Higher Education
https://doi.org/10.1007/s10734-026-01709-y

Applications of learning analytics in the study of academic performance in higher education: A meta-review with an educational perspective

Fran J. García-García[1] **· María Isabel Gómez-Núñez**[2] **· Cristian Molla-Esparza**[3]

Received: 11 August 2025 / Accepted: 8 June 2026
© The Author(s) 2026

Abstract
This meta-review synthesizes evidence from systematic reviews (SRs) examining the application of Learning Analytics (LA) to the study of academic performance in higher education from an educational perspective. Despite growing interest in this area, existing reviews vary considerably in scope and exhibit notable substantive and methodological shortcomings. A comprehensive search strategy, up until May 2025, was conducted across four major databases (Scopus, Web of Science, ERIC, and PsycINFO), and gray literature. A narrative synthesis was employed to analyze findings from 19 SRs that met the inclusion criteria. The reviewed SRs paid limited attention to key educational aspects, such as instructional contexts, pedagogical aims, disciplinary domains, underlying educational theories, and student learning stages. They were also characterized by critically low methodological quality. Considerable variability was found in how academic performance was conceptualized and operationalized, ranging from broad constructs to clearly defined indicators. Demographic and socioeconomic variables, followed by academic background and performance indicators, were the most frequently examined factors, although their predictive power was rarely reported. Supervised quantitative approaches, particularly regression and classification-based techniques. However, the predominantly narrative and heterogeneous reporting of non-comparable performance metrics precluded robust conclusions. The educational implications of these patterns are discussed, emphasizing the need for future research to incorporate contextually relevant educational dimensions, multi-source data, and multivariate and multilevel analyses. Overall, the results underscore the importance of aligning LA research with pedagogical objectives and enhancing its practical relevance in higher education.

Keywords Learning Analytics · Academic Performance · Higher Education · Meta-review

Introduction

Over the past two decades, higher education institutions have progressively integrated digital platforms into their programs to support and enhance teaching and learning processes. These platforms have enabled a systematic collection of extensive student-related

Extended author information available on the last page of the article

Published online: 22 June 2026

data, ranging from patterns of interaction with educative digital contents through to academic performance indicators. Against this backdrop, Learning Analytics (henceforth LA) has emerged as a distinct, interdisciplinary field concerned with the measurement, collection, analysis and interpretation of educational data with the aim of understanding and improving learning processes, and the environments in which they occur (Lang et al., 2022; Long et al., 2011).

Initially, LA in higher education was limited to basic functionalities within learning management systems (LMS), such as Moodle and Blackboard, focusing primarily on simple metrics such as attendance and grades. However, advances in data infrastructures, computational capabilities and artificial intelligence have facilitated the development of more sophisticated platforms, such as edX Insights, IBM Watson Education, and Learning Locker, which employ semi-automated or fully automated processes that can manage large-scale datasets. As a result, LA has evolved from a largely descriptive tool to one capable of supporting real-time analyses of student engagement, personalized learning pathways, and academic competences (Haleem et al., 2022; Mhlanga, 2024). This transformation has expanded LA's scope from passive tracking (Viberg et al., 2018) to the deployment of predictive and prescriptive models that can forecast academic outcomes and recommend tailored, data-based interventions (Ouyang et al., 2023).

Contemporary research on LA in higher education converges in three principal domains: (1) the personalization of learning experiences through the monitoring of student progression (Banihashem et al., 2022); (2) the iterative design of curricula informed by data-driven insights (Leitner et al., 2017); and (3) the optimization of institutional resources through the identification of usage patterns and systemic inefficiencies (Llopis-Albert & Rubio, 2021). Collectively, these applications underscore LA's potential to inform data-informed decision-making through pedagogical innovation (Utamachant et al., 2023; Yan et al., 2024).

Despite LA's growing prominence, literature assessing its impact, particularly through evidence synthesis methods (e.g., systematic literature reviews), reveals persistent conceptual and methodological challenges (Chan et al., 2019; Foster & Francis, 2020; Ifenthaler & Yau, 2020). While a number of reviews have confirmed LA's efficacy in identifying academic performance trends and modeling student success (Alnasyan et al., 2024; Dol & Jawandhiya, 2024; Rahul & Katarya, 2023), significant inconsistencies remain in the accuracy of models. Such discrepancies are often attributed to heterogeneity: in data sources (e.g., institutional records, LMS, tutoring systems, game-based learning environments), data types (e.g., clickstream logs, assessment scores, forum posts, sensor and biometric data), and input variables (e.g., demographic data, prior academic performance, engagement metrics, emotional states); in how academic performance is operationalized (e.g., final course grades, retention or dropout status, formative assessment outcomes); and even in applied analytical techniques (e.g., association rules, classification algorithms) (Larrabee Sonderlund et al., 2019; Namoun & Alshanqiti, 2020). Moreover, several theoretical and practical issues remain understudied. For example, many studies lack robust theoretical foundations with defined pedagogical objectives, limiting the capacity of LA to inform pedagogically meaningful interventions (Gašević et al., 2017; Reimann, 2016). Indeed, without addressing detailed characteristics related to the educational context, LA risks reinforcing techno-solutionist paradigms that prioritize algorithmic efficiency over pedagogical relevance.

In light of such challenges, and in order to address existing knowledge gaps, a comprehensive meta-review of existing synthesis studies is warranted. To date, no meta-reviews have been carried out on LA applications in the study of academic performance specifically in higher education, although two previous studies have come close to addressing this goal (Chaka, 2022; Du et al., 2019, 2021), collecting data up until 2021 at the latest. Du et al. (2019, 2021), without a predefined review protocol, conducted a meta-review covering the period from 2012 to 2017, in which 901 publications were examined. However, this meta-review combined empirical studies with only eight review papers, and lacked a clear focus on academic performance. Chaka (2022) brought improvements in methodological rigor by employing a standardized protocol, yet his review was limited to predictive methods.

Existing meta-reviews thus exhibit notable substantive and methodological limitations, particularly in terms of their thematic focus and analytical scope. While systematic reviews (henceforth SRs) are designed to synthesize evidence, they frequently diverge in terms of scope, methodological approaches, and inclusion criteria, resulting in fragmented, overlapping, or even contradictory conclusions. This underscores the need for a meta-review specifically focused on the applications of LA within a defined educational context, namely, higher education, and addressing a targeted academic domain, such as academic performance.

Ultimately, this meta-review contributes to a more integrated understanding of LA's role in higher education by examining both substantive dimensions, such as phase of the educational process in which LA was applied, the knowledge domains involved, and any underpinning pedagogical theories and objectives, and operational aspects, including the data sources and types, and the predictive variables and performance metrics employed, as well as the analytical strategies used. Furthermore, an in-depth study of LA from a more educational perspective promises to provide a theoretical framework for its use as a mechanism to analyze the academic development of higher education students. Finally, from a practical standpoint, it contributes to the optimization of LA as a set of fundamental educational procedures for personalizing university students' teaching and learning processes. The main objective of this meta-review is thus to examine SRs of the applications of LA in the study of academic performance in higher education.

Method

Meta-review protocol

This meta-review was conducted in accordance with a synthetic planning protocol preregistered in the INPLASY repository (number 2024120119 ; doi: https://doi.org/10.37766/inplasy2024.12.0119), as well as with a previously published empirically tested protocol article (Molla et al., 2025). The reporting of this meta-review was based on the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidelines (Page et al., 2021), and more specifically on the *Preferred Reporting Items for Overviews of Reviews* (PRIOR) statements (Gates et al., 2022; Pollock et al., 2019). Additional methodological recommendations specifically proposed for overviews of reviews were also taken into account (Aromataris et al., 2015).

Consistent with the principles of open science, and to enhance transparency and reproducibility, all procedures and materials referenced in this meta-review are openly accessible in the OSF repository by clicking here https://osf.io/dejnv/overview?view_only=b9e5e7314ade41e8b69f3979e3f4a403.

Information sources and search strategies

Both the selection of databases and the search strategy employed were based on the meta-review protocol study, in which a university librarian and a researcher with expertise in evidence synthesis conducted a critical review and testing process (Molla et al., 2025), assessing the adequacy of database selection, overall coverage and precision of the search syntax, the appropriateness of thesaurus descriptors, free-text keywords, truncation symbols, and Boolean operators, and the suitability of a preliminary retrieval of records. Refinements made to the search strategy can be consulted in the published meta-review protocol study.

Bibliographic searches were carried out across four electronic databases, Scopus (Elsevier), Web of Science Core Collection (via Clarivate Analytics), ERIC (via ProQuest), and PsycINFO (via ProQuest), to cover a comprehensive coverage of disciplines relevant to the subject of study (e.g., education, technology, informatics, data science). Gray literature was also explored using Google (screening the first 250 results sorted by relevance) and Dissertations & Theses Global (via ProQuest), as well as other specific sources related to LA, such as the Proceedings of the International Conference on Learning Analytics & Knowledge (LAK). Emails were sent to leading researchers and research groups with a substantial publication record on LA in order to inquire about ongoing or unpublished SRs that might meet the inclusion criteria. In addition, the reference lists of all included studies and other relevant publications were manually screened to identify additional potentially eligible reviews.

The search strategy was developed in accordance with the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension* (PRISMA-S) guidelines (Rethlefsen et al., 2021). It was designed through an iterative process that incorporated key terms from the empirical literature and consultations of relevant databases, as well as from search strategies employed in previous SRs on LA in higher education (Chaka, 2022; Ifenthaler & Yau, 2017; Pan et al., 2024). Four main groups of descriptors were therefore identified, combining both highly sensitive and specific terms related to the following dimensions: (1) Learning Analytics (e.g., learning analytics, academic analytics, data-driven education, educational data mining); (2) type of study (e.g., systematic review, scoping review, integrative review); (3) academic performance (e.g., academic grades, academic marks, academic outcomes, academic performance); and (4) higher education (e.g., tertiary education, undergraduates, university). Boolean operators (i.e., OR, AND) were applied both within and across groups of descriptors, respectively.

The search strategy was adjusted to the specific syntax, characteristics and functionalities of each bibliographic database (e.g., search fields, combinations of free text and thesaurus terms, truncation symbols), limited to publications from 1st January 2000 to 30th April 2025. Monthly alerts were set in all four databases in order to identify new records published during the review process, up until 21st May 2025. To support the reproducibility of this meta-review, the full search strategies applied in each of the databases are presented in the Supplementary Tables S1a to S1e.

Eligibility criteria and selection process

To address the stated objectives of this meta-review, studies were deemed eligible for inclusion if they: (1) were SRs, regardless of the specific review type (e.g., scoping review, integrative review); (2) were available in any document format (e.g., journal article, conference proceeding); (3) SRs focusing exclusively on higher-education settings; (4) were published in English or Spanish; and (5) were published from the year 2000 onward.

Conversely, studies were excluded if they primarily focused on: (1) bibliometric analyses; (2) non-human samples, such as software, tools, or algorithms; or (3) Massive Open Online Courses (MOOCs), unless these courses were explicitly embedded within formal higher education curricula; or (4) mixed educational levels (e.g., higher education combined with secondary or other educational contexts), unless results were reported separately for higher education.

All the retrieved records were imported into Zotero software (version 7.0.8), where duplicate references were removed prior to the screening process. Two reviewers (FGG & MIGN) independently screened all titles and abstracts against the inclusion criteria, following the predefined adjudication rules described in the meta-review protocol (Molla-Esparza et al., 2025). Thus, records advanced to full-text screening under three circumstances: (1) both reviewers judged the study to be eligible; (2) one reviewer expressed uncertainty; or (3) the reviewers reached different decisions. Screening based on titles and abstracts yielded good inter-rater reliability (Gwet's AC1 = 0.92, 95% CI: 0.91, 0.94, $p < .001$), with an observed agreement of 93% between the reviewers (Gwet, 2014). Full-text screening was conducted independently by the same two reviewers (FGG & MIGN). Cases involving disagreement or uncertainty were discussed with the aim of reaching consensus, and, when consensus could not be reached, a third reviewer (CME) was consulted to make the final decision. The full-text screening phase also demonstrated good inter-rater reliability (Gwet's AC1 = 0.65, 95% CI: 0.52, 0.78, $p < .001$), with an observed agreement of 72% between reviewers. Supplementary Table S2 indicates all the references excluded, while Supplementary Table S3 indicates all those selected for final inclusion in the meta-review after full-text screening.

The degree of overlap of primary studies across the included SRs was evaluated both using artificial intelligence tools (specifically the NotebookLM platform from Google Research), and manually. The resulting Corrected Covered Area value was 0.91%, indicating a slight overlap (Pieper et al., 2014). Consequently, there was no substantial overlap that would justify concerns regarding duplication, or warrant further examination. This assessment contributed to the interpretation of the results, particularly in terms of their validity.

Data extraction and quality assessment

Data were extracted from the included SRs using a piloted data extraction form by two independent reviewers (FGG & CME), with discrepancies between coders addressed by trying to reach consensus, and arbitration by a third reviewer (MIGN) when required. The data coding form was pilot-tested in two consecutive rounds in order to minimize potential errors and enhance the reliability of the process, and a preliminary satisfactory inter-coder agreement result was obtained (Molla-Esparza et al., 2025). Supplementary Table S4 shows the data extraction form used in the meta-review.

For coding purposes, bibliometric data included: (1) authors' names, affiliations, and subject areas; (2) publication year; (3) document type (e.g., article, conference proceeding); (4) source; (5) language; and (6) open access status. Regarding the characteristics of the SRs, the following data were coded: (7) type of SR (e.g., scoping review); (8) adherence to a reporting guideline (e.g., PRISMA); (9) research aims; (10) eligibility criteria; and (11) the number of primary studies included. To gain a deeper understanding of LA applications in the study of academic performance in higher education, the following data were also extracted: (12) higher educational levels at which LA was applied (e.g., undergraduate, postgraduate); (13) knowledge domains involved (e.g., education, medicine); (14) educational settings (e.g., online, blended); (15) educational theories; (16) objectives of the primary studies; (17) pedagogical purposes guiding LA applications; (18) phase of the educational process in which LA was applied; (19) data collection methods and sources (e.g., LMS); (20) parties generating the data and units of analysis (e.g., individual students, groups); (21) communication channels used (e.g., web-based dashboards); (22) types of data gathered (e.g., demographic, performance data, engagement data); (23) analyzed input variables (e.g., logins) and output variables (e.g., course completion rates, grade point averages), and those identified as statistically significant; (24) LA methods used (e.g., self-organizing maps, social network analysis); (25) measures of the accuracy and precision of methods (i.e., statistic performance metrics); and (26) software tools employed (e.g., R, Stata). Gwet's coefficient for data extraction was 0.90 (95% CI: 0.87, 0.93, $p < .001$), with an observed agreement of 91% between coders.

Following the same data coding process, the methodological quality of the included SRs was assessed using the *A Measurement Tool to Assess Systematic Reviews* (AMSTAR 2) guidelines (Shea et al., 2017). Minor semantic adjustments were made to adapt the tool to the LA literature (e.g., replacing 'patients' with 'users', and removing medical terms such as 'dose'). AMSTAR 2 comprises 16 domains evaluating key aspects such as research question formulation, protocol registration, search strategy comprehensiveness, duplicate screening and data extraction, specification of eligibility criteria, characterization of primary studies, risk of bias assessment, and disclosure of funding sources. Overall quality was classified in line with the AMSTAR 2 guidance, which identifies domains 2, 4, 7, 9, 11, 13, and 15 as critical. The methodological quality of each SR was thus categorized on the basis of the number and severity of its methodological weaknesses: high (no or one non-critical weakness); moderate (more than one non-critical weakness); low (one critical weakness, with or without non-critical weaknesses); and critically low (more than one critical weakness). The methodological quality assessment demonstrated good overall agreement between coders (Gwet's AC1 = .93, 95% CI: .91, .96, $p < .001$), with an observed agreement of 93%. The overall methodological quality and domain-level scores are provided in Table 1 and Supplementary Table S5, respectively.

Data synthesis

Due to the expected heterogeneity of LA applications in the study of academic performance in higher education, the data extracted from the included SRs were synthesized narratively. The narrative synthesis was guided by the research objectives of the meta-review, and was structured around the aforementioned predefined dimensions. In particular, the synthesis began with a description of search results, study characteristics (e.g., publication years, document types, reporting guidelines employed), and their methodological quality. It then addressed educational contexts and pedagogical aims targeted by LA applications,

Table 1 Main characteristics of the included systematic reviews

Study	Authors' subject area	Study type	Reporting guideline	Main aim	Inclusion criteria	Nº studies	Qual. score
Abu Saa et al. (2019)	Eng	SR-A	Kitchenham et al. (2009) and PRISMA	To identify key factors influencing student performance and learning outcomes in HE through a review of EDM literature.	Studies using data mining or ML; reporting relevant factors; full-text available; non-review; published 2009–2018; English language.	36	Cr. Low
Alhakbani & Alnassar (2022)	CS	SR-CP	PRISMA	To review benchmark studies using the OULAD dataset to examine student performance in virtual learning environments.	Peer-reviewed English-language journal articles, conference papers, or book chapters.	18	Cr. Low
Arizmendi et al. (2023)	LA	SR-A	NR	To analyze how LMS data can be mapped and used to predict student outcomes and guide interventions.	Studies up to 2020 predicting undergraduate course performance, excluding satisfaction, engagement, or multi-course outcomes.	39	Cr. Low
Blumenstein (2020)	Sci	SR-A	Kitchenham and Charters's (2007)	To assess effect sizes in LA approaches measuring student learning gain in HE.	English-language empirical studies (2011–2019) in peer-reviewed sources, involving student participants and academic subject content.	25	Cr. Low
Cardona et al. (2023)	Eng & Math	SR-A	Tranfield et al. (2003)	To review ML techniques for predicting student retention and identifying factors associated with completion rates.	Peer-reviewed English or Spanish journal or conference papers published up to 2018, applying ML to student retention prediction.	19	Cr. Low
Chan et al. (2019)	HS	SR-A	NR	To evaluate the use of e-LA in healthcare education and its relationship to learning outcomes.	Studies (2000–2017) involving health-profession students, using LA/LMS technologies, and reporting quantitative academic outcomes.	19	Cr. Low
Contreraras-B. et al. (2022)	Eng	SR-A	NR	To identify commonly used algorithms and relevant variables for academic performance prediction.	Publications (2016–2021) using ensemble algorithms for academic performance prediction; excluding dropout-only studies.	54	Cr. Low
Dhankhar et al. (2021)	Eng & CS	SR-CP	Kitchenhams (2007)	To review techniques, attributes, metrics and the state of the art in EDM and LA for student performance prediction.	Empirical studies (2016–2020) in HE engineering/university contexts implementing performance prediction models and reporting accuracy.	84	Cr. Low

Table 1 (continued)

Study	Authors' subject area	Study type	Reporting guideline	Main aim	Inclusion criteria	Nº studies	Qual. score
Fahd et al. (2022)	Eng & Bus	SR & MA-A	PRISMA	To review ML applications predicting at-risk students, academic performance and attrition in HE.	Peer-reviewed English journal and conference articles (2010–2020) applying ML to predict performance, risk and attrition in HE.	89	Cr. Low
Foster and Francis (2020)	Arts & Bus	SR-A	NR	To assess the effectiveness of Educational Analytics interventions aimed at improving student outcomes.	Primary research published 2007–2018 (journals, conferences, reports, case studies, theses).	34	Cr. Low
Ifenthaler and Yau (2020)	Edu	SR-A	Okoli and Schabram (2010)	To summarize empirical evidence on how LA supports study success, continuation, and completion.	Peer-reviewed English publications (2013–2018) with abstracts, using qualitative or quantitative analysis in the context of HE.	46	Cr. Low
Jin et al. (2024)	Edu & CS	SR-CP	PRISMA	To review recent predictive modeling studies using the Open University LA Dataset (OULAD).	Peer-reviewed English publications (2017–2024) using the OULAD dataset for predictive modeling.	17	Cr. Low
Johar et al. (2023)	Edu & LA	SR-A	PRISMA	To review LA for student engagement and performance in online-learning environments.	Empirical English-language studies (2011–2021) applying LA to online student engagement in HE.	42	Cr. Low
Mella-Norm. et al. (2022)	Edu & Sci	SR-A	PRISMA	To analyze the characteristics of predictive models based on LA in HE.	Quantitative empirical studies published up to 2021 using LA-based predictive designs with university students and LMS data.	12	Cr. Low
Pelima et al. (2024)	Eng & CS	SR-A	Kitchenham and Brereton (2013)	To review EDM research on academic performance prediction, including data sources, variables, and methodological gaps.	Peer-reviewed English-language journal studies (2018–2023) on student performance or graduation prediction in HE using ML.	70	Cr. Low
Rahul and Katarya (2023)	Eng & CS	SR-A	NR	To evaluate ML methods applied to student performance analysis.	Studies published 2010–2020 specifying the algorithms used.	109	Cr. Low

Table 1 (continued)

Study	Authors' subject area	Study type	Reporting guideline	Main aim	Inclusion criteria	Nº studies	Qual. score
Sghir et al. (2023)	Eng & CS	SR-A	PRISMA	To identify research clusters on predictive LA supporting student learning processes.	Peer-reviewed English-language studies in HE (2012–2022), using ML for learning-outcome prediction with clear methods.	74	Cr. Low
Tjandra et al. (2022)	Eng & CS	CR-CP	NR	To review student performance prediction studies, including tasks, predictors, methods, accuracy, and tools.	Studies published between 2002 and 2020 in reputable journals or conferences, and meeting methodological quality standards.	34	Cr. Low
Zulkifli et al. (2019)	Math & CS	SR-A	NR	To review predictive models of students' academic performance in HE.	English-language journal or conference studies in HE modeling student performance using simulated or real datasets (2014–2018).	69	Cr. Low

SR =Systematic Review, *MA* =Meta-analysis, *CR* =Comprehensive Review, *-A* =Article, *-CP* =Conference Proceeding, *NR* =Not reported, *PRISMA* =Preferred Reporting Items for Systematic reviews and Meta-Analyses, *ML* =Machine Learning, *LA* =Learning Analytics, *EDM* =Educational Data Mining, *HE* =Higher education, *LMS* =Learning Management Systems, *Eng* =Engineering, *Math* =Mathematics, *CS* =Computer Science, *HS* =Health Science, *Edu* =Education, *Sci* =Sciences, *Bus* =Business, *Qual. score* =AMSTAR2 quality score, *Cr. Low* =Critically low including institutional levels, disciplines, learning environments, and stated educational purposes. Furthermore, results concerning the generation, collection, and management of data for LA were gathered, reporting the systems and platforms used, as well as the types of data gathered.

The synthesis also examined how academic performance was conceptualized or operationalized, and the factors studied in relation to it. Indicators were extracted verbatim, and ambiguity in SR terminology was coded explicitly. Outcome indicators were additionally examined in terms of whether the included SRs reported clearly operationalized indicators or broader conceptual labels lacking explicit operational definitions. Then, all indicators were classified into three mutually exclusive outcome families: (1) academic performance and learning outcomes; (2) progression and retention; and (3) engagement and behavior. Factors examined in relation to academic performance were synthesized and grouped into meta-categories, such as demographics and socioeconomics, and digital interaction and log data. The synthesis also reported LA methods used and the performance of these methods. LA methods (e.g., random forest, decision trees) and their specific models or algorithms (e.g., ID3, C4.5) were extracted as reported in the included SRs and tabulated by model family (e.g., tree-based models, kernel-based models). When possible, the prevalence of each method across SRs was reported. LA model performance was also coded by extracting the specific performance metric types (e.g., accuracy, sensitivity, specificity) and the underlying modeling approach or task (e.g., classification-based), and subsequently tabulated according to families of performance metrics (e.g., threshold-dependent classification metrics, ranking-based metrics). When feasible, the prevalence of metric use across SRs was reported. In addition, when SRs provided sufficient quantitative information, performance results were summarized descriptively; otherwise, findings were synthesized narratively.

Overall, quantitative results were consistently reported in both percentages (%) and absolute counts (*n*) in order to support interpretability and ensure clarity in representing the distribution of characteristics and findings across the included SRs. Finally, findings were framed as patterns in the existing review literature and interpreted in light of the methodological quality ratings, particularly regarding the strength and robustness of the synthesized evidence base.

Results

Search results and study characteristics

The initial search of electronic databases identified a total of 1441 potentially relevant records. After the removal of duplicates, 1204 studies remained. Of these, 1124 were excluded for not meeting the inclusion criteria based on their titles and abstracts. The remaining 80 studies were retrieved for full-text assessment. Of these, 3 could not be accessed through the databases but were requested and subsequently provided by the original authors. Consequently, the full-text of 80 studies was examined, resulting in the inclusion of 16 studies that met all the established inclusion criteria. An additional 3 studies were identified in gray literature sources and incorporated into the final synthesis. Finally, a total of 19 SRs were included in this meta-review (see Fig. 1 for the PRISMA flow diagram, and Supplementary Tables S2 and S3 to consult all the references excluded and included after the full-text review).

As shown in Table 1, all SRs included in this meta-review were published between 2019 and 2024, with primary studies covering the period from 2000 to 2024. Nevertheless, there was an increase in empirical research output from 2010 onward (see Supplementary Dataset). Most of the SRs were journal articles (n = 15, 79%), and half were available via open access (n = 10, 53%). Approximately two thirds of the SRs (n = 12, 63%) were produced by researchers affiliated with engineering or computer science disciplines. Nearly all the SRs (n = 18, 95%) were standard SRs, with a minority employing comprehensive review approaches or meta-analytic techniques (n = 1, 5%). Regarding adherence to reporting guidelines, approximately 37% (n = 7) of the SRs did not report any framework. Among those that did, the PRISMA guidelines were the most frequently reported (n = 6, 32%), followed by various versions of Kitchenham's guidelines (n = 4, 21%). All SRs (n = 19, 100%) were classified as having critically low methodological quality (see Table 1 and Supplementary Table S6 for details). Additional information on the studies analyzed is provided in the Supplementary Dataset.

Educational contexts and pedagogical aims

While all the SRs used the context of higher education as an inclusion criterion regarding their primary studies, only a subset (n = 7, 37%) further specified the type of institution (e.g., universities) (n = 3, 16%), the educational level (e.g., undergraduate, postgraduate) (n = 4, 21%), or the disciplinary focus (e.g., engineering, education) (n = 1, 5%). Regard-

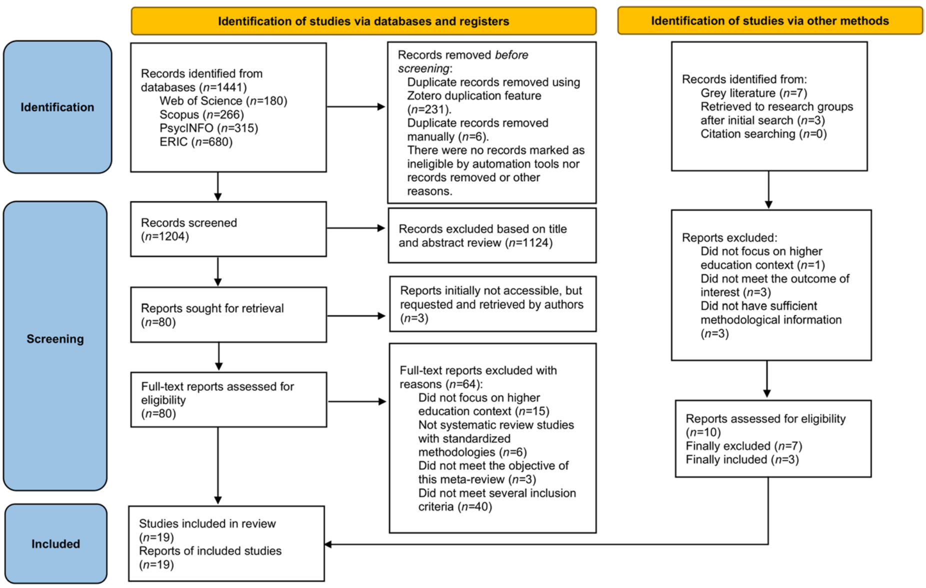

Fig. 1 PRISMA flow diagram

ing academic disciplines, three quarters of the SRs ($n = 14$, 74%) did not establish disciplinary domains as inclusion criteria, nor did they consider them as a focus of analysis of primary studies. In the SRs that explicitly examined disciplines ($n = 3$, 16%), primary studies were predominantly focused on students enrolled in engineering and technology fields, mainly engineering, computer science, and other related science, technology, engineering and mathematics (STEM) disciplines. Business-oriented and health-related programs were also represented, though to a lesser extent. These SRs also noted that many primary studies failed to specify the academic programs involved, and revealed a clear underrepresentation of social sciences, arts, and humanities, indicating a disciplinary imbalance in LA applications. Educational settings were likewise not predefined as an outcome of interest in the majority of the SRs analyzed ($n = 14$, 74%). The subset of SRs that considered this dimension ($n = 4$, 21%) indicated that applications of LA in relation to academic performance were predominantly conducted within digital learning environments, with online and blended learning modalities being the most prevalent. Moreover, none of the SRs included in this meta-review reported a synthesis of the pedagogical models or educational theories underlying the implementation of LA applications in selected primary studies.

Based on findings from fewer than half of the SRs ($n = 8$, 42%), the pedagogical purpose of the application of LA predominantly focused on improving academic performance and supporting student success. The most frequently reported objective was the assessment and improvement of student academic outcomes through predictive models, particularly those aimed at identifying students at risk of failure or dropout early in their academic trajectory. Nevertheless, the reviewed SRs also indicated that LA is used to address multiple educational objectives simultaneously, including enhancing student engagement, enabling personalized learning pathways, supporting feedback processes, and informing data-driven decision-making. Less commonly reported aims encompassed areas such as student long-term indicators of success, well-being, and satisfaction.

Although only one fifth of the SRs ($n = 4$, 21%) reported the academic or learning phase when applying LA, the available findings suggested that LA was frequently employed at the end of academic semesters, and its application, according to multiple sources, spanned the entire student life cycle from enrollment or the first-year through to graduation.

Data generation, collection and management

A third of the SRs ($n = 6$, 32%) gathered information on the systems or platforms employed for data collection in LA applications, to varying degrees of detail. Digital learning environments were consistently identified as the principal settings for data collection, particularly those enabling the tracking of student interactions. In particular, LMS emerged as the most commonly reported data source, either as a sole platform or in conjunction with other student information systems. In addition to institutional systems, other sources of data such as public datasets (e.g., the OULAD dataset) or multi-modal data strategies were occasionally employed. Furthermore, some evidence suggested that LA applications drew upon data generated by student activity ($n = 2$, 11%). However, this was an underexplored feature in the SRs included in this meta-review. Only one SR ($n = 1$, 5%) addressed the tools or platforms used to convey insights from LA to stakeholders, indicating that these are primarily implemented via customizable web-based dashboards, which display data adapted to the learning process, either in (near) real time, or on a summative basis.

Academic performance metrics

Tables 2, 3 and 4 present the findings extracted from the reviewed SRs on how academic performance has been conceptualized or operationalized in the primary studies. To create this synthesis, the metrics relating to academic performance were classified into three main categories: (1) focused on academic performance and learning outcomes (n = 14, 74% of SRs, see Table 2); (2) focused on progression and retention (n = 9, 47% of SRs, see Table 3); and (3) focused on engagement and learning behavior (n = 6, 32% of SRs, see Table 4). The synthesis of metrics across the SRs shows considerable variation, ranging from the reporting of broad constructs such as ‘learning outcomes’ and ‘student final success’ in the first category, ‘risk status’ or ‘withdrawal’ in the second category, and ‘learning behavior’ in the third category, to more precisely defined and operationalized indicators such as ‘course grades’ and ‘completion time’ in the first category, ‘dropout’, ‘retention’ and ‘course completion’ in the second category, and ‘student engagement’ in the third category. Moreover, the frequencies at which these indicators were employed in primary studies are similarly heterogeneous.

Factors studied in relation to academic performance

Most of the SRs analyzed reported and synthesized factors related to academic performance, without clarifying their significance or predictive power. Supplementary Table S7 shows the factors and indicators examined in SRs addressing predictive LA modeling of academic performance in higher education. According to the SRs examined, demographic and socioeconomic factors were the most studied in primary studies (n = 16, 84% of SRs reported it), followed by indicators related to academic background and prior performance (n = 15, 79% of SRs), psychological and personal traits (n = 12, 63% of SRs), digital interaction and log data (n = 10, 53% of SRs), behavioral engagement (n = 10, 53% of SRs), and institutional course and contextual variables (n = 5, 26% of SRs). The reporting of studied factors across the SRs was highly variable, once again, ranging from general or ambiguous predictors (reported verbatim, e.g., demographics, behavioral data, psychosocial factors) to more clearly defined and operational indicators (reported verbatim, e.g., gender, academic records, participation in class). In addition, most of the factors synthesized were employed to predict the three aforementioned categories of academic performance. However, these factors should be interpreted as frequently examined variables in the review literature, not as established predictors of academic performance.

With respect to the statistical significance or predictive contribution of factors examined in relation to academic performance, only three of the included SRs explicitly addressed this issue, although with notable limitations. Abu Saa et al. (2019) report ‘factors found significant’, and Cardona et al. (2023) refer to ‘factors with high impact’; however, in both cases it is unclear whether these terms denote statistical significance or a more general

Table 2 Conceptualizations of academic performance or learning outcomes ($n = 14$, 73.7% of included SRs)

Meta-category	Reported by n/N SRs	SRs	Indicator metric	Metric type	Operational definitions as reported in SRs	Primary studies using the indicators
Ambiguous conceptualizations on academic performance	8/19	Alhakbani & Alnassar (2022)	NR	NR	Academic performance or success without operational definition.	NR
		Blumenstein (2020)	NR	NR	Learning outcomes as grades, assessments, and pass rates.	NR
		Ifenthaler and Yau (2020)	NR	NR	'Student performance' reported without operational definition.	9/46 (19.6%)
		Ifenthaler and Yau (2020)	NR	NR	'Low performance' or 'under-achieving students'.	4/26 (15.4%)
		Jin et al. (2024)	NR	NR	'Student performance' reported without operational definition.	NR
		Mella-Norambuena et al. (2022)	NR	NR	'Academic performance' reported without operational definition.	12/12 (100%)
		Pelima et al. (2024)	NR	NR	'Academic performance' reported without operational definition.	23/70 (32.9%)
		Rahul and Katarya (2023)	NR	NR	'Student performance' reported without operational definition.	NR
		Rahul and Katarya (2023)	NR	NR	'Student's knowledge' reported without operational definition.	NR
		Sghir et al. (2023)	NR	NR	'Student performance' reported without operational definition.	57/74 (77%)
Regarding student overall qualifications	6/19	Arizmendi et al. (2023)	Above/below threshold	Nominal	Performance classified relative to a course-specific threshold.	NR
		Cardona et al. (2023)	GPA	Scale	GPA used as a performance-related variable.	NR
		Chan et al. (2019)	Final course grade	Scale	Exams and coursework combined without operational definition.	14/19 (73.7%)
		Dhankhar et al. (2021)	A-E, pass/fail, division	Nominal/Ordinal	Indicator metrics as examples of 'course grade range'.	NR

Table 2 (continued)

Meta-category	Reported by *n*/*N* SRs	SRs	Indicator metric	Metric type	Operational definitions as reported in SRs	Primary studies using the indicators
		Dhankhar et al. (2021)	Grades, marks, scores	Scale	Course 'grades', 'marks', and 'scores'.	NR
		Dhankhar et al. (2021)	GPA/CGPA	Scale	GPA and CGPA reported together.	NR
		Fahd et al. (2022)	GPA	Scale	Final GPA or course results reported without further detail.	NR
		Foster and Francis (2020)	GPA, learning gains	Scale	'Performance' as class-average improvement/individual grade gains.	10/34 (29.4%)
Regarding tasks and exams performance	3/19	Alhakbani & Alnassar (2022)	Final exam success	NR	Exam 'success' used as a benchmark label.	NR
		Ifenthaler and Yau (2020)	Correctness of answers	Nominal	Correctness of answers in quizzes, tests, or tasks.	1/46 (2.2%)
		Dhankhar et al. (2021)	Time to complete tasks	Scale	Task completion time given as an example for 'assignment' metrics.	NR

Note 1. *SRs* = Systematic Reviews, *NR* = Not reported, *GPA* = Grade Point Average, *CGPA* = Cumulative Grade Point Average

Note 2. Proportions reported in the column 'Primary studies using the indicators' reproduce the information as presented in the original SRs. SRs that provide summaries of predictive models but not primary studies are not considered primary study counts

Table 3 Conceptualizations of progression and retention ($n = 9$, 47.4% of included SRs)

Meta-category	Reported by n/N SRs	SRs	Indicator metric	Metric type	Operational definitions as reported in SRs	Primary studies using the indicators
Regarding at-risk status	7/19	Alhakbani & Alnassar (2022)	NR	NR	'At-risk' label without operational definition.	NR
		Cardona et al. (2023)	NR	NR	Risk of 'discontinuation' without operational definition.	NR
		Dhankhar et al. (2021)	NR	NR	'At-risk of dropout/attrition' without operational definition.	NR
		Ifenthaler and Yau (2020)	NR	NR	'Loss of academic status' without operational definition.	2/46 (4.3%)
		Jin et al. (2024)	NR	NR	'At-risk status' without operational definition.	NR
		Pelima et al. (2024)	NR	NR	'At-risk students' without operational definition.	16/70 (22.9%)
		Sghir et al. (2023)	NR	NR	'At risk students' without operational definition.	20/74 (27%)
Regarding retention and completion	4/19	Dhankhar et al. (2021)	Graduate and retention	Nominal	'Graduate/retention' without operational definition.	NR
		Foster and Francis (2020)	NR	NR	'Retention' without operational definition.	2/34 (5.9%)
		Ifenthaler and Yau (2020)	Course completion	NR	'Course completion' without operational definition.	7/46 (15.2%)
		Ifenthaler and Yau (2020)	Student retention	NR	'Student retention' without operational definition.	1/46 (2.2%)
		Pelima et al. (2024)	NR	NR	'Dropout and graduation' without operational definition.	14/70 (20%)
Regarding dropout and withdrawal	3/19	Alhakbani & Alnassar (2022)	Early dropout	NR	Early dropout benchmark without specification.	NR
		Ifenthaler and Yau (2020)	Attrition	NR	'Attrition' without operational definition.	1/46 (2.2%)
		Rahul and Katarya (2023)	NR	NR	'Student drop-out' without operational definition.	NR

Note 1. *SRs* = Systematic Reviews, *NR* = Not reported

Note 2. Proportions reported in the column 'Primary studies using the indicators' reproduce the information as presented in the original SRs. SRs that provide summaries of predictive models but not primary studies are not considered primary study counts

Table 4 Conceptualizations of engagement and learning behavior ($n = 6$, 31.6% of included SRs)

Meta-category	Reported by *n*/*N* SRs	SRs	Indicator metric	Metric type	Operational definitions as reported in SRs	Primary studies using the indicators
Ambiguous conceptualizations on student engagement	5/19	Foster and Francis (2020)	NR	NR	'Engagement' without operational definition.	9/24 (37.5%)
		Jin et al. (2024)	NR	NR	'Student engagement' as a high-level thematic outcome.	NR
		Johar et al. (2023)	NR	NR	'Cognitive engagement' without operational definition.	11/42 (26.2%)
		Johar et al. (2023)	NR	NR	'Emotional engagement' without operational definition.	11/42 (26.2%)
		Johar et al. (2023)	NR	NR	'Behavioral engagement' without operational definition.	9/42 (21.4%)
		Johar et al. (2023)	NR	NR	'Social engagement' without operational definition.	4/42 (9.5%)
		Johar et al. (2023)	NR	NR	'Collaborative engagement' without operational definition.	4/42 (9.5%)
		Pelima et al. (2024)	NR	NR	'Learning engagement' without operational definition.	17/70 (24.3%)
		Sghir et al. (2023)	NR	NR	'Student's engagement' without operational definition.	5/74 (6.8%)
Regarding learning behavior	1/19	Ifenthaler and Yau (2020)	NR	NR	'Student online behavior' reported without operational definition.	1/46 (2.2%)

Note 1. *SRs* = Systematic Reviews, *NR* = Not reported

Note 2. Proportions reported in the column "Primary studies using the indicators" reproduce the information as presented in the original SRs. SRs that provide summaries of predictive models but not primary studies are not considered primary study counts notion of predictive relevance or effect size. Chan et al. (2019) explicitly refer to statistical significance, noting that some primary studies did not yield significant results, yet they do not specify which predictors failed to reach statistical significance.

Learning Analytics methods, model performance, and software

LA methods were examined in 84% ($n=16$) of the included SRs. However, only 37% ($n=7$) provided summary statistics or tabular summaries of the methods used in primary studies. Across these reviews, a wide range of analytical approaches were reported, reflecting substantial heterogeneity in both modeling strategies and reporting practices. In particular, supervised predictive modeling approaches were most frequently reported, including a variety of classification- and regression-based models (see Table 5). Unsupervised methods (e.g., K-Means) were reported to a lesser extent. The synthesis of specific models or algorithms was limited across most LA methods, with linear and generalized linear regression models constituting an exception, as these were more frequently reported at a specific model level. Overall, there was substantial variation across SRs in the reported use of these methods among primary studies, precluding direct comparisons.

SRs analyzed the use of various classification approaches in primary studies, highlighting considerable variability in their application, with Naïve Bayes classifiers being the most frequently reported technique. Model performance was examined in 53% ($n=10$) of the included SRs, with only 37% ($n=7$) reporting descriptive or tabulated summaries of results. Therefore, the synthesis of LA methods and model performance should be read as a description of reporting patterns across SRs, rather than as evidence that any specific analytical approach performs better than others.

As shown in Supplementary Table S7, the SRs synthesized a heterogeneous set of performance metrics used in primary studies, spanning multiple metric families. Threshold-dependent classification performance metrics were the most frequently reported. Specifically, accuracy was the predominant metric; however, its reporting varied considerably ($n=6$; 86% of those examined model performance). Most SRs did not specify accuracy results by model type or underlying analytical task (e.g., classification, prediction) ($n=4$; 67% of those synthesized accuracy results). In addition, several reviews reported accuracy as the percentage of primary studies employing the metric ($n=3$; 50%), whereas others provided performance ranges or summary statistics ($n=3$; 50% of those synthesized accuracy results). The heterogeneity of accuracy values reported across SRs precluded cross-review comparability. Other threshold-dependent classification performance metrics were reported less frequently, including precision, sensitivity, specificity, and F-measure. Ranking-based, threshold-independent metrics (e.g., Matthews correlation coefficient, area under the receiver operating characteristic curve), as well as continuous prediction error metrics (e.g., root mean square error, mean absolute error), were reported far less commonly across SRs.

The SRs that gathered data on software used to implement LA ($n=4$, 21%) indicated that many primary studies failed to report this aspect. Nevertheless, a wide range of programs was identified, with Weka, Python, and R being the most frequently reported. Other tools, including RapidMiner, MATLAB, Orange, KNIME, SQL-based environments, SPSS and Java were mentioned less frequently.

Table 5 Summary of learning analytics methods to predict student academic performance

Analytical approach	Model family	Analytical model categories	SRs	Specific models or algorithms	Primary studies using the model in each SR
Predictive modeling approaches	Tree-based models	Random forest	Dhankhar et al. (2021)	NR	19/84 (22.6%)
			Pelima et al. (2024)	NR	27/70 (38.6%)
			Sghir et al. (2023)	NR	7/74 (9.5%)
			Tjandra et al. (2022)	NR	3/33 (9.1%)
		Decision trees	Abu Saa et al. (2019)	ID3 algorithm	4/34 (11.8%)
			Abu Saa et al. (2019)	C4.5 algorithm	4/34 (11.8%)
			Dhankhar et al. (2021)	NR	42/84 (50%)
			Sghir et al. (2023)	NR	4/74 (5.4%)
			Tjandra et al. (2022)	NR	9/33 (27.3%)
	Kernel-based models	Support Vector Machines	Abu Saa et al. (2019)	NR	8/34 (23.5%)
			Dhankhar et al. (2021)	NR	20/84 (23.8%)
			Pelima et al. (2024)	NR	31/70 (44.3%)
			Tjandra et al. (2022)	NR	4/34 (11.8%)
	Linear and generalized lineal models	Regression-based models	Mella-Norm. et al. (2022)	Linear	7/12 (58%)
			Pelima et al. (2024)	Logistic	13/70 (18.6%)
			Abu Saa et al. (2019)	Logistic	6/34 (17,6%)
			Chan et al. (2019)	Multiple linear or logistic	5/19 (26.3%)
			Dhankhar et al. (2021)	NR	32/84 (38.1%)
			Sghir et al. (2023)	Multiple linear	2/74 (2.7%)
			Sghir et al. (2023)	Single linear	2/74 (2.7%)
			Tjandra et al. (2022)	NR	1/33 (3%)
	Rule-based models	Rule-based models	Dhankhar et al. (2021)	NR	1/84 (1.2%)
			Tjandra et al. (2022)	NR	2/33 (6.1%)
	Probabilistic classifiers	Naive Bayes	Abu Saa et al. (2019)	NR	13/34 (38.2%)
			Dhankhar et al. (2021)	NR	20/84 (23.8%)
			Sghir et al. (2023)	NR	3/74 (4.1%)
			Tjandra et al. (2022)	NR	6/33 (18.2%)
	Distance-based models	K-Nearest Neighbor	Abu Saa et al. (2019)	NR	5/34 (14.7%)
			Dhankhar et al. (2021)	NR	11/84 (13.1%)
			Mella-Norm. et al. (2022)	NR	1/12 (8.3%)
			Pelima et al. (2024)	NR	9/70 (12.9%)
			Sghir et al. (2023)	NR	2/74 (2.7%)
			Tjandra et al. (2022)	NR	4/33 (12.1%)
	Neural Networks	Neural Networks	Abu Saa et al. (2019)	Multi-layer perceptron	4/34 (11.8%)
			Pelima et al. (2024)	Multi-layer perceptron	14/70 (20%)

Table 5 (continued)

Analytical approach	Model family	Analytical model categories	SRs	Specific models or algorithms	Primary studies using the model in each SR
			Abu Saa et al. (2019)	NR	4/34 (11.8%)
			Dhankhar et al. (2021)	NR	33/84 (39.3%)
			Mella-Norm. et al. (2022)	NR	2/12 (16.7%)
			Sghir et al. (2023)	NR	10/74 (13.5%)
			Tjandra et al. (2022)	NR	4/33 (12.1%)
	Partitioning-based algorithms	K-Means	Sghir et al. (2023)	NR	7/74 (9.5%)
	Ensemble methods	Ensemble methods	Dhankhar et al. (2021)	Various (e.g., stacking, boosting)	12/84 (14.3%)
			Sghir et al. (2023)	Gradient boosting	5/74 (6.8%)
			Sghir et al. (2023)	NR	3/74 (4.1%)
Statistical inferential methods	Correlational tests	Correlation tests	Chan et al. (2019)	Pearson or Spearman	3/19 (15.8%)
	Group comparison tests	T-test	Chan et al. (2019)	NR	4/19 (21.1%)
		Mann–Whitney U	Chan et al. (2019)	Not applicable	4/19 (21.1%)
		Analysis of Variance	Chan et al. (2019)	Univariate or multivariate	4/19 (21.1%)

Note 1. This table provides a descriptive synthesis of the analytical methods and algorithms reported in the included SRs. The classification is intended to characterize the analytical methods used in LA research, rather than to propose a definitive or exhaustive taxonomy

Note 2. Analytical model categories and specific models or algorithms are reported as described in the original SRs. When the underlying model or algorithm was not specified, this is indicated as not reported (NR)

Note 3. Values are reported as n/N (%), where n represents the number of primary studies within each SR that employed a given model or algorithm, and N represents the total number of primary studies included in that review

Note 4. Analytical model categories are classified under the columns 'Analytical approach' and 'Model families' according to their primary analytical purpose and their underlying mathematical and algorithmic structure, respectively

Note 5. Predictive modeling approaches were distinguished from statistical inferential methods, which were included when explicitly reported by the SRs as analytical strategies

Discussion

The aim of this meta-review was to examine SRs on the use of LA in the study of academic performance in higher education. Compared to the two previous meta-reviews in the field (Chaka, 2022; Du et al., 2019, 2021), this work presents a more comprehensive search strategy, and adopted a broader, more integrated educational perspective by synthesizing both pedagogical and operational dimensions. This was achieved through the use of a standardized methodology and adherence to a pre-registered protocol, with a view to updating and mapping the available evidence. More specifically, this meta-review offers a classification and appraisal of how synthesis studies have addressed: (1) the educational context in which LA has been used and its pedagogical purposes; (2) the conceptualization and measurement of academic performance; (3) the factors examined in relation to academic outcomes; and (4) the analytical strategies or techniques employed to process and interpret educational data.

The findings also reflect a sharp increase in the publication of SRs on LA from 2000 to 2025. This growth reflects the rapid expansion and diversification of LA as a research field (Lang et al., 2022; Viberg et al., 2018), but also underscores the need to periodically meta-synthesize these reviews in order to prevent theoretical fragmentation, map new advances, and guide future research and educational practice.

A notable characteristic of the SRs examined was the predominance of authors from engineering and computer science disciplines (63%), in contrast to only 21% from education disciplines. This disciplinary imbalance reflects LA's origins in technical fields and its reliance on computational techniques (Long et al., 2011). Yet this technical perspective often prioritizes data collection capabilities and algorithmic accuracy, while overlooking the pedagogical aims and stages of learning with which LA should be aligned, for example. The effectiveness of LA as a mechanism to personalize educational experiences and improve academic performance depends on the explicit articulation of pedagogical goals and specific moments in the teaching and learning process in which interventions are most opportune (Fariani et al., 2023). This idea is associated with classic frameworks such as Biggs' (1996) constructive alignment, which remains widely applied today in support of the achievement of intended learning outcomes. LA, without such pedagogical grounding, risks remaining a superficial reporting tool that merely informs stakeholders about student progression, without translating educational data into actionable insights for curricula design, teaching, and learning decisions.

This lack of pedagogical anchoring was evident in several of the SRs analyzed. For instance, nearly 90% of the SRs did not gather or report the educational level of students in the primary studies they reviewed. The limited reporting of educational aspects constrains the potential to adapt LA interventions to the specific developmental needs and self-regulatory skills of students at different educational stages. First-year undergraduates, for example, often require structured guidance, with the scaffolding of a set of guidance activities, to develop foundational self-regulatory skills, such as time management and goal-setting. In contrast, more advanced students may benefit from adaptive feedback mechanisms that promote autonomy, support complex problem-solving and a more precise articulation of how the acquired learning can be transferred to a relevant disciplinary or professional context. Without clearly identifying the educational stage, it is more challenging to align LA interventions with learners' specific cognitive, socio-emotional and pedagogical needs, thereby limiting their effectiveness and educational relevance.

Similarly, approximately 74% of the SRs analyzed did not gather or report the disciplinary context of the primary studies. Notably, the few SRs that did report disciplinary contexts identified a strong bias toward STEM fields, consistent with the affiliations of the SRs authors themselves, leaving social sciences, humanities and education largely unexamined. Failing to specify the academic disciplines addressed in primary studies further constrains the interpretation and transferability of findings, as learning processes, assessment practices and pedagogical approaches can differ significantly across fields such as engineering and social and health sciences. It is therefore recommendable that future synthesis studies collect, analyze and report the aforementioned aspects in relation to the applications of LA in higher education.

Another underexplored dimension in the SRs concerned the specific educational settings in which LA was applied, which were reported in only 21% of the SRs. Among these, the majority indicated that LA was predominantly employed in digital contexts, likely due to the richer data streams that these environments produce, as well as their increased prominence following the COVID-19 pandemic. However, these settings demand higher levels of learner autonomy and self-regulation. This may place certain students at a disadvantage, particularly first-year students, if the LA interventions do not adequately consider such demands. Therefore, understanding the interplay between educational settings, student academic levels and LA interventions is crucial for their effective design and implementation.

Equally striking was the complete lack of reported theoretical grounding of the primary studies across all of the SRs. This lack has been widely criticized in the literature (Gašević et al., 2017; Reimann, 2016). The limited reporting of theoretical underpinnings may also help explain why 58% of the SRs did not gather information on the pedagogical aims of primary studies. When the underlying educational purpose in the application of LA is ill-defined, LA's usefulness seems to be reduced to mere prediction and risk detection. Future SRs would benefit from explicitly identifying and reporting the educational theories underpinning primary studies, not merely as background context, but as analytical frameworks guiding the application of LA methods, the interpretation of results, and the derivation of educational implications. In educational interventions mediated by LA applications that lack an explicit theoretical foundation, there is a risk for data to be treated primarily as outcomes that reinforce a technocentric approach, rather than as tools to support learning processes, inform instructional decision-making, or foster student development on the basis of meaningful educational insights. Therefore, integrating theoretical frameworks into SRs would enhance interpretative depth, support learner-responsive LA design, and strengthen the connection between research and educational practice in higher education (Utamachant et al., 2023; Yan et al., 2024).

Regarding data collection and dissemination, most SRs highlighted LMS as the primary data source in primary studies, reflecting the convenience and ubiquity of such platforms. Yet LMS data captures only certain dimensions of learning (e.g., clicks, logins, submissions), while mostly neglecting affective and collaborative aspects, and offline learning behaviors. The promise of multimodal LA lies in its ability to integrate diverse data sources (Giannakos & Cukurova, 2023), including biometric and observational data, although this potential remains to date underexploited. Moreover, some studies pointed out the limitations of certain metrics, such as time-on-task and clickstreams, as not always being predictive of performance in specific tasks (López-Francés et al., 2022; Tellakat et al., 2019). Equally problematic is the failure of most of the SRs to document the parties generating the data (students, instructors, or automated systems), and how the data were assessed. Treating heterogeneous data sources as equivalent introduces interpretive risks and compromises the validity of performance metrics. These limitations underscore the need for future research to move beyond decontextualized LMS-based indicators by incorporating contextually relevant educational dimensions, integrating multi-source data, and adopting multivariate and multilevel analytical approaches that more adequately reflect the complexity of learning processes and its outcomes.

A similar heterogeneity was observed in how academic performance itself is defined and measured. The SRs alternately reported final grades, engagement indicators, retention rates, and subjective assessments, conflating distinct constructs under the umbrella of "performance". Clearer conceptualizations of performance, possibly distinguishing between short-term task outcomes, long-term retention, and engagement, would enhance comparability, and guide more opportune LA applications.

The heterogeneity in the conceptualization and measurement of academic performance in the primary studies reported in the SRs reflects inconsistent operationalizations of outcomes, disparate measurement methods, and an overreliance on quantitative metrics, often unsupported by robust theories of change (Foster & Francis, 2020; Ifenthaler & Yau, 2020). Such inconsistency obscures what students are actually engaging with in learning environments, and weakens the explanatory power of analyses (Chan et al., 2019; Mella-Norambuena et al., 2022). Unless these conceptual weaknesses are addressed, LA risks falling short of its potential to generate actionable insights that genuinely enhance teaching and learning.

In this regard, both academic performance outcomes and predictors have been predominantly synthesized from a technical, data-mining perspective, with limited consideration of educational theories in their conceptualizations. Some exceptions do exist however. For instance, Blumenstein's (2020) SR took into account the Future of Learning framework (Redecker et al., 2011) and conceptualizations of learning gains in higher education (Vermunt et al., 2018) in discussing its results and proposing the LA Learning Gain Design Model. Similarly, Johar et al. (2023) used the framework of Redmond et al. (2018) to explain student engagement. However, such theoretical integrations remain rare, highlighting a persistent gap that limits both the interpretability and educational relevance of LA findings.

From an educational perspective, it is important to distinguish between static predictors, such as personal traits, demographics, socioeconomic characteristics, and prior academic background, and process-oriented variables, such as student engagement, digital interaction and learning behaviors, self-regulatory behaviors, and institutional and course-related contexts. This distinction reflects different pedagogical mechanisms and levels of educational action associated with LA applications in higher education. While static variables may support the early identification of potential risk profiles, process-oriented indicators may provide more actionable insights for understanding learning processes, informing instructional decision-making, and designing timely formative interventions. In this regard, the reporting of predictive factors in the SRs was often incomplete or ambiguous. Although demographic and socioeconomic variables were frequently listed, few of the SRs specified their predictive significance. The continued emphasis on static demographic factors over dynamic, context-sensitive variables, such as real-time behavioral and affective data, limits insights into the learning process (Sghir et al., 2023). Such approaches overlook social, cognitive and emotional factors essential to explaining student learning behaviors (Ifenthaler & Yau, 2020; Mella-Norambuena et al., 2022). This raises concerns about over-reliance on easily available but potentially less educationally meaningful variables, while neglecting process-oriented variables (e.g., participation in class, time spent with materials, study habits), or contextually situated indicators (e.g., external assessments, learning environments, teaching-related variables). While static variables may be useful for identifying risk profiles, they offer limited explanatory value for understanding learning processes or informing instructional adaptation. In contrast, process-oriented variables may be more closely aligned with pedagogical theory and may provide greater potential for designing timely, formative, and learner-responsive LA interventions (Musso et al., 2020). However, these interpretations should be understood as educationally grounded hypotheses derived from patterns in the review literature. Furthermore, little attention was paid to the risk of bias or the inequity inherent in demographic predictors, omissions that warrant correction in future research. This limited reporting hinders a nuanced understanding of which factors truly influence academic performance, and may lead to simplistic conclusions. This meta-review therefore underscores the need for more critical and comprehensive analyses in future SRs.

Methodologically, the predominance for supervised, quantitative techniques (primarily regression and classification models) suggests that evidence in the reviewed literature was operationalized mainly through predictive modeling approaches. Clustering methods and mixed-method approaches remain underutilized, despite their potential to uncover latent patterns and provide rich insights into learning processes. A more balanced methodological repertoire, including qualitative and hybrid approaches, could enhance both the interpretability and relevance of findings. In addition, the results highlight substantial limitations in how both LA methods and model performance are synthesized and reported across SRs. The frequent synthesis of analytical models grouped under broad categories, with limited specification of underlying algorithms, hampers methodological comparisons and the derivation of robust conclusions. Similarly, heterogeneous reporting practices and limited contextualization of model performance by model type, analytical task, or validation strategy restrict the interpretability of reported results and preclude meaningful cross-study synthesis. Addressing these issues, particularly through more transparent and standardized reporting of analytical models, modeling tasks, and performance metrics, is essential to strengthen the evidence base of LA.

Although this meta-review offers a comprehensive synthesis of how LA has been applied in the study of academic performance in higher education, several limitations should be considered. Synthesizing evidence from previously published reviews entails inherent risks of bias, particularly given the variability, among the included SRs, in terms of scope, narrative focus, terminology, and reporting practices. This heterogeneity complicated data extraction, especially for key aspects such as academic performance conceptualizations and predictive factors used in LA models. While the approach taken enables a comprehensive overview of the field, it also limits the direct comparability of findings, which should be interpreted with caution. The low methodological quality of our analytical sample is another important aspect to highlight in our research. None of the nineteen included SRs conducted a formal risk-of-bias assessment of primary studies, or evaluated potential publication bias. Consequently, the strength of the evidence base is constrained not only by variability in reporting, but also by the absence of critical quality appraisals, limiting the interpretability and robustness of the synthesized findings. A further limitation of this meta-review is the inclusion of only English- and Spanish-language SRs, which may also have introduced selection bias. Nonetheless, the meta-review involved a broad geographical coverage and diverse institutional contexts. Additionally, the limited number of SRs reporting on certain context subgroups (e.g., educational levels, disciplines) precluded more granular comparisons. The predominance of prediction-focused objectives across most SRs, while methodologically coherent, risks narrowing the educational implications of LA by overlooking formative and developmental purposes. Notwithstanding such limitations, this meta-review makes a significant contribution to the literature and has notable strengths. It employs an updated and exhaustive search strategy, improving upon previous meta-reviews, and benefits from a minimal overlap in primary studies across the SRs analyzed. These features enhance its internal validity, broaden its coverage, and support a more nuanced and diverse view of the available evidence. Moreover, in identifying the absence of educational perspective in existing SRs, this meta-review contributes by evidencing a critical gap in the LA literature, and delineates clear directions for theoretically informed evidence synthesis in higher education.

Overall, this meta-review highlights patterns in the existing review literature on the applications of LA in the study of academic performance in higher education. The included SRs devote limited attention to essential educational dimensions, including instructional contexts, pedagogical aims, disciplinary domains, theoretical underpinnings, and stages of student learning. Considerable variation was found in the conceptualization and operationalization of academic performance. While demographic and academic background variables were among the most frequently examined factors, their predictive value was rarely reported in the SRs. Supervised quantitative methods were also frequently reported, but performance metrics varied substantially and were not sufficiently comparable to support robust conclusions about model performance. The findings of this study emphasize the need for future synthesis studies to systematically incorporate and report more educationally relevant aspects.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10734-026-01709-y.

Acknowledgements Not applicable.

Author contributions **Fran J. García-García**: Conceptualization, Validation, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Project administration, Funding acquisition. **María Isabel Gómez-Núñez**: Validation, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Cristian Molla-Esparza**: Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Supervision.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The research has been supported by the Spanish Ministry of Science, Innovation and Universities (under grant number PID2022-141403NB-I00, with MCIN/AEI/https://doi.org/10.13039/501100011033/FEDER-EU funding), and by the Government of the Autonomous Region of Valencia, Spain (under grant number CIGE/2023/53).

Data availability The author confirms that all data generated or analyzed during this study are included in this published article. Specifically, the data supporting the findings of this study are openly available under the Open Science Framework by clicking here.

Declarations

Ethics approval Not required.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abu Saa, A., Al-Emran, M., & Shaalan, K. (2019). Factors affecting students' performance in higher education: a systematic review of predictive data mining techniques. *Technology Knowledge and Learning*, *24*(3), 567–598. https://doi.org/10.1007/s10758-019-09408-7

Alnasyan, B., Basheri, M., & Alassafi, M. (2024). The power of deep learning techniques for predicting student performance in virtual learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *6*, 100231. https://doi.org/10.1016/j.caeai.2024.100231

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, *13*(3), 132–140. https://doi.org/10.1097/XEB.0000000000000055

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. A. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, *37*, 100489. https://doi.org/10.1016/j.edurev.2022.100489

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, *32*(3), 347–364. https://doi.org/10.1007/bf00138871

Blumenstein, M. (2020). Synergies of Learning Analytics and Learning Design: A Systematic Review of Student Outcomes. *Journal of Learning Analytics*, *7*(3), 13–32. https://doi.org/10.18608/jla.2020.73.3

Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2023). Data Mining and Machine Learning Retention Models in Higher Education. *Journal of College Student Retention: Research Theory & Practice*, *25*(1), 51–75. https://doi.org/10.1177/1521025120964920

Chaka, C. (2022). Educational data mining, student academic performance prediction, prediction methods, algorithms and tools: An overview of reviews. *Journal of E-Learning and Knowledge Society*, *18*(2), 58–69. https://doi.org/10.20368/1971-8829/1135578

Chan, A. K. M., Botelho, M. G., & Lam, O. L. T. (2019). Use of learning analytics data in health care–related educational disciplines: Systematic review. *Journal of Medical Internet Research*, *21*(2), e11241. https://doi.org/10.2196/11241

Dol, S. M., & Jawandhiya, P. M. (2024). Systematic Review and Analysis of EDM for Predicting the Academic Performance of Students. *Journal of The Institution of Engineers (India): Series B*, *105*(4), 1021–1071. https://doi.org/10.1007/s40031-024-00998-0

Du, X., Yang, J., Zhang, M., Hung, J. L., & Shelton, B. E. (2019). Learning Analytics Research: Using Meta-Review to Inform Meta-Synthesis. In K. Arai, R. Bhatia, & S. Kapoor (Eds.), *Proceedings of the future technologies conference (FCT) 2018* (pp. 1097–1108). Springer. https://doi.org/10.1007/978-3-030-02686-8_81

Du, X., Yang, J., Shelton, B. E., Hung, J. L., & Zhang, M. (2021). A systematic meta-review and analysis of learning analytics research. *Behaviour & Information Technology*, *40*(1), 49–62. https://doi.org/10.1080/0144929X.2019.1669712

Fariani, R. I., Junus, K., & Santoso, H. B. (2023). A Systematic Literature Review on Personalised Learning in the Higher Education Context. *Technology Knowledge and Learning*, *28*(2), 449–476. https://doi.org/10.1007/s10758-022-09628-4

Foster, C., & Francis, P. (2020). A systematic review on the deployment and effectiveness of data analytics in higher education to improve student outcomes. *Assessment & Evaluation in Higher Education*, *45*(6), 822–841. https://doi.org/10.1080/02602938.2019.1696945

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice*, *3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Gates, M., Gates, A., Pieper, D., Fernandes, R. M., Tricco, A. C., Moher, D., Brennan, S. E., Li, T., Pollock, M., Lunny, C., Sepúlveda, D., McKenzie, J. E., Scott, S. D., Robinson, K. A., Matthias, K., Bougioukas, K. I., Fusar-Poli, P., Whiting, P., Moss, S. J., & Hartling, L. (2022). Reporting guideline for overviews of reviews of healthcare interventions: Development of the PRIOR statement. *BMJ*, e070849. https://doi.org/10.1136/bmj-2022-070849

Giannakos, M., & Cukurova, M. (2023). The role of learning theory in multimodal learning analytics. *British Journal of Educational Technology*, *54*(5), 1246–1267. https://doi.org/10.1111/bjet.13320

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* (4th edition). Advanced Analytics, LLC. https://doi.org/10.1002/9781118445112.stat06882

Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, *3*, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004

Ifenthaler, D., & Yau, J. Y. K. (2017). Utilising learning analytics for study success in higher education: A systematic review. In M. Campbell, J. Willems, C. Adachi, D. Blake, I. Doherty, S. Krishnan, S. Macfarlane, L. Ngo, M. O'Donnell, S. Palmer, L. Riddell, I. Story, H. Suri, & J. Tai (Eds.), *Open oceans: Learning without borders. Proceedings ASCILITE 2018 Geelong* (pp. 406–411). ASCILITE https://doi.org/10.14742/apubs.2018.1958

Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, *68*(4), 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Johar, N., Kew, S., Tasir, Z., & Koh, E. (2023). Learning analytics on student engagement to enhance students' learning performance: a systematic review. *Sustainability*, *15*(10), 7849. https://doi.org/10.3390/su15107849

Lang, C., Siemens, G., Wise, A. F., Gašević, D., & Merceron, A. (Eds.). (2022). *The handbook of learning analytics*. SoLAR. https://doi.org/10.18608/hla22

Larrabee Sonderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, *50*(5), 2594–2618. https://doi.org/10.1111/bjet.12720

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

Llopis-Albert, C., & Rubio, F. (2021). Application of Learning Analytics to Improve Higher Education. *Multidisciplinary Journal for Education Social and Technological Sciences*, *8*(2), 1–18. https://doi.org/10.4995/muse.2021.16287

Long, P., Siemens, G., Conole, G., & Gašević, D. (2011). *LAK 2011: 1st international conference on learning analytics and knowledge*. https://doi.org/10.1145/2090116

López-Francés, I., Garcia-Garcia, F. J., Gargallo López, B., & Molla-Esparza, C. (2022). Implications of Message Length and Delay in Undergraduate Online Discussions. *RIED-Revista Iberoamericana de Educación a Distancia*, *25*(2), 259–276. https://doi.org/10.5944/ried.25.2.32810

Mella-Norambuena, J., Badilla-Quintana, M. G., & López-Angulo, Y. (2022). Predictive models based on the use of learning analytics in higher education: A systematic review. *Texto Livre-Linguagem e Tecnologia*, *15*, e36310. https://doi.org/10.35699/1983-3652.2022.36310

Mhlanga, D. (2024). Digital transformation of education, the limitations and prospects of introducing the fourth industrial revolution asynchronous online learning in emerging markets. *Discover Education*, *3*(1), 32. https://doi.org/10.1007/s44217-024-00115-9

Musso, M. F., Hernández, C. F. R., & Cascallar, E. C. (2020). Predicting key educational outcomes in academic trajectories: A machine-learning approach. *Higher Education*, *80*, 875–894. https://doi.org/10.1007/s10734-020-00520-7

Namoun, A., & Alshanqiti, A. (2020). Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. *Applied Sciences*, *11*(1) , 237. https://doi.org/10.3390/app11010237

Ouyang, F., Xu, W., & Cukurova, M. (2023). An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning*, *18*(1), 39–66. https://doi.org/10.1007/s11412-023-09387-z

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, 71. https://doi.org/10.1136/bmj.n71

Pan, Z., Biegley, L., Taylor, A., & Zheng, H. (2024). A Systematic Review of Learning Analytics. *Journal of Learning Analytics*, *11*(2), 52–72. https://doi.org/10.18608/jla.2023.8093

Pieper, D., Antoine, S. L., Mathes, T., Neugebauer, E. A., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, *67*(4), 368–375. https://doi.org/10.1016/j.jclinepi.2013.11.007

Pollock, M., Fernandes, R. M., Pieper, D., Tricco, A. C., Gates, M., Gates, A., & Hartling, L. (2019). Preferred Reporting Items for Overviews of Reviews (PRIOR): A protocol for development of a reporting guideline for overviews of reviews of healthcare interventions. *Systematic Reviews*, *8*, 335. https://doi.org/10.1186/s13643-019-1252-9

Rahul, & Katarya, R. (2023). A Systematic Review on Predicting the Performance of Students in Higher Education in Offline Mode Using Machine Learning Techniques. *Wireless Personal Communications*, *133*(3), 1643–1674. https://doi.org/10.1007/s11277-023-10838-x

Redecker, C., Leis, M., Leendertse, M., Punie, Y., Gijsbers, G., Kirschner, P., Stoyanov, S., & Hoogveld, B. (2011). *The future of learning: Preparing for change*. Publications Office of the European Union. https://doi.org/10.2791/64117

Redmond, P., Heffernan, A., Abawi, L., Brown, A., & Henderson, R. (2018). An Online Engagement Framework for Higher Education. *Online Learning*, *22*(1), 183–204. https://doi.org/10.24059/olj.v22i1.1175

Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice*, *2*(2), 130–142. https://doi.org/10.1080/23735082.2016.1210198

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., Group, P. R. I. S. M. A. S., Blunt, H., Brigham, T., Chang, S., Clark, J., Conway, A., Couban, R., De Kock, S., Farrah, K., Fehrmann, P., Foster, M., Fowler, S. A., & Young, S. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, *10*(1), 39. https://doi.org/10.1186/s13643-020-01542-z

Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent Advances in Predictive Learning Analytics: A Decade Systematic Review (2012–2022). *Education and Information Technologies*, *28*(7), 8299–8333. https://doi.org/10.1007/s10639-022-11536-0

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, j4008 https://doi.org/10.1136/bmj.j4008

Tellakat, M., Boyd, R. L., & Pennebaker, J. W. (2019). How do online learners study? The psychometrics of students' clicking patterns in online courses. *PLOS ONE*, *14*(3), e0213863. https://doi.org/10.1371/journal.pone.0213863

Utamachant, P., Anutariya, C., & Pongnumkul, S. (2023). i-Ntervene: Applying an evidence-based learning analytics intervention to support computer programming instruction. *Smart Learning Environments*, *10*, 37. https://doi.org/10.1186/s40561-023-00257-7

Vermunt, J. D., Ilie, S., & Vignoles, A. (2018). Building the foundations for measuring learning gain in higher education: A conceptual framework and measurement instrument. *Higher Education Pedagogies*, *3*(1), 266–301. https://doi.org/10.1080/23752696.2018.1484672

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, *89*, 98–110. https://doi.org/10.1016/j.chb.2018.07.027

Yan, L., Echeverria, V., Jin, Y., Fernandez-Nieto, G., Zhao, L., Li, X., Alfredo, R., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Evidence-based multimodal learning analytics for feedback and reflection in collaborative learning. *British Journal of Educational Technology*, *55*(5), 1900–1925. https://doi.org/10.1111/bjet.13498

Molla-Esparza, C., Gómez-Núñez, M. I., & García-García, F. J. (2025). Applications of learning analytics in the study of academic performance in higher education: A pilot-tested meta-review protocol. *International Journal of Educational Research Open*, *8*, 100433. https://doi.org/10.1016/j.ijedro.2024.100433

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

Authors and Affiliations

Fran J. García-García[1] · **María Isabel Gómez-Núñez**[2] · **Cristian Molla-Esparza**[3]

✉ Cristian Molla-Esparza
Cristian.Molla@uv.es

Fran J. García-García
garfran9@uv.es

María Isabel Gómez-Núñez
mariaisabel.gomez@unir.net

[1] Department of Educational Theory, Faculty of Philosophy and Education Sciences, University of Valencia, Av. de Blasco Ibáñez, 30, El Pla del Real, Valencia 46010, Spain

[2] Department of Research Methods and Diagnostics in Education, Faculty of Education Sciences and Humanities, International University of La Rioja, Avenida de la Paz, 137, Logroño 26006, La Rioja, Spain

[3] Department of Research Methods and Diagnostics in Education, Faculty of Philosophy and Education Sciences, University of Valencia, Avenida de Blasco Ibáñez, 30, El Pla del Real, Valencia 46010, Spain