


RESEARCH

Open Access



i-Ntervene: applying an evidence-based learning analytics intervention to support computer programming instruction

Piriya Utamachant^{1*} , Chutiporn Anutariya¹ and Suporn Pongnumkul²

*Correspondence:
piriya.utamachant@ait.ac.th

¹ICT Department, School of Engineering and Technology, Asian Institute of Technology, Pathum Thani, Thailand

²National Electronics and Computer Technology Center, Pathum Thani, Thailand

Abstract

Apart from good instructional design and delivery, effective intervention is another key to strengthen student academic performance. However, intervention has been recognized as a great challenge. Most instructors struggle to identify at-risk students, determine a proper intervention approach, trace and evaluate whether the intervention works. This process requires extensive effort and commitment, which is impractical especially for large classes with few instructors. This paper proposes a platform, namely *i-Ntervene*, that integrates Learning Management System (LMS) automatic code grader, and learning analytics features which can empower systematic learning intervention for large programming classes. The platform supports instructor-pace courses on both Virtual Learning Environment (VLE) and traditional classroom setting. The platform iteratively assesses student engagement levels through learning activity gaps. It also analyzes subject understanding from programming question practices to identify at-risk students and suggests aspects of intervention based on their lagging in these areas. Students' post-intervention data are traced and evaluated quantitatively to determine effective intervention approaches. This evaluation method aligns with the evidence-based research design. The developed *i-Ntervene* prototype was tested on a Java programming course with 253 first-year university students during the Covid-19 pandemic in VLE. The result was satisfactory, as the instructors were able to perform and evaluate 12 interventions throughout a semester. For this experimental course, the platform revealed that the approach of sending extrinsic motivation emails had more impact in promoting learning behavior compared to other types of messages. It also showed that providing tutorial sessions was not an effective approach to improving students' subject understanding in complex algorithmic topics. *i-Ntervene* allows instructors to flexibly trial potential interventions to discover the optimal approach for their course settings which should boost student's learning outcomes in long term.

Keywords: Learning intervention platform, Learning analytics intervention, Learning intervention, Instructional intervention, Evidence based interventions, Intervention in programming education

Introduction

High attrition rate has been acknowledged as a significant challenge in tertiary education (Aina et al., 2022; Barbera et al., 2020; OECD, 2019). A large instructor-pace course usually suffers from: (i) students with insufficient academic backgrounds who struggle catching up, and (ii) students who lack self-regulation learning skills and who present inattentive learning behavior, resulting in low performance and eventually dropping out from the program (Rodriguez-Planas, 2012). The problem is most obvious for students who are advancing or transitioning to a new curriculum. Among those, the failure of the first programming course in an undergraduate computer science curriculum, known as CS1, has been recognized as a significant situation (Azcona et al., 2019). Firstly, a survey study in 2007 reported a 33% CS1 failure rate worldwide (Bennedsen & Caspersen, 2007). This finding concurred with a study in 2014 which evaluated data extracted from relevant CS1 articles (Watson & Li, 2014). Later, another survey study conducted in 2017 reported a slightly improved rate at 28% CS1 failure rate worldwide (Bennedsen & Caspersen, 2019) which remained high. Recent research studies within the community have continued to focus on the issue of CS1 attrition, indicating the persistence of this situation (Obiado et al., 2023; Takacs et al., 2022). In Thailand, the increasing of failure rate of programming courses in undergrad curriculums is also a well-known issue and has been discussed among universities' lecturers. As an example, the particular CS2 course which we experimented on has records of more than 50% failure rate in the past 5 years.

Educators have continued their efforts to address the issue. Recent studies have broadened their scope to analyze potential root causes and introduce novel approaches aimed at improving student's learning, which could lead to an increase in success rates. Other interesting research includes analyzing impact of students' demographics and background (Layman et al., 2020; Stephenson et al., 2018), investigating students' learning motivations (Aivaloglou & Hermans, 2019; Barr & Kallia, 2022; Loksa et al., 2022; Santana et al., 2018), predicting and identifying students who have high potential of non-progression (Cobos & Ruiz-Garcia, 2020; Tempelaar et al., 2020), and analyzing problematic learning behaviors (Nam Liao et al., 2021; Salguero et al., 2021).

Among various approaches, Learning intervention is a highly effective methodology composing of multiple research areas. Learning intervention plays an important role in aiding at-risk or underachieving students before they fail. It can provide support to maintain their motivation, improve behavioral disorders, and promote better learning outcomes (Szabo et al., 2017; Kew & Tasir, 2017; Sacr et al., 2018; Sclater & Mullan, 2017; Dvorak & Jia, 2016; Hsu et al., 2016). Successful learning intervention relies on instructors regularly assessing students' situations and performing proper learning intervention in time (Othman et al., 2011). However, intervention is a great challenge for most instructors due to several factors (Wong & Li, 2020). First, a huge amount of complex data with limited competent personnels make it difficult to analyze, monitor and identify learning issues in a timely manner especially for courses with low instructor resources (Neil et al., n.d.; Sonderlund et al., 2018; Zhang et al., 2018; Gašević et al., 2016; Kennedy et al., 2014). An effectiveness evaluation of the applied interventions is even more demanding, as the process requires a complicated experimental design which is usually unachievable in most common courses (Huang et al., 2016; Szabo et al., 2017; Wong,

2017). An important gap is the lack of a comprehensive model with a strong evidence-based approach to support instructors for such process (Rienties et al., 2017).

With the use of LMS and automatic code grader in modern programming education, most evidences of student learning are recorded in real-time. Appropriate analytics techniques could evaluate this data to identify student's deficiencies in subject understanding and learning activities and then visualize reports that support instructors to perform precise intervention in timely manner. Students' learning data after applied intervention can be analyzed and compared with pre-intervention data to evaluate effectiveness of the intervention aligning with quasi-experimental design which complies with Evidence-Based Intervention (EBI) guidelines (U.S. Department of Education, 2016). Based on this concept, the paper proposes an integrated platform of an LMS, an automatic code grader and learning analytics features that can facilitate effective intervention processes for programming courses, named *i-Ntervene*. The platform systematically collects and assesses students' learning engagement and subject understanding deficiencies to identify students who are at risk of failing the course or dropping out and suggests their lagging aspects which guides instructor's intervention decision. Afterward, *i-Ntervene* statistically evaluates the applied intervention results using pre and post data. These features should boost students' learning outcomes for large programming courses by empowering a small number of instructors to implement a complete loop of instructional interventions throughout the semester.

This research is organized into 4 main sections. The first section provides an overview context of the research problem and introduces a solution proposal. Section [Related work](#) describes the relevant background knowledge and literature. Section [Learning analytics intervention](#) introduces a model of instructional intervention cycle and proposes a platform, *i-Ntervene*, that accommodates instructors to achieve such effective intervention. Section [LA intervention methods and evaluation](#) demonstrates an experiment of *i-Ntervene* prototype on an actual programming course along with results and findings. Finally, Sect [Evidence-based interventions \(EBI\)](#), presents conclusions and implications for future research.

Related work

Learning analytics intervention

Learning Analytics (LA) refers to the systematic process of collecting, analyzing, and interpreting data derived from educational settings and activities with the aim of extracting valuable insights. These insights are then used to make well-informed decisions that enhance learning outcomes and educational achievements (Siemens & Gasevic, 2012). Chatti et al. (2012) proposed a reference model for LA that presents a practical perspective based on four dimensions:

- (i) Environment, and Context: This dimension focuses on the types of data collected, managed, and utilized for analysis, considering the specific educational environment and context.
- (ii) Stakeholders: The stakeholders targeted by the analysis include students, teachers, institutions, and researchers, each with their own unique goals and expectations.

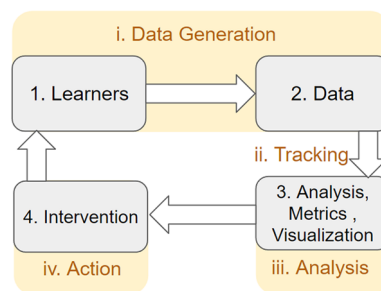


Fig. 1 Learning analytics (LA) cycle for effective Intervention proposed by Doug Clow (1–4), and Khahil & Ebner (i–iv)

- (iii) Objectives of LA: The objectives of LA involve making informed decisions about teaching and learning based on the specific focus of the stakeholders. These objectives can encompass various aspects such as monitoring, analysis, prediction, intervention, tutoring, mentoring, assessment, feedback, adaptation, personalization, recommendation, and reflection.
- (iv) Analysis Techniques or Methods: This dimension encompasses the techniques and methods employed for data analysis, including statistics, information visualization, data mining, and social network analysis.

Learning intervention refers to structured methods for identifying areas of weakness and helping students increase academic proficiency (Lynch, 2019). With the adoption of LMS, fine-grained student interactions were recorded, allowing Learning Analytics (LA) techniques to unfold their learning behaviors and cognition (Bodily & Verbert, 2017), which greatly enhance the learning interventions. The common goal of LA Intervention is to increase student academic success (Wong & Li, 2020; Kew & Tasir, 2017; Heikkinen et al., 2022). It highly influences student success and has been widely used in classroom instruction to improve and foster learning skills in both behavioral and performance areas (Heikkinen et al., 2022; Szabo et al., 2017).

Clow (2012) introduced an LA cycle model aiming at facilitating effective interventions. This model closely aligns with the one proposed by Khalil and Ebner (2015). Figure 1 depicts the conformity of the 2 models. The LA Intervention process in the model begins with data collection in accordance with the defined objectives (1–2). A research by Wong and Li (2020) reported that the most frequently used data is learning behaviors, e.g., learning activities, forum discussion, and study performance, followed by data obtained from surveys and related sources, e.g., demographic information, learning psychological, and academic history. Subsequently, the collected data is analyzed based on defined metrics, followed by visualizations for the stakeholders (3). The stakeholders then utilize this information to make informed decisions and implement intervention actions (4) that target enhancements of cognitive abilities, learning psychological aspects, and learning behaviors to ultimately improve learning success.

A valuable application of LA involves the detection of At-Risk Students (ARS) who are likely to fail or drop out of a course or program. This allows for preemptive interventions to be implemented. A notable successful case is the implementation

of an *Early Warning and Response System* introduced by Johns Hopkins University.¹ Through multiple longitudinal analyses across multiple states, this research established a set of robust indicators known as the *ABC indicators*—Attendance, Behavior, and Course performance (Balfanz et al., 2019) to identify ARS. International organizations like UNESCO and OECD have also promoted various early warning systems, with reports of successful implementation in the United States and many other countries (Bowers, 2021; UNESCO, 2022).

Recent systematic review papers (Ifenthaler & Yau, 2020; Wong & Li, 2020) discussed Learning Analytic Intervention research conducted over the past decade. These studies concluded that LA Interventions have a positive impact on student achievement. The key intervention practices for students involve the provision of feedback, recommendations, and visualization of learning data to enhance their self-awareness and self-regulation. On the other hand, for instructors, the key practices aim at identifying at-risk students and visualizing learning data, which can support their decision-making process to implement appropriate intervention actions and fine-tune their instructional delivery. The utilization of data-driven LA enables precise interventions which can lead to positive outcomes such as improved learning and teaching effectiveness, increased student participation and engagement, and enhanced awareness of progress and performance. As a result, these outcomes play a significant role in fostering student success and achievement in the learning process.

LA intervention methods and evaluation

In practice, there is no universal intervention method or strategy that fits every circumstance. Intervention methods should target specific students and contexts according to the significance and urgency of the problem. Effort, ease of use and scalability are also important factors to consider (Choi et al., 2018; Kizilcec et al., 2020). Frequent methods used included visualizations, personalized reports and recommendations via email or text message while individual contact such as face-to-face meetings and phone calls were rarely used (Heikkinen et al., 2022; Wong & Li, 2020; Kew & Tasir, 2017). In computer education, a literature review on approaches to teaching introductory programming during from 1980–2014 (Hellas et al., 2014) reported 5 intervention categories, i.e., collaboration and peer support, bootstrapping practices, relatable content and contextualization, course setup assessment and resourcing, hybrid approaches. The authors found interventions increase pass rates by 30% on average but no statistically significant differences between methods.

In terms of intervention effectiveness evaluation, qualitative methods, such as interviews, think-aloud and focus groups were used to verify the usability and students' perception, while quantitative methods such as trace data and pre-post intervention surveys were usually used to evaluate the efficiency (Heikkinen et al., 2022). However, some research results are questionable with significant limitations including simple evaluation designs, convenience sampling and small study populations

¹ Everyone Graduates Center, Johns Hopkins University. (n.d.). Early Warning and Response Systems, <https://new.every1graduates.org/tools-and-models/early-warning-and-response-systems/>

(Sonderlund et al., 2018). A solid evaluation requires evidence-based research with a proper LA method and experimental design (Araka et al., 2020; Rienties et al., 2017).

Evidence-based interventions (EBI)

Evidence-Based Interventions (EBI) are practices or programs with emphasis on evidence to show that the interventions are effective at producing accurate results and improving outcomes when implemented. The U.S. Department of Education (DoE) published a guideline encouraging the use of EBI to help educators successfully choose and implement interventions that improve student outcomes (U.S. Department of Education, 2016).

The guideline advises instructors to first understand where students are in their learning process. This may involve a review of available historical evidence, pre-tests, questionnaires and interviews. For the next step, evidence of students' learning should be considered to determine what learning strategies and interventions are likely to improve student understanding and skills. Finally, the progress students make in their learning over time is monitored in order to evaluate the effectiveness of teaching strategies and interventions (Masters, 2018).

In the instructional domain, 'evidence' usually refers to educational research or theories, but it may also extend beyond the 'scientific' (Biesta, 2010), which renders some forms of evidence as better than others. Different forms of evidence have been ranked based on trustworthiness. The hierarchy of evidence proposed by Triana (2008) rates Randomized Controlled Trials (RCT) as the highest level, followed by quasi-experimental study, pre-post comparison, and cross-sectional study. Other methodologies such as case studies, professional or expert testimonies and user opinions were also considered evidence but with less significance.

To be more practical, a study by Rienties et al. (2017) proposed 5 approaches that provide evidence of LA Intervention impact, each with its own strengths and limitations as presented in Table 1.

Table 1 Evidence-based experimental approaches on LA intervention

Experimental approaches	Strengths	Limitations
Randomized control trial	Ideal experimental design providing the most reliable evidence	Requires huge amount of effort Risk of inflicting ethical issues
A/B testing	Useful to compare impact of different types of interventions on the same objective	Requires certain effort in designing and implementation May raise concerns on ethical issue
Comparison with previous implementation	A natural way to compare impact of interventions as most courses tend to repeat with similar instructional settings. This approach has minimal ethical concern	Delayed result Test subjects (students) are different Difficult to preserve the same setting and environment between implementations
Quasi-experiment	Easy to implement and evaluate by comparing pre and post intervention. No ethical concern	The result is not as reliable as RCT
Switching replications design	An adaptation of Quasi-experiment. Compare results of 2 groups applying the same intervention switching role of experimental and control group	Students' attitude may already have changed by the time of switching the intervention

Educational practice is more than a set of interventions. Instructors should also include their professional judgement with research evidence as necessary when making decisions (Kvernbekk, 2015). In other words, evidence-based practice does not exclude an instructor's practical experience or knowledge, nor the beliefs that teachers hold based on anecdotal evidence (Kiemer & Kollar, 2021).

Challenges of LA intervention

There are major challenges of LA Intervention stated in recent research (Rienties et al., 2017; Sonderlund et al., 2018; Wong & Li, 2020). These include:

1. *Variability and complexity of courses and students*: The combinations of these factors make the situation rather unique and render a 'best-practice' intervention infeasible.
2. *Conditions for implementation*: Particular conditions can limit the implementation such as data availability, channel of access to students, and instructor's previous experience.
3. *Scalability*: Certain LA Interventions are effective for a certain size of class, pedagogy, and subject content. Changes of these factors may cause an otherwise successful intervention to not achieve the expected result.
4. *Timeliness of intervention*: An intervention needs to be applied in time when the problem is raised, which means data collection and analysis process must be swift, accurate and meaningful.
5. *Effectiveness Evaluation*: A solid evaluation requires evidence-based research which involves appropriate design and effort to implement. Without proper evaluation, the impact of the intervention is indecisive and may lead to uncertainty of the effectiveness when replicating with different courses or environments.

The Challenges 1 to 3 rely on instructors to carefully choose interventions based on their course conditions, and problems at hand. However, to choose and implement an intervention in time (Challenge 4), an instructor requires up-to-date analytical data during course instruction in order to identify ARS. After applying intervention, the instructor also needs to know how well it works based on reliable evidence in order to make adjustments or make a further decision (Challenge 5). The 2 latter procedures require excessive efforts which are infeasible to execute iteratively in a common course without supporting tools. The Open University's publication (Rienties et al., 2017) addresses the need to develop an evidence-based framework for LA to determine the effectiveness of an intervention and its conditions. To the best of our knowledge, the research community still lacks such a supporting system.

LA Intervention systems and tools

Research on systems and tools which accommodate LA Interventions by analyzing temporal leaning engagement data have been growing. Table 2 highlights recent research and summarizes their key characteristics including the proposed methods, the implemented system/platform as well as the intervention evaluation.

As depicted by Table 2, most research and tools aim to provide meaningful information to encourage students, support instructors in making intervention decisions, and

Table 2 Recent LA Intervention research with temporal student engagement analysis

Research	Context	ARS identification methods proposed/used	Platform implementation	Intervention/effectiveness evaluation
Tempelaar et al. (2020)	Statistic Course 1027 students	Analyze temporal students' engagement (#Attempt, #Examples, #Hints, #Views, Time spent) by analyzing 3 time-phases after delivered each topic: 1) prepare for tutorial, 2) prepare for quiz, and 3) prepare for exam. Students who exhibited lower engagement in the early phases can be identified as ARS	Not available	Not available
Foster and Siddle (2019)	Large scale, Open university	Send alert messages to students who had been inactive on the platform every 14 days. The number of alert messages generated by each student served as an effective indicator for identifying at-risk students	Student dashboard	Alert message to student & personal tutor/Number of Alert messages has high prediction on fail students
Cobos and Ruiz-Garcia (2020)	Web application course, 84 students	Cluster students weekly based on engagements (#Question, #VDO, #Attendance, Time on course) to identify ARS. Additional feature for instructors to record student interventions for tracking	LMS integration (edX)	Email/Student Interview with positive feedback
Majumda et al. (2019)	Reading courses, 3 universities	Calculate score from student engagements (#Event, Total Time, Complete Rate, Unique Week/Day), classify students to 3 groups. The weaker group is identified as at-risk students	Student, instructor Dashboard, xAPI	Email/No evaluation reported
Herodutou et al. (2019)	Large scale, Open university	Utilize machine learning methods with demographic and student engagement data (#access to forum, content, resource, wiki, glossary) to identify ARS that likely not submit the next assessment. The result reveals that students under the instructors that used the system gained higher scores	Instructor Dashboard	Upon instructor decision/No evaluation reported

Table 2 (continued)

Research	Context	ARS identification methods proposed/used	Platform implementation	Intervention/effectiveness evaluation
Şahin and Yurdugül (2019)	Computer Network Course, 79 students	Visualize student's learning engagement and rank on the dashboard to enhance self-awareness and motivations. No ARS identification. Student interview results indicated the system is useful	Student dashboard LMS plugin (Moodle)	Not available
Azcona et al. (2019)	CS1 & CS2 course, 266 students	Weekly predict ARS using Machine Learning methods. The prediction features include demographics (Age, Location, School GPA, previous course grade) and weekly engagements (%SolvedQuestion, #Lab work, #Attendance, Time spent, #Material access, Weekday access)	Customized VLE	Weekly message based on student engagement status/Evaluation based on score comparison from summative assessments
Froissard et al. (2015)	Large scale, Public university	Instructors can customize engagement parameters to identify ARS including. The parameters include assessment activity, forum activity, gradebook and login. The system calculates risk score and display on instructor dashboard	Instructor dashboard LMS plugin (Moodle)	Email/No evaluation reported
F-Nervene (This research)	CS2 course, 253 students	Iteratively identify ARS in 2 aspects, learning engagement (#Attendance, #In-class activities, #Assignment, #Supplementary material access) and subject understanding (Assignment score) by analyzing temporal gap between what student performed and instructor's expectation. The system provides temporal gap visualization in both individual students and class level	Instructor dashboard	Upon instructor decision/Intervention tracking/Systematic evaluation based on improvement of temporal engagement and assessment scores

facilitate communication between instructors and students. However, the main challenges arise from the need for different intervention approaches in courses with unique settings, which could change over time. Instructors must iteratively implement interventions and evaluate them based on evidence, but resource limitations and a large student population make this difficult. This creates a gap in the research community, as there is no practical system to support instructors in overcoming these challenges. This paper proposes *i-Ntervene*, an integration of LMS, automatic code grader, and learning analytics. It utilizes students’ temporal learning engagements and subject understandings to identify at-risk students, track their progress, and evaluate the effectiveness of interventions. This empowers instructors to conduct comprehensive interventions and adapt their approaches to each unique situation.

i-Ntervene: a platform for instructional intervention in programming courses

A successful LA Intervention design requires a sound pedagogical approach along with appropriate LA cycles to provide feedback for both cognitive and behavioral processes (Sedrakyan et al., 2020). In course instruction, intervention is known to be an ongoing process and should be developed iteratively in order to find the most effective way to enhance student’s learning (Sonnenberg & Bannert, 2019). This paper firstly introduces a model of instructional intervention series, depicted in Fig. 2, based on the LA cycle model, and then proposes *i-Ntervene* which is an integrated platform that reinforces effective intervention cycles for instructor-led programming courses, as illustrated in Fig. 3.

For the instructional intervention cycle, two types of data are collected as evidence in each cycle: (a) Learning behaviors, such as class attendance, class engagement, assignment practice and self-study, and (b) Learning performance either formative or summative assessments, which indicate the level of a student’s understanding of a subject. Instructors analyze the cycle’s evidence and (1) decide on a proper intervention, and (2) apply the intervention in the subsequent cycle. (3) Changes in learning performance and behaviors from the previous cycle provide solid evidence that can be used to evaluate the effectiveness of the interventions revealing which approaches work and which should be avoided. This ongoing process could enhance the efficiency of course instruction and student success rate.

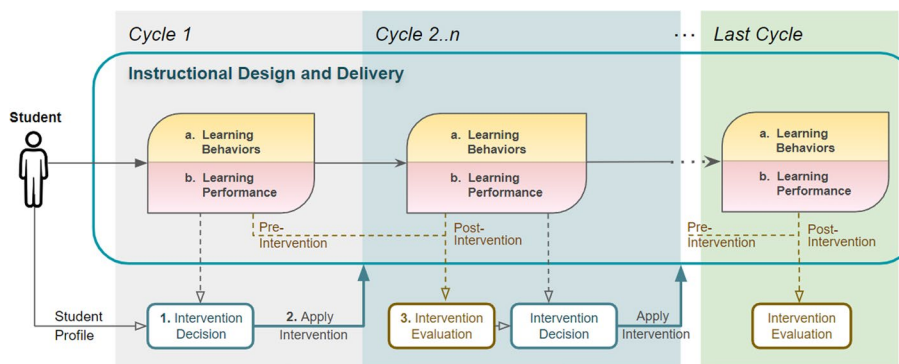


Fig. 2 An instructional intervention cycle in a course

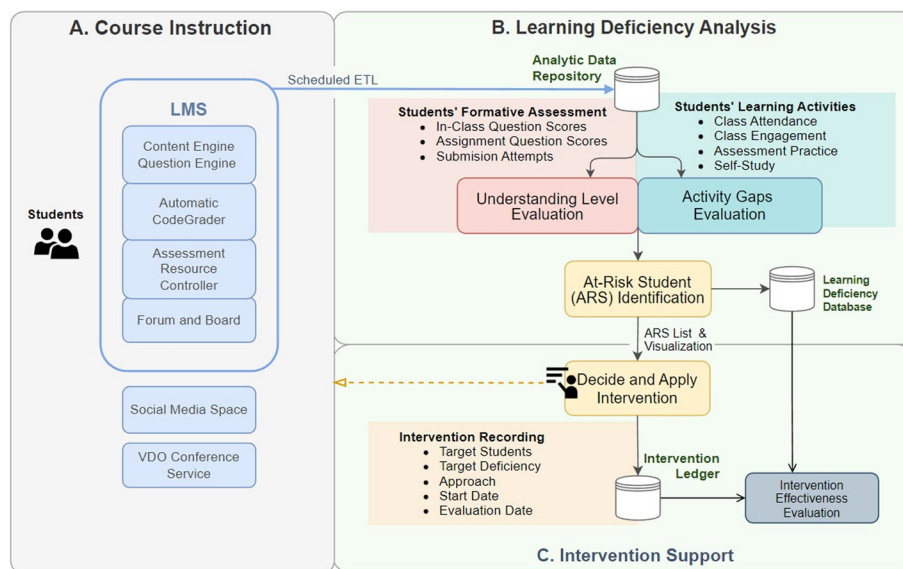


Fig. 3 i-Ntervene architecture

Concerning the challenges of LA Intervention discussed in Subsection [Related work.LA intervention methods and evaluation](#), data collection and analysis require excessive effort which is infeasible to run in iterations for most courses. This paper proposes an intervention platform of LMS, learning analytics and reporting, aiming to strengthen effective intervention cycles. Figure 3 depicts the platform architecture which comprises 3 main components: (A) Course Instruction which provides essential learning environment, (B) Learning Deficiency Analysis which identifies ARS in each intervention cycle by evaluating students' learning activities, and their subject understanding, and (C) Intervention Support which records intervention information and evaluates their effectiveness according to EBI practice. The following subsections elaborate these 3 components in detail.

Course instruction

This component provides the functionalities and environment necessary for programming instruction. A Learning Management System (LMS) offers essential e-Learning features and records learning activities, which serve as the main data source of i-Ntervene. The key to success in programming education is 'learn-by-doing' where students are required to practice on their own (Bannedsen & Caspersen, 2007; Neil et al., n.d.). Thus, most student learning activities are centered around coding questions. Automatic code grading integrated with LMS is compulsory as it validates and provides immediate feedback on submitted code and also records details of students' attempts for further analysis.

The Covid-19 pandemic has accelerated the need for a Virtual Learning Environment. i-Ntervene includes features of online assessment and video conferencing for remote face-to-face instruction. In recent years, many software applications have been developed to control resources on students' computers during quizzes and exams. This

software can often incorporate with LMS, e.g., Safe Browser Exam,² Proctortrack.³ There are also multiple choices of well-known video conference services that can work with LMS such as Zoom,⁴ Google Meet⁵ and Microsoft Team.⁶

Learning deficiency analysis

This research adopts the ABC indicators from Johns Hopkins University's study, described in Sect. [Learning analytics intervention](#), to identify At-Risk Students (ARS) in a course. The Attendance and Behavior indicators are represented by 4 student activities, i.e., class attendance, in-class engagement, assignment practice and self-study, while the course performance indicator is represented by students' question practice scores. During course instruction, students may expose learning deficiencies in the form of (1) Lacking in learning engagement which can be captured from their learning activities, e.g., decreasing class attendances or low participation in class activities, and (2) Lagging in subject understanding which can be measured from formative and summative assessments, e.g., assignments, lab practices and quizzes.

The i-Ntervene learning analytics component evaluates students' deficiency in both aspects to identify ARS whom instructors should focus on and intervene in each cycle. First, the scheduled ETL extracts students' learning activities and question practice data from the LMS and code grader log, cleans and transforms it to the specific format to load into the Analytics Data Repository. The dataset is then evaluated with expected values defined by the instructor on the corresponding cycles, resulting in identification of student learning deficiencies, i.e., Activity Gaps and Understanding Level, which are criteria used to justify whether a student is at-risk of falling behind. The component's final outputs are lists of ARS and visualizations to support intervention decisions in each cycle.

Activity gaps evaluation

Student misbehaviors frequently become more intense and more resistant without early intervention (Wills et al., 2019). Such misbehaviors are usually expressed in the form of low learning engagement. With the capability of LMS, i-Ntervene captures 4 generic learning activities to evaluate the learning engagement, including class attendance, in-class engagements, assignment practice, and self-study. Table 3 describes how students' interactions are observed and categorized into learning activities.

Since learning activities are usually planned by the instructor for each instructional period, the number of activities a student performed is compared with the instructor's expectation, resulting in Activity Gaps which unveil students that potentially require interventions. Table 4 depicts the calculations.

Class attendance, class engagement and assignment practice are core learning activities designed by instructors. Hence, these Activity Gaps can be used to evaluate student's learning behavior directly. On the other hand, self-studying is a freewill activity

² <https://safeexambrowser.org>.

³ <https://www.proctortrack.com>

⁴ <https://www.zoom.us>

⁵ <https://meet.google.com/>

⁶ <https://www.microsoft.com/en/microsoft-teams>

Table 3 Student interactions on LMS categorized into 4 learning activities by i-Ntervene

Learning activities	Interaction on LMS	Observing timeframe
Class Attendance	A student logs in or joins the VDO conference session	Classroom start time ± grace periods for early and late
Class Engagement	A student interacts with a learning material (e.g., pop-up question, coding question, and video) in which the instructor introduced in classroom	Within classroom period
Assignment Practice	A student submits code on questions in which the instructor assigned as homework	Outside classroom period
Self-Study	A student watches video or submits code on questions in which the instructor did not assign	Outside classroom period

Table 4 Student’s Activity Gaps calculation for each intervention cycle

Activity gaps	Definition and calculation
Class Attendance Gap (Atdn Gap) $\in[-1,0]$	Ratio of Student’s class attendance in a cycle ranging from -1 for no attendance to 0 for fully attended $AtdnGap = \left(\frac{\text{Numberofclassstudentattended}}{\text{Numberofclassattendancethatinstructorexpected}} \right) - 1$
Class Engagement Gap (Engm Gap) $\in[-1,0]$	Ratio of student’s engagement in classroom in a cycle. The ratio ranges from -1 for no interaction on any material to 0 if a student interacted on all material of his attended classes. In case that a student did not attend any class, Engm gap will be defined as ‘not available’ (na) $EngmGap = \left(\frac{\text{Numberofmaterialastudentworkedoninclass}}{\text{Totalmaterialsintroducedintheclassthastudentattended}} \right) - 1$
Assignment Practice Gap (Asgm Gap) $\in[-1,0]$	Ratio of assignments in which students practiced in a cycle ranging from -1 for no practice to 0 for practiced on all assigned material $AsgmGap = \left(\frac{\text{Numberofassignmentstudentpracticed}}{\text{Numberofassignmentthatinstructorexpected}} \right) - 1$
Self-Study Rate $\in[0,\infty]$	Amount of non-assignment material in which students interacted in a cycle ranging from 0 for none to any positive value

reflecting student’s intrinsic motivation and might not be used to indicate who requires interventions. Self-study rate is good evidence of student learning motivation and a potential predictor of learning performance. It can be particularly useful in certain analytics, e.g., analyzing exam preparation patterns, tracing learning motivation trend or evaluating some aspects of instructional delivery.

Subject understanding level (UL) evaluation

Summative assessments such as exams and quizzes are well-established tools for evaluating students’ learning. The key disadvantage is that they are usually presented at the end of a topic or a semester when very limited corrective action can be taken. Formative assessments are recognized as a better tool because they consistently provide ongoing feedback that help instructors recognize where students are struggling and can provide appropriate intervention in time.

In STEM education, problem practice is a primary key to promote student learning (Utamachant et al., 2020). It is mandatory for students to consistently work on coding questions to assimilate programming concepts and achieve competencies. Hence, outcomes of student’s coding practice are reliable formative assessments that can indicate their understanding in each period. Programming code that a student submits can be assessed in 5 key aspects (Combéfis, 2022; Restrepo-Calle et al., 2018) as depicted in Table 5.

Table 5 Key assessment aspects on student code

Assessment aspects	Criteria
Syntactic correctness	All commands in code must be correct. Automatic code graders can simply verify this using a compiler
Semantic/logical correctness	Code's outputs must agree with the expected solution described by specifications. Most automatic code graders can validate the code outputs using defined test cases
Efficiency/performance	Code must provide outputs within a specific duration and computer resource. Most automatic code graders can specify CPU time and memory limits to validate this aspect
Quality/maintainability	This aspect involves coding style, guidelines and best practices which focus on readability and reusability. It is a consideration for advanced programming course. This aspect cannot be judged by automatic code graders
Plagiarism	Code should not replicate nor closely resemble another student's code. Some automatic code graders have a feature to check for code authenticity, but it could yield high false positives in cases of simple coding questions

Most fundamental programming courses have a common objective for students to become familiar with coding syntax and be able to adapt and apply programming logic to solve problems (Comb  fis, 2022). Thus, the evaluation of subject understanding should prioritize syntactic and logical correctness. The automatic code grader in i-Ntervene first compiles a student's code to validate the syntax. The code that fails this process is counted as a 'Syntax Error' attempt and feedback is given to the student. The code that is successfully compiled is run against defined test cases for logical evaluation and is scored in the proportion of correct test cases.

i-Ntervene considers a student's attempts on one question as a *Question Practice*. On every intervention cycle, all question practices are evaluated to identify those that the student has trouble understanding. The evaluation procedure consists of 2 steps: i) Classifies all question practices into 3 question types, i.e., Resolved, Struggled and Gave-up, and ii) Calculates ratio of each question type to justify individual Understanding Level of that cycle.

I. Student's Question Practice Classification. Figure 4 depicts the decision tree for classifying student's question practices. Input of the tree is a student's question practice outcome, i.e., highest score, total attempts, time duration and error type. A student's question practice that has highest score above the pass score (a) is considered a *Resolved* question, which demonstrates student's satisfactory understanding. Otherwise, the question practice indicates a problem which can either be a result of the student's poor understanding or they did not exhibit the effort necessary to solve the questions. The tree determines student's effort spending on the question (b), i.e., answering attempts and duration. If the student did not spend effort up to the instructor's expectation, then the student's question practice is considered a *Gave-up* question. On the other hand, if the student had spent a proper amount of effort, then his question practice is determined as a *Struggled* question. The tree further classifies the struggled question based on error types. If the student's answering attempts contain syntax errors over Syntax Error Limit, the student's question practice is identified as *Syntax Struggled*, otherwise it is defined as *Logic Struggled*.

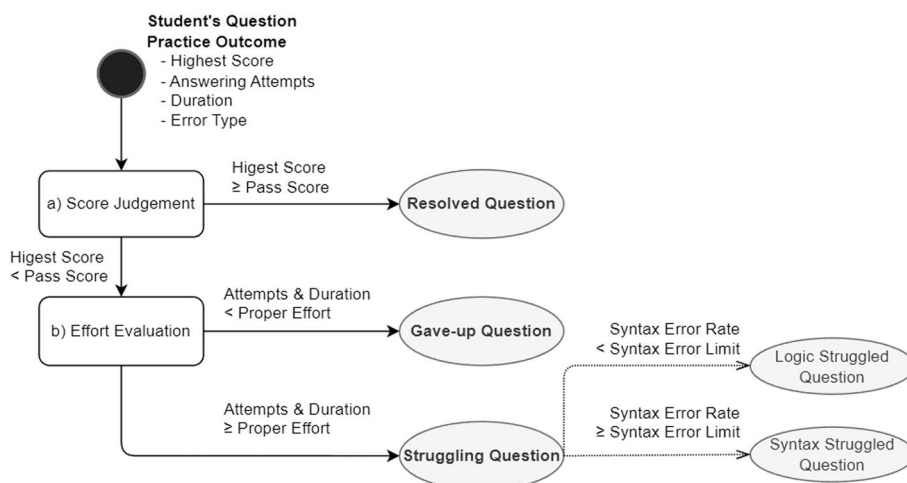


Fig. 4 The decision tree to justify a student's question practice

Parameters in Fig. 4 are described as follow:

- Pass Score: A passing score to consider that a student can successfully resolve the question.
- Proper Effort: Expected number of attempts and duration that a student should have spent to solve the question.
- Syntax Error Rate: A ratio of syntax error attempts to total attempts.
- Syntax Error Limit: An upper limit of syntax error rate to consider syntax struggling.

The current study treats runtime errors as a part of Logic Struggling as they are usually caused by mistakes in code logic, resulting in access violation of computer resources, e.g., memory reference, file system, process timeout. The rule reaches the final stage at detecting Syntax struggling from Logic struggling. It would be beneficial to break down Logic Struggling further into more fine-grained subtypes for precise intervention, but it is very challenging as there could be countless possible kinds of mistakes that could cause a student's code to not achieve all test cases successfully. This could be another topic of research in programming education.

II. Understanding Level Calculation. Student's Understanding Level in an intervention cycle is calculated from the proportions of Resolved, Struggled and Gave-up question practices. A formula that could be general for common courses is proposed below. Instructors could adjust it to fit their course design:

$$Student's\ Understanding\ Level = \left(\frac{Resolved\ Questions - (Struggled\ Questions + Gave\ Up\ Questions)}{Total\ Practiced\ Questions} \right)$$

The result of the formula ranges from -1 to 1 representing a student who is seriously struggling and a student who completely understands. Note that Gave-up questions are added up to Struggled questions in the calculation because, in our experience, giving up is usually the consequence of struggling. In other words, students who struggle to resolve one question are likely to spend less effort and give up on succeeding questions of the same topic. However, Gave-up question practices could indicate students

with low effort regulation motivation and this would be useful for analysis in certain circumstances.

A student's Understanding Level can be broken down to Syntax Understanding Level and Logic Understanding Level using ratio in the formula:

$$\text{SyntaxUnderstandingLevel} = 1 - \left(\frac{\text{SyntaxStruggledQuestions}}{\text{TotalStruggledQuestions}} \right)$$

$$\text{LogicUnderstandingLevel} = 1 - \text{SyntaxUnderstandingLevel}$$

The Syntax and Logic Understanding Level, ranging from 0 to 1, indicates the potential root cause of the student's problem which enables the instructor to apply more precise interventions. Students with low Syntax Understanding Level can be helped simply by advising the correct syntax. On the other hand, low Logic Understanding Level students may need more sophisticated approaches such as tutorial sessions, workshop coaching or peer learning.

At-risk student (ARS) identification

In order to enhance student learning, the priority is to early detect At-Risk Students (ARSs) who exhibit signs of falling behind the instructor's expectations in term of both learning behaviors and content understanding. These students are considered at risk of failing the course or dropping out. To be more specific, this study identifies ARS at the end of cycles based on 2 criteria: (i) high Activity Gap which indicates misbehaving students who do not engage in learning activities up to the instructor's expected level, and (ii) low Understanding Level which indicates struggling students who could not resolve assigned questions to the instructor's expectations. Figure 5 depicts the identification rule.

The learning behavior intervention should place priority on students who have persistent behavior deficiencies rather than a single recent occurrence of misbehavior. This reduces false positives caused by personal incidents which may happen to anyone in short durations. For example, a student who is sick and misses activities for 1 week is likely to get back on track after recovery. Giving that student behavior intervention would be unproductive. To reduce such circumstances, the research employs a trending concept by averaging Activity Gaps of 2 or more consecutive periods, known as *Moving Average (MA)*. Instructors should carefully choose an MA period that fits the nature of their course. Too short of a period may induce more false positive cases, while too long of a period could delay the intervention and cause instructors to miss the best opportunity for taking effective action.

The effective intervention should be implemented based on a student's particular situation and problem (Gašević et al., 2016; Zhang et al., 2020). Figure 5 illustrates a common rule to classify intervention types for ARS:

- (i) Students who have no activity in the system for a specific period, are considered dropped-out and excluded at the first step to avoid wasting unnecessary intervention effort.
- (ii) Students' activity gaps are evaluated according to Sect. [Learning analytics intervention](#). Students whose gaps' MA surpasses the acceptable threshold defined by the

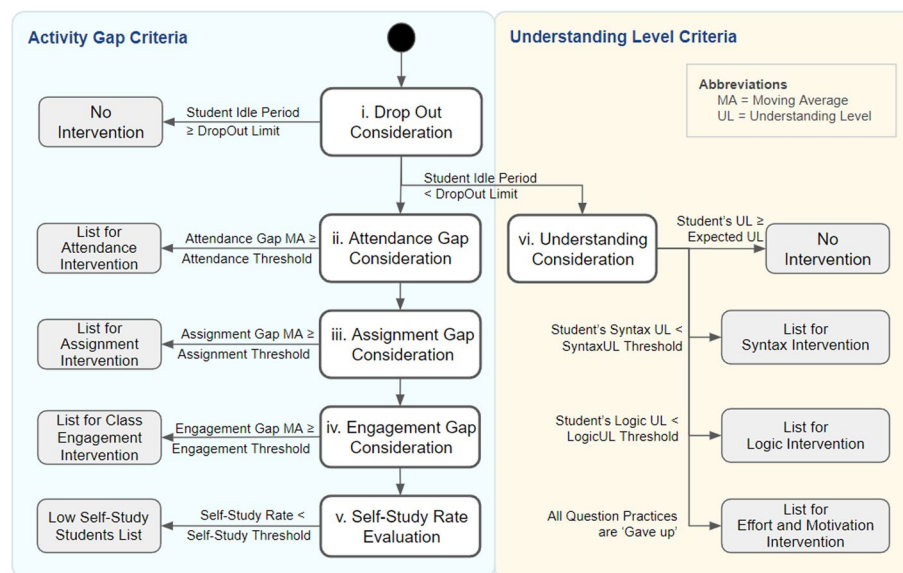


Fig. 5 ARS identification rule

instructor in each period are listed for intervention. For example, if the instructor expects students to attend the class not less than 50% then Attendance Threshold should be set at 0.5. At the end of an intervention cycle, students with Attendance Gap MA below 0.5 will be listed for Attendance Intervention.

- (iii) Self-Study rate is based on the number of non-assignment materials that a student hands in. In cases that Instructors define a threshold, i-Intervene can list out low self-study rate students for ad hoc review and take supplemental actions.
- (iv) Instructors should treat students having understanding problems as soon as indicated since they are likely to get more confused over time and this usually results in drop out or poor learning performance. At the end of each intervention cycle, i-Intervene calculates Understanding Level of students as described in Subsection [LA intervention methods and evaluation](#). Students whose Understanding Level surpasses the instructor's Expected UL demonstrate good understanding of the subject matter in that period and will not be listed for intervention. In the meanwhile, students with Understanding Level below the instructor's expectation will be checked on Syntax UL and Logic UL and listed for an intervention accordingly. Lastly, students with all of their question practices categorized as Gave-up will be listed for Effort and Motivation intervention.

Intervention support

Intervention decision

Designing intervention is a complex process that involves many variables (Klang et al., 2022). It is the instructor's task to determine a proper intervention method for At Risk Students (ARS) which relies on instructor's expert judgement rather than algorithm tuning (Arnold & Pistilli, 2012).

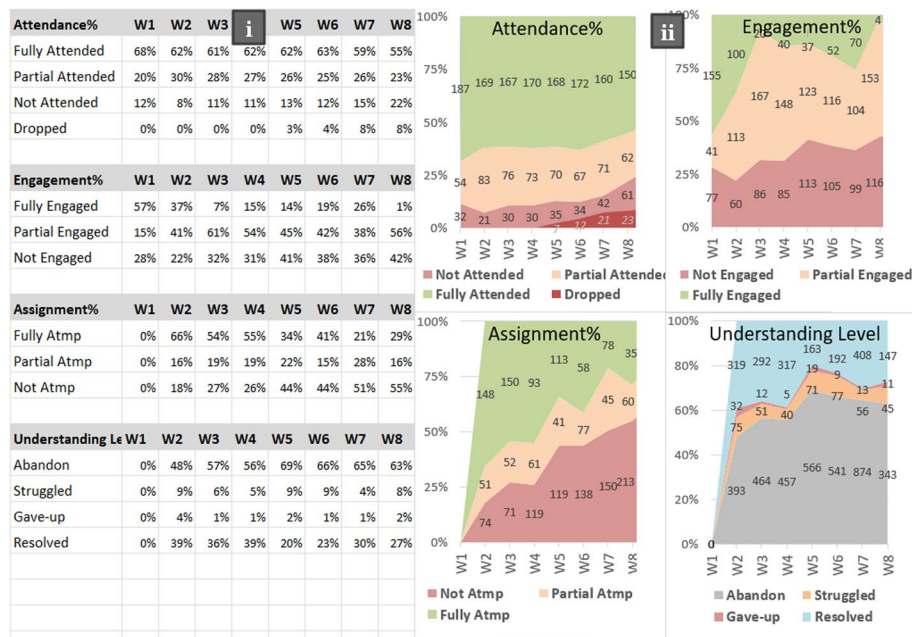


Fig. 6 Activity participation in class level

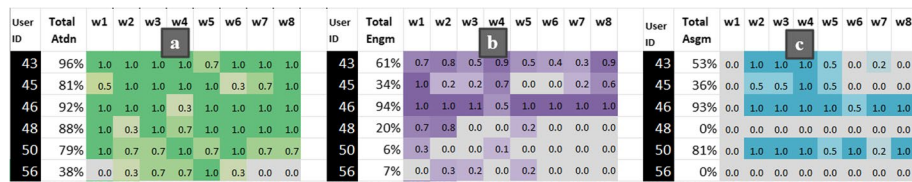


Fig. 7 Activity gaps in individual level. a Attendance gap, b Engagement gap, c Assignment gap

To support intervention decisions, i-Ntervene displays the trend of student’s activities participation and Understanding Level along with ARS lists. Figure 6 illustrates students’ participation and understanding in class level. The spreadsheet (i), exported from i-Ntervene, shows the percentage of students in each cycle up until the present. Stack charts on the right (ii) visualize the spreadsheet data. Values in the stack areas represent the number of students. The information provides an overview of the instructional situation and trends endorse instructors to perform preemptive intervention to the whole class. The spreadsheet in Fig. 7 illustrates individual activity gaps which provide detail on patterns of misbehavior of each student and enhances instructor’s intervention decision at an individual and group level.

After determining an intervention approach, the instructor needs to record the information in i-Ntervene including, target students, target deficiency, intervention approach, start and end date. Figure 8 illustrates an instructor’s screen for recording and evaluating Understanding Interventions.

Effectiveness evaluation

The gold standard for evaluating educational interventions is Randomized Control Trials (RCT) (Outhwaite et al., 2019). However, such an approach is rarely feasible in

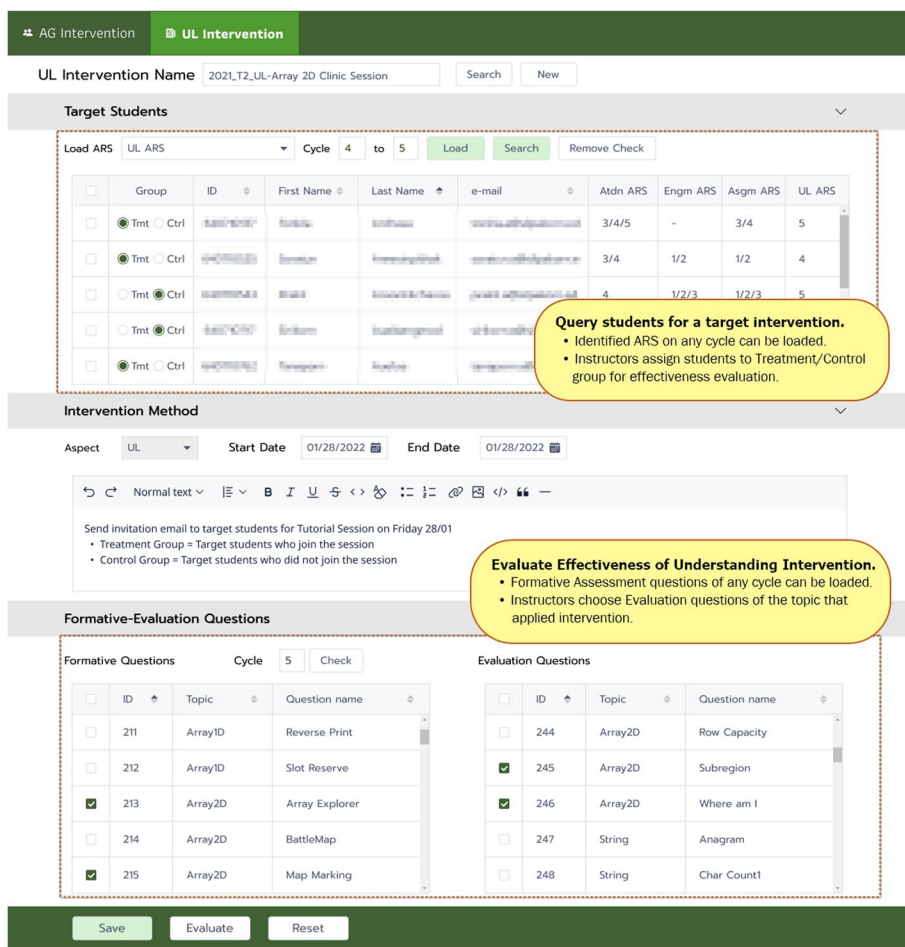


Fig. 8 Recording and evaluation screen of understanding intervention.

a real-world course. RCT has the following limitations (Rienties et al., 2017; Sullivan, 2011): (1) uncontrollable and unpredictable variables, (2) insufficient or unqualified population to establish valid control and treatment groups, (3) ethical issue due to unequal learning opportunities, and (4) tremendous experience resources required to support the whole process.

To be applicable for most courses, i-Intervene adopts a quasi-experimental design principle by systematically evaluating intervention effectiveness based on the improvement in Activity Gaps or Understanding Level of pre and post intervention. In other words, the effectiveness is calculated from the ratio of students who show improvement after the intervention as compared to the total of intervened students.

For an intervention targeting misbehaviors, i.e., class attendance, class engagement, assignment and self-study, i-Intervene compares Activity Gap of intervened students during the evaluation period with their Activity Gap before intervention. Ratio of students who have Activity Gap improvement over the expected effect size to the total intervened students illustrate the effectiveness.

$$MisbehaviorInterventionEffectiveness =$$

$$\left(\frac{\text{Number of Students with Activity Gap Improved exceed Effect size during Evaluation Period}}{\text{Total Number of Intervened Students}} \right)$$

Note : Effect size is specified by the instructor.

Interventions on learning behavior require time in order to have an effect that can be captured from learning activity (Li et al., 2018; Wong & Li, 2020). A key to success is to diligently determine a proper evaluation period which can be different from case to case.

For an intervention on subject understanding deficiency, its effectiveness is calculated by comparing the average question scores of pre- and post-intervention of the same topic between At Risk Students (ARS) who has received the intervention (treatment group) and ARS who has not (control group). The instructors can easily assign students to each group based on the ARS list suggested by i-NTervene. Also, the instructors can flexibly choose post-intervention questions for evaluation which can compose multiple assignments, quizzes or exams. Figure 8 illustrates the user interface of the prototype.

$$UL\text{InterventionEffectiveness} = \left(\frac{\text{AvgScoreofIntervenedARS} - \text{AvgScoreofUnintervinedARS}}{\text{AvgFullScore}} \right) \dagger$$

†: Avg Score : the average score from Post-Intervention questions

The effectiveness result can be positive and negative values which indicate the favorable and unfavorable effect of the UL intervention. The associated p-values affirm the certainty of the average score difference.

Since each intervention is different in context, e.g., approaches, students, situations, there is no target effectiveness value that can indicate success or failure in general. The instructor needs to determine whether the value is satisfactory or compare with values from other approaches to justify a better one. After running intervention effectiveness evaluations for a period of time, instructors should be able to compile a set of interventions that works for their courses. This intellectual knowledge is beneficial to other courses in the same domain or in a similar environment.

Experimentation

Course setting and challenges

A prototype of i-NTervene was developed and tested on a Java programming course with the main objective to build basic Object-Oriented (OO) programming competency for first-year undergraduate students in computer science curriculum. The course was delivered to 253 enrolled students for the duration of 16 weeks with 8 learning topics. The learning effort for each week was 2 lecture-hours and 4 lab-hours. There were 6 summative assessments covering every topic. The instructor assigned weekly programming questions as homework. The results of these assignments indicate their understanding regularly which is an ideal formative assessment. The prototype was built on Moodle⁷ with H5P plugin⁸ for video interactive content and CodeRunner plugin⁹ for auto-grading programming questions. The course was conducted online due to the pandemic, using

⁷ <https://moodle.org>

⁸ <https://h5p.org>

⁹ <https://coderunner.org.nz>

Zoom⁴ for virtual classroom sessions and Safe Browser Exam² to control the remote exam environment.

This Java course has been suffering from high drop out and fail rates with the rates exceeding 50% of total enrollments in recent years. The instructors have observed and concluded 2 major root causes:

Low Self-Regulated Learning (SRL): Most students came directly from high school. Many of them struggled to adapt to university learning, which requires high-level of SRL in order to succeed. They did not spend enough effort in learning, which showed in class absence, not paying attention in class, and disregarding assignments. This situation correlates with an SRL research on college students (Kitsantas et al., 2008) and is a mutual concern with many higher education institutes (Hellings & Haelarmans, 2022).

Insufficient background knowledge: From a pre-course survey, 41% of respondents had no computer programming background, which is a big challenge to developing their OO programming competency within only one semester.

Recognizing the challenges, the instructors had a strong intention to improve student learning with effective interventions, which would not have been possible if implemented manually without a proper tool due to the large ratio of students to instructors (253:2). We believed i-Intervene could support the instructors in making decisions and perform iterative interventions in a timely manner and also help evaluate the intervention effectiveness so that the successful interventions can be reused.

i-Intervene implementation

Setup parameters described in the previous section were configured by the instructors with the intervention cycle set for every week. Table 6 depicts parameters to justify students' question practices (cf. Subsection [Learning analytics intervention](#)). The instructors selected assignment questions with similar difficulty on each topic in this experiment semester. Therefore, all questions in a topic share the same parameter values. In case the instructors desire to enrich students' experience by providing questions with variety of difficulty, the parameters will need to be tuned individually.

For ARS identification (cf. Sect. [LA intervention methods and evaluation](#)), the instructors specified the values of Activity Gaps and Understanding Level for every cycle as shown in Table 7. These parameters can be adjusted to accommodate courses with different settings.

Unfortunately, in this experiment, students chose to develop their code on local IDEs, such as, NetBeans¹⁰ or Eclipse¹¹ as it provides rich features for debugging. Complete codes were copied from their local IDEs to CodeRunner IDE in Moodle just for grading against predefined testcases to gain scores. Therefore, most of the submitted code was correct in syntax. Students who struggled in syntax basically discarded the question and did not submit any code. The prototype was unable to identify Syntax Struggled question practice in this situation, so we needed to treat struggled question practices as a single type, i.e., Logic Struggled.

¹⁰ <https://netbeans.apache.org>

¹¹ <https://www.eclipse.org/ide/>

Table 6 Parameters to justify Student’s Question Practice

Learning topics	Pass score	Proper effort		Syntax error limit
		Attempts	Duration (min)	
Array1D, string	0.7	5	5	0.8
Array2D	0.7	10	5	0.8
Method	1	5	5	0.8
Class programming, class inheritance	0.7	5	5	0.5
Recursion	1	10	5	0.8

Table 7 Activity gaps and understanding level criteria for at risk student identification

Activity gap	Value	Explanation
Drop-out threshold	4 Cycles	Students with no activity on LMS for 4 consecutive weeks were assumed to have already dropped out and excluded from the intervention process
Moving average	2 Cycles	Students who had average Activity Gaps of 2 recent weeks less than the thresholds are listed for ARS in that aspect
Attendance threshold	– 0.2	
Assignment threshold	– 0.2	
Engagement threshold	– 0.4	
Self-study rate	na	No intervention on self-study
Expected effect size	0.01	Students who gain more than 1% improvement of Activity Gap Moving Average during the evaluation period are counted toward the intervention effectiveness
Expected UL	0.5	Students who archived pass score on more than half of their hands-on questions in that cycle are not list for intervention
Syntax UL threshold	0.66	For students who do not meet Expected Understanding Level in that cycle, if 2/3 of their question practices were classified as Syntax Struggled, they are listed for Syntax intervention. Otherwise, they are listed for Logic intervention
Logic UL threshold	0.33	

Instructional intervention execution and evaluation results

i-Ntervene provided ARS lists for intervention in Week#2 but the instructors decided to wait until the grace enrollment period ended and began to consider ARS lists in Week#4. With 253 enrolled students and only 2 instructors, they chose to perform interventions at the level of class and specific student groups using email, verbal, and social media space communication. This approach was adopted in many studies with a large number of students (Choi et al., 2018; Dodge et al., 2015; Kimberly & Pistilli, 2012; Lu et al., 2017; Milliron et al., 2014).

For learning behavior interventions, the instructor delivered messages regarding their lacking aspects via ARS emails or announce the messages in the course’s social media space and in the classroom as detailed in Tables 8 and 9. The effectiveness of learning behavior intervention was evaluated from ratio of ARSs who had activity gap improvement to total intervened ARSs in the successive week. For subject understanding, the instructors set up tutorial sessions after complete teaching of each topic and invited ARSs to attend at their own will. The effectiveness of understanding interventions was assessed from average question scores between ARSs who attended tutorial

Table 8 Result of class attendance interventions

Applied Week#/ evaluation Week#	Class attendance intervention methods	% Effectiveness (improved gap ARS/total intervened ARS)
5/6	Method 1: Sending email to ARSs to address their absence and encourage to attend the class	45% (19/42)
7/9	Method 2: Sending email to ARSs to address their absence with the previous semester stats evidencing the importance of class attendance	71% (20/28)
11/12	Method 3: Reminding minimum attendance rate to pass the course via social media and classroom announcement	83% (15/18)

sessions (treatment group) and ARSs who did not (control group). The evaluation questions selected from quizzes of the same topic after tutorial sessions. Mann–Whitney U tests were applied to test for statistical significance due to non-normal data distribution.

With support from i-Ntervene, the instructors were able to implement 12 complete intervention cycles (7 learning behavior and 5 subject understanding) during 16 weeks of the semester. Tables 8, 9 and 10 depict the implemented intervention details and evaluation results.

Findings from intervention results

The results demonstrate the capability of i-Ntervene to support intervention cycles in an instructor-led programming course. Regarding the effectiveness of behavior intervention, the instructor gained insight into what methods work better than others which can be compared quantitatively. Tables 7 and 8 show that the method of sending emails together with the last semester stats (Methods 2 and 5) was more effective than sending only an encouragement email, since the behavior improvement rate was higher on both class attendance (45% to 71%) and assignment practice (23% to 46%). However, the most effective intervention was the critical condition announced by the instructors in the later weeks that students will fail the course if the minimum attendance rate was not met (Methods 3 and 7). The effective ratio of class attendance and assignment practice

Table 9 Result of assignment practice interventions

Applied Week#/ evaluation Week#	Assignment practice intervention methods	% Effectiveness (improved gap ARS/total intervened ARS)
6/7	Method 4: Sending email to ARSs to address their assignment missing and encourage to practice more	23% (14/62)
7/9	Method 5: Sending email to ARSs to address their assignment missing with the previous semester stats emphasize the importance of assignment practice	46% (13/28)
10/11	Method 6: Sending email to individual ARSs with customized message addressing their assignment and in-class engagement patterns	36% (55/152)
12/13	Method 7: Reminding minimum assignment practice rate to pass the course via Social Media Space and classroom announcement	55% (36/65)

Table 10 Result of Understanding Intervention using Tutorial Sessions

Learning topics Tutorial date, Exam date (duration)	Attended: unattended	Statistical test
Array1D 7/1/2022, 29/1/2022 (22 days)	Number of students = 14: 106 Average Scores = 2.93: 2.80	p = 0.45
Array2D 28/1/2022, 29/1/2022 (1 days)	Number of students = 37: 28 Average Scores = 1.51: 1.43	p = 0.22
String 4/2/2022, 12/2/2022 (8 days)	Number of students = 14: 74 Average Scores = 7.86: 6.18	p = 0.47
Class basic 11/2/2022, 12/2/2022 (1 days)	Number of Students = 22: 37 Average Scores = 8.24: 3.81	p = 0.015*
Array of objects 25/02/2022, 23/3/2022 (25 days)	Number of Students = 9: 62 Average Scores = 6.66: 2.74	p = 0.020*

* Statistically Significant (Mann–Whitney U Test)

were up to 83% and 55% respectively. We note that the channel of communication could yield a certain impact on intervention effectiveness evaluation. Some students informed us that they did not check their university emails on regular basis, which means some of them could miss the intervention. The instructors decided to change communication from emails to the course's social media and chat group in later interventions.

Regarding the effectiveness of intervention on subject understanding, i-Ntervene revealed that providing tutorial sessions as the intervention for struggling students was only effective for certain topics. Table 9 shows that students who attended the early topics tutorial sessions, i.e., Array1D, Array2D, and String gained only a slight margin and insignificantly better in average score compared to those who did not. On the other hand, for later topics, i.e., Class Basic and Array of Objects, the tutorial session yielded substantial improvement on average scores with statistical significance ($p = 0.015$ and 0.020 , respectively). The instructors analyzed and concluded that a single tutorial session may not be sufficient to improve student understanding on complex algorithmic topics such as Array1D, Array2D and String since they require mastery of programming logics to handle variety of conditions and complex loops. On the contrary, the tutorial session method is effective for conceptual and structural topics such as Class Basic, Class Inheritance and Array of Objects. Based on these important findings drawn from i-Ntervene, the instructors decided to revise the instructional design of complex algorithmic topics by assigning more consistent question practices and providing multiple labs/workshops with teaching assistance as interventions.

Discussions

Although i-Ntervene shares similar objectives with several existing studies by focusing on instructor intervention support (Cobos & Ruiz-Garcia, 2020; Majumda et al., 2019; Herodutou et al., 2019; Azcona et al., 2019; Froissard et al., 2015) in identifying At-Risk Students (ARS) and displaying proper visualization to support decision making, i-Ntervene specifically aims to tackle the main challenges faced by a course with small number of instructors but having a large class size and a lengthy instruction period. While most existing research merely considered a *single* intervention, i-Ntervene stands out by supporting instructors to perform *multiple* rounds of interventions

throughout the course period with its comprehensive process. It traces temporal students' engagement and understanding to systematically identify ARSs, records the applied intervention methods used, and then evaluates the effectiveness of the implemented interventions based on the improvement. This feature consistently informs instructors on *who* should be intervened, *which aspects* require attention, and *how effective* the implemented intervention is throughout the iterations. Thus, instructors are equipped with supporting tools and well-informed data and analysis results to precisely choose the right intervention techniques for the right students. Moreover, continuous monitoring, fine-tuning and adjustment of the applied interventions can be done along the course period. The experimental results in Sect. [LA intervention methods and evaluation](#) clearly illustrate i-Ntervene's capabilities as claimed. For the experimented course, there were only 2 instructors instructing a total of 253 students for the period of 16 weeks. With the support of i-Ntervene, a total of 12 complete intervention cycles were successfully implemented, evaluated, and adjusted.

Concerning instructional intervention execution, most interventions usually aim at improving students' comprehension and learning behavior. For the comprehension aspect, i-Ntervene tracks students' activities and evaluates their subject understanding from their assignment scores. Common intervention methods are providing supplemental materials or arranging additional tutoring sessions on the problematic topics. The effectiveness of the interventions can be evaluated directly by the improvement of assessment scores. However, when it comes to the learning behavior interventions, instructors are often faced with the challenge of managing students' learning motivations and preferred learning styles, which are psychological in nature and cannot be evaluated tangibly. The current version of i-Ntervene does not yet integrate a tool to assess them. Therefore, instructors need to rely on their intuition and prior experience to select appropriate intervention methods.

In this experimentation, the instructors decided to use email as the method for behavior interventions due to a large number of students. We could observe some traces of learning motivation that were affected positively by the implemented intervention methods. The first observation is the email intervention sent in Week 7 (Methods 2 and 5). The message in the email stressed the importance of class attendance and assignment practice for learning success which could activate *task value motivation*. i-Ntervene assessed the effectiveness of this method and discovered that it yielded significantly better results compared to the simple encouragement message sent during Week 5 (Methods 1 and 4). This could be a result of increasing task value motivation on the intervened students. The second observation is the email intervention in Week 10 (Method 6). The email message contains individual engagement statistics aiming to stimulate students' self-awareness. i-Ntervene reveals higher effectiveness of this method over the baseline (Methods 1 and 4) suggesting an improvement of *metacognitive and self-regulation* among the intervened students. Lastly, the email intervention sent in Week 11/12 (Methods 3 and 7) informed the critical conditions that students must meet to pass the course. The message aimed to trigger *extrinsic goal motivation* so the improvement gained over the baseline could be a part of increasing in extrinsic goal motivation. Although these interpretations are not definitive, they do provide some indication of potential impact of learning

motivation on the effectiveness of selected intervention approaches. We believe that incorporating learning motivation data with the existing learning engagement data can strengthen the temporal dataset which can enhance precision of at-risk student identification significantly.

In addition to the temporal data, some LA Intervention research have indicated the usefulness of static data to enhance the analytics such as students' demographic, background knowledge (Herodutou et al., 2019; Azcona et al., 2019) and learning style (Shaidullina et al., 2023). Integrating such data into the future version of i-Ntervene can enhance the platform analytics feature and provide valuable insights to instructors, hence supporting them to select an optimal intervention method that can address issues at the core and can substantially improve the intervention efficiency.

Conclusion

The paper proposes i-Ntervene, an LA Intervention platform that supports iterative learning intervention for programming courses. Student learning interactions and outcomes are systematically collected and analyzed for two types of learning deficiencies. Firstly, learning behaviors deficiencies are evaluated from activity gaps, i.e., class attendance, in-class engagement, assignment practice, and self-study. Secondly, student's subject understanding level is evaluated from their question practices. Based on each student's specific deficiency area, i-Ntervene identifies at-risk students and provides proper visualizations to support instructors for intervention decision. Furthermore, i-Ntervene analyzes the effectiveness of the implemented interventions at the defined period by comparing the improvement rate of pre- vs. post-intervention on the specific area complying with Evidence-Based Intervention (EBI) specification.

i-Ntervene prototype was experimented on a Java programming course with 253 first-year university students. Although there were only 2 instructors, with the support of the platform, they could successfully perform 7 learning behavior interventions and 5 subject understanding interventions on at-risk students. The effectiveness evaluation quantitatively revealed the performance of every intervention, which enabled the instructors to determine what approach worked in these particular course settings. This allows them to keep optimizing the course's intervention in the long term.

Currently, i-Ntervene limited to learning activities stored in LMS. In the future, student engagement data from classrooms could be included to enrich at-risk-student identification. Many educational science research aim to study classroom engagement by investigating interactions that actually happened in class. A well-known method is Flender's Interaction Analysis, which manually collects classroom communication every few seconds and performs analysis to improve instructional delivery in class (Amatari, 2015; Sharma & Tiwari, 2021). This feature can be adopted to enhance classroom participation analysis allowing the platform to evaluate comprehensive student engagements in both online and offline aspects. In addition to including offline learning feature, an important area for future work in this study is to incorporate information that has been proven in educational research community to impact student success. The information includes temporal data, such as learning motivation and emotion, as well as static data, such as student demographics, background knowledge and learning style. This should

enhance at-risk-student identification and provide more informative support for selecting appropriate intervention approaches.

Abbreviations

ARS	At-Risk Students
Asgm	Assignments
Atdn	Class Attendance
CS1	Introduction to programming courses
CS2	Basic data structure courses
EBI	Evidence-Based Intervention
Engm	Class Engagement
IDE	Integrated Development Environment
LA	Learning Analytics
LMS	Learning Management System
MA	Moving Average
OO	Object-Oriented
RCT	Randomized Controlled Trials
UL	Understanding Level
VLE	Virtual Learning Environment

Acknowledgements

We would like to thank Asst.Prof.Dr. Pinyo Taeprasartsit, Asst.Prof.Dr. Ratchadaporn Kanawong and Asst. Prof. Dr. Tasanawan Soonklang, Department of Computer Science, Silpakorn University, Thailand for providing opportunity and supports on the platform experiment.

Author contributions

PU contributed to the research conceptualization, design, data collection, analysis and interpretation of results, and writing manuscript. CA and SP provided research consulting. All authors read and approved the final manuscript.

Funding

This work was supported by the grants from Asian Institute of Technology (AIT) and National Science and Technology Development Agency (NSTDA) according to the Thailand Graduate Institute of Science and Technology (TGIST) scholarship agreement SCA-CO-2562-9681-TH.

Availability of data and materials

The data sets generated and analyzed during the current study are not publicly available due to anonymity issues but are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 17 April 2023 Accepted: 18 July 2023

Published online: 25 July 2023

References

- Aina, C., Baici, E., Casalone, G., & Pastore, F. (2022). The determinants of university dropout: A review of the socio-economic literature. *Socio-Economic Planning Sciences*, 79, 101102. <https://doi.org/10.1016/j.seps.2021.101102>
- Aivaloglou, E., & Hermans, F. (2019). Early programming education and career orientation: The effects of gender, self-efficacy, motivation and stereotypes. In *Proceedings of the 2019 ACM Conference on International Computing Education Research (ICER'19)* (pp. 679–685). <https://doi.org/10.1145/3287324.3287358>
- Amatari, V. (2015). The instructional process: A review of flanders' interaction analysis in a classroom setting. *International Journal of Secondary Education*, 3, 43. <https://doi.org/10.11648/jijsedu.20150305.11>
- Araka, E., Maina, E., Gitonga, R., & Oboko, R. (2020). Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—systematic review (2008–2018). *Research and Practice in Technology Enhanced Learning*. <https://doi.org/10.1186/s41039-020-00129-5>
- Arnold, K., & Pistilli, M. (2012). Course signals at Purdue: Using learning analytics to increase student success. *ACM International Conference Proceeding Series*, 10(1145/2330601), 2330666.
- Azcona, D., Hsiao, I. H., & Smeaton, A. F. (2019). Detecting students-at-risk in computer programming classes with learning analytics from students' digital footprints. *User Modeling and User-Adapted Interaction*, 29, 759–788. <https://doi.org/10.1007/s11257-019-09234-7>
- Balfanz, R., Hall, D., Verstraete, P., Walker, F., Hancock, M., Liljengren, J., Waltmeyer, M., Muskauskis, L., & Madden, T. (2019). *Indicators & Interventions*. School of Education for the Everyone Graduates Center, Johns Hopkins University.
- Barbera, S. A., Berkshire, S. D., Boronat, C. B., & Kennedy, M. H. (2020). Review of undergraduate student retention and graduation since 2010: Patterns, predictions, and recommendations for 2020. *Journal of College Student Retention: Research, Theory & Practice*, 22(2), 227–250. <https://doi.org/10.1177/1521025117738233>

- Barr, M., & Kallia, M. (2022). Why students drop computing science: using models of motivation to understand student attrition and retention. In *Proceedings of the 22nd Koli Calling International Conference on Computing Education Research (Koli Calling'22)* (pp. 1–6). Association for Computing Machinery. <https://doi.org/10.1145/3564721.3564733>.
- Bennedsen, J., & Caspersen, M. (2007). Failure rates in introductory programming. *SIGCSE Bulletin*, 39, 32–36. <https://doi.org/10.1145/1272848.1272879>
- Bennedsen, J., & Caspersen, M. (2019). Failure rates in introductory programming: 12 years later. *ACM Inroads*, 10(2), 30–36. <https://doi.org/10.1145/3324888>
- Biesta, G. J. J. (2010). Why “what works” still won't work: From evidence-based education to value-based education. *Studies in Philosophy and Education*, 29, 491–503. <https://doi.org/10.1007/s11217-010-9191-x>
- Bodily, R., Verbert, K. (2017). Trends and issues in student-facing learning analytics reporting systems research. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK'17)* (pp. 309–318). New York: Association for Computing Machinery. <https://doi.org/10.1145/3027385.3027403>
- Bowers, A. J. (2021). 9. Early warning systems and indicators of dropping out of upper secondary school: The emerging role of digital technologies. OECD Digital Education Outlook. <https://doi.org/10.1787/589b283f-en>
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5/6), 318–331. <https://doi.org/10.1504/IJTEL.2012.051815>
- Choi, S. P. M., Lam, S. S., Li, K. C., & Wong, B. T. M. (2018). Learning analytics at low cost: At-risk student prediction with clicker data and systematic proactive interventions. *Educational Technology & Society*, 21(2), 273–290.
- Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/2330601.2330636>
- Cobos, R., & Ruiz-Garcia, J. (2020). Improving learner engagement in MOOCs using a learning intervention system: A research study in engineering education. *Computer Applications in Engineering Education*. <https://doi.org/10.1002/cae.22316>
- Combéfis, S. (2022). Automated Code Assessment for Education: Review. *Classification and Perspectives on Techniques and Tools, Software*, 1(1), 3–30. <https://doi.org/10.3390/software1010002>
- Dodge, B., Whitmer, J., & Frazee, J. P. (2015). Improving undergraduate student achievement in large blended courses through data-driven interventions. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 412–413).
- Dvorak, T., & Jia, M. (2016). Do the timeliness, regularity, and intensity of online work habits predict academic performance? *Journal of Learning Analytics*, 3(3), 318–330.
- Foster, E., & Siddle, R. (2019). The effectiveness of learning analytics for identifying at-risk students in higher education. *Assessment & Evaluation in Higher Education*, 45, 1–13. <https://doi.org/10.1080/02602938.2019.1682118>
- Froissard, C., Richards, D., Atif, A., & Liu, D. (2015). An enhanced learning analytics plugin for Moodle: Student engagement and personalized intervention. In *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE) Conference* (pp. 259–263). Retrieved from <https://ascilite.org/conferences/perth2015/index.php/program/ascilite2015/paper/view/194>
- 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.
- Heikinen, S., Saqr, M., Malmberg, J., et al. (2022). Supporting self-regulated learning with learning analytics interventions—a systematic literature review. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-022-11281-4>
- Hellas, A., Airaksinen, J., & Watson, C. (2014). A systematic review of approaches for teaching introductory programming and their influence on success. In *ICER 2014—Proceedings of the 10th Annual International Conference on International Computing Education Research*. <https://doi.org/10.1145/2632320.2632349>.
- Hellings, J., & Haelermans, C. (2022). The effect of providing learning analytics on student behaviour and performance in programming: a randomised controlled experiment. *Higher Education*. <https://doi.org/10.1007/s10734-020-00560-z>
- Herodotou, C., Rienties, B., Boroowa, A., et al. (2019). A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective. *Educational Technology Research and Development*, 67, 1273–1306. <https://doi.org/10.1007/s11423-019-09685-0>
- Hsu, T. Y., Chiou, C. K., Tseng, J. C. R., & Hwang, G. J. (2016). Development and evaluation of an active learning support system for context-aware ubiquitous learning. *IEEE Transactions on Learning Technologies*, 9(1), 37–45.
- Huang, C. S. J., Yang, S. J. H., Chiang, T. H. C., & Su, A. Y. S. (2016). Effects of Situated mobile learning approach on learning motivation and performance of EFL students. *Educational Technology & Society*, 19(1), 263–276.
- Ifenthaler, D., & Yau, J. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-020-09788-z>
- Kennedy, G., Corrin, L., Lockyer, L., Dawson, S., Williams, D., Mulder, R., Khamis, S., & Copeland, S. (2014). Completing the loop: Returning learning analytics to teachers.
- Kew, S., Tasir, Z. (2017). A systematic review of learning analytics intervention contributing to student success in online learning (pp. 62–68). <https://doi.org/10.1109/LaTiCE.2017.18>
- Khalil, M., & Ebner, M. (2015). Learning analytics: Principles and constraints. In *Proceedings of world conference on educational multimedia, hypermedia and telecommunications 2015* (pp. 1326–1336).
- Kiemer, K., & Kollar, I. (2021). Source selection and source use as a basis for evidence-informed teaching: Do pre-service teachers' beliefs regarding the utility of (non-)scientific information sources matter? *Zeitschrift Für Pädagogische Psychologie*, 35, 1–15. <https://doi.org/10.1024/1010-0652/a000302>
- Kimberly, E. A., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge—LAK'12* (pp. 267–270)
- Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college. *Journal of Advanced Academics*, 20(1), 42–68.

- Kizilcec, R., Reich, J., Yeomans, M., Dann, C., Brunskill, E., Lopez, G., Turkay, S., Williams, J., & Tingley, D. (2020). Scaling up behavioral science interventions in online education. *Proceedings of the National Academy of Sciences*, 117, 201921417. <https://doi.org/10.1073/pnas.1921417117>
- Klang, N., Åsman, J., Mattsson, M., Nilholm, C., & Folkeryd, J. (2022). Intervention combining cooperative learning and instruction in reading comprehension strategies in heterogeneous classrooms. *Nordic Journal of Literacy Research*, 8, 44–64. <https://doi.org/10.23865/njlr.v8.2740>
- Kvernbekk, T. (2015). Evidence-based practice in education. *Functions of Evidence and Causal Presuppositions*. <https://doi.org/10.4324/9780203774830>
- Layman, L., Song, Y., & Guinn, C. (2020). Toward predicting success and failure in CS2: A mixed-method analysis. In *Proceedings of the 2020 ACM Southeast Conference (ACM SE'20)* (pp. 218–225). Association for Computing Machinery. <https://doi.org/10.1145/3374135.3385277>
- Li, K. C., Ye, C. J., & Wong, B. T. M. (2018). Status of learning analytics in Asia: Perspectives of higher education stakeholders. In *Technology in education: Innovative solutions and practices* (pp. 267–275).
- Loksa, D., Margulieux, L., Becker, B. A., Craig, M., Denny, P., Pettit, R., & Prather, J. (2022). Metacognition and self-regulation in programming education: Theories and exemplars of use. *ACM Transactions on Computing Education (TOCE)*, 22(4), 39. <https://doi.org/10.1145/3487050>
- Lonn, S., Aguilar, S. J., & Teasley, S. D. (2015). Investigating student motivation in the context of a learning analytics intervention during a summer bridge program. *Computers in Human Behavior*, 47, 90–97.
- Lu, O. H. T., Huang, J. C. H., Huang, A. Y. Q., & Yang, S. J. H. (2017). Applying learning analytics for improving students' engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, 25(2), 220–234.
- Lynch, M. (2019). Types of classroom intervention. <https://www.theedadvocate.org/types-of-classroom-interventions/>
- Majumdar, R., Akçapınar, A., Akçapınar, G., Flanagan, B., & Ogata, H. (2019). LAVIEW: Learning analytics dashboard towards evidence-based education. In *Proceedings of the 9th international conference on learning analytics and knowledge* (pp. 500–501). ACM. <https://doi.org/10.1145/3303772.3306212>
- Masters, G. (2018). The role of evidence in teaching and learning. Teacher columnist—Geoff Masters. <https://research.acer.edu.au/columnists/39>
- Milliron, M. D., Malcolm, L., & Kil, D. (2014). Insight and action analytics: Three case studies to consider. *Research and Practice in Assessment*, 9, 70–89.
- Bretana, N.A., Robati, M., Rawat, A., Panday, A., Khatri, S., Kaushal, K., Nair, S., Cheang, G., Abadia, R. (n.d.). *Predicting student success for programming courses in a fully online learning environment*, UniSA STEM, University of South Australia
- Nam Liao, S., Shah, K., Griswold, W. G., & Porter, L. (2021). A quantitative analysis of study habits among lower- and higher-performing students in CS1. In *Proceedings of the 26th ACM conference on innovation and technology in computer science education (ITiCSE'21)* (Vol. 1, pp. 366–372). <https://doi.org/10.1145/3430665.3456350>
- Nielsen, T. (2018). The intrinsic and extrinsic motivation subscales of the motivated strategies for learning questionnaire: A rasch-based construct validity study. *Cogent Education*, 5, 1. <https://doi.org/10.1080/2331186X.2018.1504485>
- Obaido, G., Agbo, F. J., Alvarado, C., & Oyelere, S. S. (2023). Analysis of attrition studies within the computer sciences. *IEEE Access*, 11, 53736–53748. <https://doi.org/10.1109/ACCESS.2023.3280075>
- OECD (2019). How many students complete tertiary education? In *Education at a Glance 2019: OECD indicators*. Paris: OECD Publishing. <https://doi.org/10.1787/62cab6af-en>
- Othman, H., Hamid, A., Budin, S., & Rajab, N. (2011). The effectiveness of learning intervention program among first year students of biomedical science program. *Procedia—Social and Behavioral Sciences*, 18, 367–371. <https://doi.org/10.1016/j.sbspro.2011.05.052>
- Outhwaite, L., Gulliford, A., & Pitchford, N. (2019). A new methodological approach for evaluating the impact of educational intervention implementation on learning outcomes. *International Journal of Research & Method in Education*, 43, 1–18. <https://doi.org/10.1080/1743727X.2019.1657081>
- Pritchard, Alan (2014) [2005]. *Learning styles. Ways of learning: Learning theories and learning styles in the classroom* (3rd edn., pp. 46–65). New York: Routledge
- Rienties, B. & Cross, S., & Zdráhal, Z. (2017) Implementing a learning analytics intervention and evaluation framework: What works? https://doi.org/10.1007/978-3-319-06520-5_10.
- Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children*, 82(2), 144–169.
- Rodriguez-Planas, N. (2012). Mentoring, educational services, and incentives to learn: What do we know about them? *Evaluation and Program Planning*, 35, 481–490. <https://doi.org/10.1016/j.evalprogplan.2012.02.004>
- Restrepo-Calle, F., Ramirez-Echeverry, J. J., & González, F. (2018). Continuous assessment in a computer programming course supported by a software tool. *Computer Applications in Engineering Education*. <https://doi.org/10.1002/cae.22058>
- Sacr, M., Fors, U., Tedre, M., & Nouri, J. (2018). How social network analysis can be used to monitor online collaborative learning and guide an informed intervention. *PLoS ONE*, 13(3), e0194777.
- Şahin, M., & Yurdugül, H. (2019). An intervention engine design and development based on learning analytics: The intelligent intervention system (In2S). *Smart Learning Environments*, 6, 18. <https://doi.org/10.1186/s40561-019-0100-7>
- Salguero, A., Griswold, W. G., Alvarado, C., & Porter, L. (2021). Understanding sources of student struggle in early computer science courses. In *Proceedings of the 17th ACM conference on international computing education research (ICER 2021)* (pp. 319–333). Association for Computing Machinery. <https://doi.org/10.1145/3446871.3469755>
- Santana, B., Figueredo, J., & Bittencourt, R. (2018). Motivation of engineering students with a mixed-contexts approach to introductory programming. *IEEE Frontiers in Education Conference (FIE)*. <https://doi.org/10.1109/FIE.2018.8659158>
- Slater, N., & Mullan, J. (2017). Learning analytics and student success-Assessing the evidence. Retrieved from <http://analytics.jiscinvolve.org/wp/files/2015/07/jisc-la-network-ed-foster-ntu.pdf>
- Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, 107, 2020. <https://doi.org/10.1016/j.chb.2018.05.004>

- Shaidullina, A. R., Orekhovskaya, N. A., Panov, E. G., Svintsova, M. N., Petyukova, O. N., Zhuykova, N. S., & Grigoryeva, E. V. (2023). Learning styles in science education at university level: A systematic review. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), 02293. <https://doi.org/10.29333/ejmste/13304>
- Sharma, M., & Tiwari, P. (2021). A study of class interaction analysis using Flanders's FIAC. *International Journal of Scientific Research in Science, Engineering and Technology*. <https://doi.org/10.32628/IJSRSET218432>
- Siemens, G., & Gasevic, D. (2012). Guest editorial-learning and knowledge analytics. *Journal of Educational Technology & Society*, 15(3), 1–2.
- Sonderlund, A. L., Hughes, E., & Smith, J. (2018). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*.
- Sonnenberg, C., & Bannert, M. (2019). Using Process Mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Computers in Human Behavior*, 96(2019), 259–272. <https://doi.org/10.1016/j.chb.2018.06.003>
- Stephenson, C., Derbenwick Miller, A., Alvarado, C., Barker, L., Barr, V., Camp, T., Frieze, C., Lewis, C., Cannon Mindell, E., Limbird, L., Richardson, D., Sahami, M., Villa, E., Walker, H., & Zweben, S. (2018). Retention in computer science undergraduate programs in the U.S.: Data challenges and promising interventions. New York: Association for Computing Machinery
- Sullivan, G. (2011). Getting off the "gold standard": randomized controlled trials and education research. *Journal of Graduate Medical Education*, 3, 285–289. <https://doi.org/10.4300/JGME-D-11-00147.1>
- Szabo, C., Falkner, N., Knutas, A., & Dorodchi, M. (2017). Understanding the Effects of lecturer intervention on computer science student behaviour (pp. 105–124). <https://doi.org/10.1145/3174781.3174787>.
- Takacs, R., Kárász, J. T., Takács, S., Horváth, Z., & Oláh, A. (2022). Successful steps in higher education to stop computer science students from attrition. *Interchange*, 53, 1–16. <https://doi.org/10.1007/s10780-022-09476-2>
- Tempelaar, D., Nguyen, Q., & Rienties, B. (2020). Learning Analytics and the Measurement of Learning Engagement. In D. Ifenthaler & D. Gibson (Eds.), *Adoption of Data Analytics in Higher Education Learning and Teaching Advances in Analytics for Learning and Teaching*. Cham: Springer. https://doi.org/10.1007/978-3-030-47392-1_9
- Triana, R. (2008). Evidence-based policy: A realist perspective, by Ray Pawson. *Journal of Policy Practice*, 7(4), 321–323. <https://doi.org/10.1080/15588740802262039>
- UNESCO (2022). Early warning systems for school dropout prevention in Latin America and the Caribbean. https://unesdoc.unesco.org/ark:/48223/pf0000380354_eng
- Utamachant, P., Anutariya, C., Pongnumkul, S., & Sukvaree, N. (2020). Analyzing online learning behavior and effectiveness of blended learning using students' assessing timeline. *International Symposium on Project Approaches in Engineering Education*, 10, 64–71.
- U.S. Department of Education (2016). Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments. Retrieved from <https://www2.ed.gov/policy/elsec/leg/essa/guidanceusesinvestment.pdf>
- Watson, C. & Li, F.W.B. (2014) Failure rates in introductory programming revisited. In *Proceedings of the 2014 conference on Innovation technology in computer science education (ITICSE'14)* (pp. 39–44). New York: Association for Computing Machinery (ACM). <https://doi.org/10.1145/2591708.2591749>
- Wills, H. P., Caldarella, P., Mason, B. A., Lappin, A., & Anderson, D. H. (2019). Improving student behavior in middle schools: Results of a classroom management intervention. *Journal of Positive Behavior Interventions*, 21(4), 213–227. <https://doi.org/10.1177/1098300719857185>
- Wong, B. T. (2017). Learning analytics in higher education: An analysis of case studies. *Asian Association of Open Universities Journal*, 12(1), 21–40.
- Wong, B. T., & Li, K. C. (2020). A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education*, 7, 7–28. <https://doi.org/10.1007/s40692-019-00143-7>
- Zhang, J.-H., Zhang, Y.-X., Zou, Q., & Huang, S. (2018). What learning analytics tells Us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. *Educational Technology & Society*, 21(2), 245–258.
- Zhang, J.-H., Zou, L.-C., Miao, J.-J., Zhang, Y.-X., Hwang, G.-J., & Zhu, Y. (2020). An individualized intervention approach to improving university students' learning performance and interactive behaviors in a blended learning environment. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2019.1636078>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)