríguez, M., Llamas-Nistal, M., & Mikic-Fonte, F. (2019). Predictors and early warning systems in higher education: A systematic literature review. In, M. Caeiro-Rodríguez, Á. Hernández-García & P. J. Muñoz-Merino (Eds.), *Proceedings of the Learning Analytics Summer Institute Spain 2019* (LASI-SPAIN 2019), 27–28 June 2019, Vigo, Spain (pp. 84–99).

Lizzio, A., & Wilson, K. (2013). Early intervention to support the academic recovery of first-year students at risk of non-continuation. *Innovations in Education and Teaching International*, *50*(2), 109–120. https://doi.org/10.1080/14703297.2012.760867

Ma, J., Pender, M., & Welch, M. (2019). *Education pays 2019: The benefits of higher education for individuals and society*. CollegeBoard. https://research.collegeboard.org/media/pdf/education-pays-2019-full-report.pdf

Macfarlane, B. (2015). Student performativity in higher education: Converting learning as a private space into a public performance. *Higher Education Research and Development*, *34*(2), 338–350. https://doi.org/10.1080/07294360.2014.956697

Manrique, R., Nunes, B. P., Marino, O., Casanova, M. A., & Nurmikko-Fuller, T. (2019). An analysis of student representation, representative features and classification algorithms to predict degree dropout. In *Proceedings of the 9th International Conference on Learning Analytics and Knowledge (LAK '19)*, 4–8 March 2019, Tempe, AZ, USA (pp. 401–410). ACM Press. https://doi.org/10.1145/3303772.3303800

Mastrodicasa, J., & Metellus, P. (2013). The impact of social media on college students. *Journal of College and Character*, *14*(1), 21–30. https://doi.org/10.1515/jcc-2013-0004

Masui, C., Broeckmans, J., Doumen, S., Groenen, A., & Molenberghs, G. (2014). Do diligent students perform better? Complex relations between student and course characteristics, study time, and academic performance in higher education. *Studies in Higher Education*, *39*(4), 621–643. https://doi.org/10.1080/03075079.2012.721350

McMahon, B. M., & Sembiante, S. F. (2020). Re-envisioning the purpose of early warning systems: Shifting the mindset from student identification to meaningful prediction and intervention. *Review of Education*, *8*(1), 266–301. https://doi.org/10.1002/rev3.3183

Michaeli, S., Kroparo, D., & Hershkovitz, A. (2020). ~~View of~~ Teachers' use of education dashboards and professional growth. *International Review of Research in Open and Distributed Learning*, *21*(4), 61–78. https://www.irrodl.org/index.php/irrodl/article/view/4663/5416

Molenaar, I., Horvers, A., Dijkstra, R., & Baker, R. S. (2020). Personalized visualizations to promote young learners' SRL: The learning path app. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20)*, 23–27 March 2020, Frankfurt, Germany (pp. 330–339). ACM Press. https://doi.org/10.1145/3375462.3375465

Molenaar, I., & Knoop-van Campen, C. A. N. (2019). How teachers make dashboard information actionable. *IEEE Transactions on Learning Technologies*, *12*(3), 347–355. https://doi.org/10.1109/TLT.2018.2851585

Na, K. S., & Tasir, Z. (2017). Identifying at-risk students in online learning by analysing learning behaviour: A systematic review. In *Proceedings of the 2017 IEEE International Conference on Big Data Analysis (ICBDA 2017)*, 16–17 November, Kuching, Malaysia (pp. 118–123). https://doi.org/10.1109/ICBDAA.2017.8284117

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2017). The possibility of predicting learning performance using features of note taking activities and instructions in a blended learning environment. *International Journal of Educational Technology in Higher Education*, *14*(6), 1–14. https://doi.org/10.1186/s41239-017-0048-z

Nik Nurul Hafzan, M. Y., Safaai, D., Asiah, M., Mohd Saberi, M., & Siti Syuhaida, S. (2019). Review on predictive modelling techniques for identifying students at risk in university environment. In L. M. Hee (Ed.), *Proceedings of the Engineering Application of Artificial Intelligence Conference 2018 (EAAIC 2018)*, 3–5 December 2018, Kota Kinabalu, Malaysia (MATEC Web of Conferences 255, Article 03002, pp. 1–8). EDP Sciences. https://doi.org/10.1051/matecconf/201925503002

Ocumpaugh, J., Baker, R. S., Pedro, M. O. C. Z. S., Hawn, M. A., Heffernan, C., Heffernan, N., & Slater, S. A. (2017). Guidance counselor reports of the ASSISTments college prediction model (ACPM). In *Proceedings of the 7th International Conference on Learning Analytics and Knowledge (LAK '17)*, 13–17 March 2017, Vancouver, BC, Canada (pp. 479–488). ACM Press. https://doi.org/10.1145/3027385.3027435

Oh, P. S., & Oh, S. J. (2011). What teachers of science need to know about models: An overview. *International Journal of Science Education*, *33*(8), 1109–1130. https://doi.org/10.1080/09500693.2010.502191

O'Neill, L., Hartvigsen, J., Wallstedt, B., Korsholm, L., & Eika, B. (2011). Medical school dropout – Testing at admission versus selection by highest grades as predictors. *Medical Education*, *45*(11), 1111–1120. https://doi.org/10.1111/j.1365-2923.2011.04057.x

O'Shea, S., Lysaght, P., Roberts, J., & Harwood, V. (2016). Shifting the blame in higher education: Social inclusion and deficit discourses. *Higher Education Research and Development*, *35*(2), 322–336. https://doi.org/10.1080/07294360.2015.1087388

Owen, J. P. (2009). *Teachers' views of what schools are lacking in dropout prevention* [Unpublished master's thesis]. California State University.

Owens, H., Christian, B., & Polivka, B. (2017). Sleep behaviors in traditional-age college students: A state of the science review with implications for practice. *Journal of the American Association of Nurse Practitioners*, *29*(11), 695–703. https://doi.org/10.1002/2327-6924.12520

Paura, L., & Arhipova, I. (2014). Cause analysis of students' dropout rate in higher education study program. *Procedia: Social and Behavioral Sciences*, *109*, 1282–1286. https://doi.org/10.1016/j.sbspro.2013.12.625

Pérez, L. X. (1998). Sorting, supporting, connecting, and transforming: Intervention strategies for students at risk. *Community College Review*, *26*(1), 63–78. https://doi.org/10.1177/009155219802600105

Plak, S., Cornelisz, I., Meeter, M., & van Klaveren, C. (2022). Early warning systems for more effective student counseling in higher education: Evidence from a Dutch field experiment. *Higher Education Quarterly, 76*(1)*,* 131-152. https://doi.org/10.1111/hequ.12298

Richman, E. L. (2013). *The academic success of college students with attention deficit hyperactivity disorder and learning disabilities* [Doctoral dissertation]. University of North Carolina. https://doi.org/10.17615/e791-xx31

Rovira, S., Puertas, E., & Igual, L. (2017). Data-driven system to predict academic grades and dropout. *PLoS ONE*, *12*(2), 1–21. https://doi.org/10.1371/journal.pone.0171207

Salehian Kia, F., Hatala, M., Baker, R. S., & Teasley, S. D. (2021). Measuring students' self-regulatory phases in LMS with behavior and real-time self report. In *Proceedings of the 11th International Conference on Learning Analytics and Knowledge (LAK '21)*, 12–16 April 2021, Irvine, CA, USA (pp. 259–268). ACM Press. https://doi.org/10.1145/3448139.3448164

Schalk, C. N., W., Duan, X., Woodard, V. W., Young, K. M. H. & Ambrose, G. A. (2021). Visual and predictive analytics for early identification of non-thriving students in an introductory chemistry course (Poster presentation). *Midwest Scholarship of Teaching & Learning (SoTL) Annual Conference*. Virtual.

Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, *64*, 939–955. https://doi.org/10.1007/s11423-016-9458-1

Seaver, W. B., & Quarton, R. J. (1976). Regression discontinuity analysis of dean's list effects. *Journal of Educational Psychology*, *68*(4), 459–465. https://doi.org/10.1037/0022-0663.68.4.459

Selvig, D., Holaday, L. W., Purkiss, J., & Hortsch, M. (2015). Correlating students' educational background, study habits, and resource usage with learning success in medical histology. *Anatomical Sciences Education*, *8*(1), 1–11. https://doi.org/10.1002/ase.1449

Seppälä, E. M., Bradley, C., Moeller, J., Harouni, L., Nandamudi, D., & Brackett, M. A. (2020). Promoting mental health and psychological thriving in university students: A randomized controlled trial of three well-being interventions. *Frontiers in Psychiatry*, *11*(Article 590), 1–14. https://doi.org/10.3389/fpsyt.2020.00590

Stiles Hanlon, A. (2018). *Athletes to scholars: Campus ethos and non-cognitive factors that matter to student-athletes at California community colleges* [Doctoral dissertation]. California State University. http://hdl.handle.net/20.500.12680/rv042t75d

Strauss, V. (2019, January 23). Why we should stop labeling students as 'at risk'—and the best alternative. *The Washington Post*. https://www.washingtonpost.com/education/2019/01/23/why-we-should-stop-labeling-students-risk-best-alternative/

Suardiaz-Muro, M., Morante-Ruiz, M., Ortega-Moreno, M., Ruiz, M. A., Martín-Plasencia, P., & Vela-Bueno, A. (2020). Sueño y rendimiento académico en estudiantes universitarios: Revisión sistemática [Sleep and academic performance in university students: A systematic review]. *Revista de Neurologia*, *71*(2), 43–53. https://doi.org/10.33588/RN.7102.2020015

Sun, K., Mhaidli, A. H., Watel, S., Brooks, C. A., & Schaub, F. (2019). It's my data! Tensions among stakeholders of a learning analytics dashboard. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*, 4–9 May 2019, Glasgow, Scotland, UK (Paper No. 594, pp. 1–14). ACM Press. https://doi.org/10.1145/3290605.3300824

Syed, M., Anggara, T., Lanski, A., Duan, X., Ambrose, G. A., & Chawla, N. V. (2019). Integrated closed-loop learning analytics scheme in a first year experience course. In *Proceedings of the 9th International Conference on Learning Analytics and Knowledge (LAK '19)*, 4–8 March 2019, Tempe, AZ, USA (Article 4, pp. 521–530). ACM Press. https://doi.org/10.1145/3303772.3303803

Tamada, M. M., de Magalhães Netto, J. F., & de Lima, D. P. R. (2019). Predicting and reducing dropout in virtual learning using machine learning techniques: A systematic review. In *Proceedings of the 2019 IEEE Frontiers in Education Conference (FIE)*, 16–19 October 2019, Covington, KY, USA (pp. 1–9). IEEE Computer Society. https://doi.org/10.1109/FIE43999.2019.9028545

Taylor, D. B., & Harrison, G. J. (2018). Supporting biomedical students struggling with second-choice-syndrome to thrive rather than just survive first year. *Journal of College Student Retention: Research, Theory and Practice*, *20*(2), 176–196. https://doi.org/10.1177/1521025116654162

Torenbeek, M., Jansen, E., & Suhre, C. (2013). Predicting undergraduates' academic achievement: The role of the curriculum, time investment and self-regulated learning. *Studies in Higher Education*, *38*(9), 1393–1406. https://doi.org/10.1080/03075079.2011.640996

Tsai, Y.-S., Rates, D., Moreno-Marcos, P. M., Muñoz-Merino, P. J., Jivet, I., Scheffel, M., Drachsler, H., Delgado Kloos, C., & Gašević, D. (2020). Learning analytics in European higher education: Trends and barriers. *Computers and Education*, *155*, 103933. https://doi.org/10.1016/j.compedu.2020.103933

Tsai, Y.-S., Moreno-Marcos, P. M., Jivet, I., Scheffel, M., Tammets, K., Kollom, K., & Gašević, D. (2018). The SHEILA framework: Informing institutional strategies and policy processes of learning analytics. *Journal of Learning Analytics*, *5*(3), 5–20. https://doi.org/10.18608/jla.2018.53.2

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2021). More than figures on your laptop: (Dis)trustful implementation of learning analytics. *Journal of Learning Analytics*, *8*(3), 81–100. https://doi.org/10.18608/jla.2021.7379

Tsiakmaki, M., Kostopoulos, G., Koutsonikos, C., Pierrakeas, S. & Ragos, O. (2018). Predicting university students' grades based on previous academic achievements. In *Proceedings of the 9th International Conference on Information, Intelligence, Systems and Applications (IISA)*, 23–25 July 2018, Zakynthos, Greece (pp. 1–6). IEEE Computer Society. https://doi.org/10.1109/IISA.2018.8633618

Valentine, J. C., Hirschy, A. S., Bremer, C. D., Novillo, W., Castellano, M., & Banister, A. (2011). Keeping at-risk students in school: A systematic review of college retention programs. *Educational Evaluation and Policy Analysis*, *33*(2), 214–234. https://doi.org/10.3102/0162373711398126

van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *International Journal of Computer-Supported Collaborative Learning*, *14*, 261–289. https://doi.org/10.1007/s11412-019-09299-x

Vezzoli, Y., Mavrikis, M., & Vasalou, A. (2020). Inspiration cards workshops with primary teachers in the early co-design stages of learning analytics. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20),* 23–27 March 2020, Frankfurt, Germany (pp. 73–82). ACM Press. https://doi.org/10.1145/3375462.3375537

Vuorikari, R., & Castaño Muñoz, J. (Eds.), Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., & Vuorikari, R. (2016). Research evidence on the use of learning analytics: Implications for education policy, EUR 28294. Publications Office of the European Union. https://doi.org/10.2791/955210

Waddington, R. J., Nam, S., Lonn, S., & Teasley, S. D. (2016). Improving early warning systems with categorized course resource usage. *Journal of Learning Analytics*, *3*(3), 263–290. https://doi.org/10.18608/jla.2016.33.13

Wagner, P., Schober, B., & Spiel, C. (2008). Time students spend working at home for school. *Learning and Instruction*, *18*(4), 309–320. https://doi.org/10.1016/j.learninstruc.2007.03.002

Weinstein, C. E. (1982). Training students to use elaboration learning strategies. *Contemporary Educational Psychology*, *7*(4), 301–311. https://doi.org/10.1016/0361-476X(82)90013-3

Whitman, M. (2020). "We called that a behavior": The making of institutional data. *Big Data and Society*, *7*(1). https://doi.org/10.1177/2053951720932200

Williams, L., Titus, K. J., & Pittman, J. M. (2021). How early is early enough: Correlating student performance with final grades. In *Proceedings of the Conference on Computing Education Practice 2021 (CEP '21)*, 7 January 2021, Durham, UK (pp. 13–16). ACM Press. https://doi.org/10.1145/3437914.3437972

Willis, J. E., Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: A typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development*, *64*, 881–901. https://doi.org/10.1007/s11423-016-9463-4

Woods-Weeks, G. (2017). *Graduates of an early college high school: Perceptions of college readiness* [Doctoral dissertation]. East Carolina University. https://hdl.handle.net/10342/6497

Wright, N. A. (2020). Perform better, or else: Academic probation, public praise, and students decision-making. *Labour Economics*, *62*, 101773. https://doi.org/10.1016/j.labeco.2019.101773

Yukselturk, E., Ozekes, S., & Türel, Y. K. (2018). Predicting dropout student: An application of data mining methods in an online education program. *European Journal of Open, Distance and E-Learning*, *17*(1), 118–133. https://doi.org/10.2478/eurodl-2014-0008

Zabalou, S. J. (2021). Perceptions of high school teachers on factors impacting African American male dropout [Unpublished doctoral dissertation]. St. Thomas University.

Zhang, Y., Fei, Q., Quddus, M., & Davis, C. (2014). An examination of the impact of early intervention on learning outcomes of at-risk students. *Research in Higher Education Journal*, *26*, pp. 1–12.