Computers and Education: Artificial Intelligence 2 (2021) 100027

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Automatic feedback in online learning environments: A systematic literature review

Anderson Pinheiro Cavalcanti [a,*], Arthur Barbosa [b], Ruan Carvalho [b], Fred Freitas [a], Yi-Shan Tsai [c], Dragan Gašević [c,d,e], Rafael Ferreira Mello [b]

[a] Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil
[b] Departamento de Computação, Universidade Federal Rural de Pernambuco, Recife, Brazil
[c] Centre for Learning Analytics, Faculty of Information Technology, Monash University, Australia
[d] School of Informatics, University of Edinburgh, United Kingdom
[e] Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia

ARTICLE INFO

Keywords:
Educational feedback
Automatic feedback
Systematic review
Online learning environments

ABSTRACT

Feedback is an essential component of scaffolding for learning. Feedback provides insights into the assistance of learners in terms of achieving learning goals and improving self-regulated skills. In online courses, feedback becomes even more critical since instructors and students are separated geographically and physically. In this context, feedback allows the instructor to customize learning content according to the students' needs. However, giving feedback is a challenging task for instructors, especially in contexts of large cohorts. As a result, several automatic feedback systems have been proposed to reduce the workload on the part of the instructor. Although these systems have started gaining research attention, there have been limited studies that systematically analyze the progress achieved so far as reported in the literature. Thus, this article presents a systematic literature review on automatic feedback generation in learning management systems. The main findings of this review are: (1) 65.07% of the studies demonstrate that automatic feedback increases student performance in activities; (2) 46.03% of the studies demonstrated that there is no evidence that automatic feedback eases instructors' workload; (3) 82.53% of the studies showed that there is no evidence that manual feedback is more efficient than automatic feedback; and (4) the main method used for automatic feedback provision is the comparison with a desired answer in some subject (such as logic circuits or programming).

1. Introduction

Online learning has grown tremendously in recent years as an alternative or complementary option to traditional education which is primarily based on face-to-face teaching. One of the factors of this growth is that students are active users of technology and that, in general, they use technologies more than they believe they are obliged to (Jones et al., 2010). According to Sung and Mayer (2012), online learning has grown because it is more flexible than traditional educational environments. For the purpose of facilitating online learning, various platforms such as learning management systems (LMSs) have emerged in the past decades (Ouadoud et al., 2017). The use of LMSs has increased in recent years due to the use of information and communication technologies as an educational support tool (Reiser & Dempsey, 2012). These environments have several resources (e.g., chat, forum, and wiki) that allow numerous interactions between instructors, students, and content (Joksimović et al., 2015). Despite the advantages of online learning, there are some challenges for instructors. Among these, it is particularly notable that instructors struggle to follow the progress and activities of a cohort that is potentially unlimited in size (Hernández-García et al., 2015).

Feedback is an essential component in the teaching-learning process as it allows students to identify gaps and assess their learning progress (Butler & Winne, 1995). According to Sadler (1989), feedback needs to provide specific information related to a learning task or process that fills a gap between the desired and the real understanding of the content or the development of abilities. Through feedback, students seek to hone some inadequate or poor knowledge or skills that hinder their learning

* Corresponding author.
E-mail addresses: apc@cin.ufpe.br, cavalcanti.pinheiro@gmail.com (A.P. Cavalcanti), arthudiego@gmail.com (A. Barbosa), ruan.carvalho@ufrpe.br (R. Carvalho), fred@cin.ufpe.br (F. Freitas), yi-shan.tsai@monash.edu (Y.-S. Tsai), dragan.gasevic@monash.edu (D. Gašević), rafael.mello@ufrpe.br (R.F. Mello).

https://doi.org/10.1016/j.caeai.2021.100027
Received 1 April 2021; Received in revised form 30 June 2021; Accepted 5 July 2021
Available online 17 July 2021

progress. Several studies have shown that useful feedback brings benefits to learning (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Parikh et al., 2001). For instance, Black and Wiliam (1998) analyzed more than 250 feedback studies and concluded that feedback produced significant gains in student learning and satisfaction. Recently, the study of Henderson et al. (2019) analyzed seven case studies, through multiple stages of thematic analysis, case comparison, and reliability verification, and proposed 12 main conditions that support effective feedback. These conditions highlight the importance of carefully designing feedback processes and have been organized into three categories: capacity, projects, and culture.

In online learning contexts, feedback plays a crucial role due to the lack of face-to-face interaction among the participants of the course (Ypsilandis, 2002). As instructors and students are separated in space and/or time in online contexts, the instructor must provide high-quality feedback to assist students in their learning and motivation (Nicol & Macfarlane-Dick, 2006). Tseng and Tsai (2007) found that reinforcing feedback is useful to promote the quality of the student's project, especially in online peer assessment environment. However, the large size of student cohort in online learning environments can make it challenging for the instructor to provide useful and sufficient feedback to students. In light of this, several automatic tools have been proposed to enhance feedback practice (Belcadhi, 2016; Gulwani et al., 2014; Marin et al., 2017).

There has been a lacuna in studies that systematically analyze automatic feedback systems in online environments. One exception is a technical report on studies about automatic feedback generation for programming exercises (Keuning et al., 2018). One key finding of this study is that existing tools often do not give feedback on how to solve problems and take the next steps. This has also made it difficult for instructors to quickly adapt tools and resources to their own needs. The difference of the study presented in Keuning et al. (2018) from the systematic literature review presented in this paper is that we do not limit automatic feedback to programming exercise tools only. Instead, we include all the automatic feedback generation systems in online learning environments. Moreover, previous literature reviews in the field of Educational Technology showed the importance of analyzing feedback systems (Chen, Zou, Cheng, & Xie, 2020,b).

In this context, this paper presents a *systematic literature review focusing on tools and resources that enable automatic feedback in learning management systems*. It allows the identification, evaluation, and interpretation of all available research relevant to a research question, subject, or event of interest (Kitchenham, 2004). Moreover, a literature review should conduct a critical evaluation of research studies that address a specific issue and must have a well-defined structure so that the results are not biased. Finally, the rigor of a systematic literature review needs to be strengthened by reducing random effects and ensuring reproductivity (Becheikh et al., 2006).

The current systematic literature study followed the guidelines and model of systematic review protocol proposed by Keele et al. (2007), which included three main steps:

1. The *planning step* identified the goals of the systematic literature review and defined the review protocol;
2. The *execution step*, which was the main stage of the review, and included the following activities: (i) formulated focused research questions, (ii) searched for and selected the primary studies, (iii) defined the papers needed to answer the research questions, and (iv) extracted the data and synthesized the results.
3. The *reporting step* presented the summarized results with interpretation and discussion.

In summary, we selected 63 studies based on relevant keywords related to feedback and online learning environments, published between 2009 and 2018. In order to present a concise analysis, we have extracted 19 features from papers selected. These features were grouped into categories such as basic information (e.g., year and title), goals, results, and methods to provide feedback of the papers selected. The results and their implications are further discussed in this paper.

2. Method

2.1. Research questions

Automatic feedback emerges as a solution to an instructor's heavy workload due to the need to support a large number of students enrolled in online courses. However, it is necessary to analyze whether the studies that propose an automatic feedback approach help the instructor and/or the student. To do so, we defined the overarching research question:

RESEARCH QUESTION: What are the approaches used for generating automatic feedback in online learning environments?

Based on this overarching research question, we divided our work into four sub-questions:

RESEARCH QUESTION 1 (RQ1): Does automated feedback in online learning environments improve student performance in activities?

This question aims to identify whether the papers selected support the expectation that automated feedback approaches improve student performance in activities compared to conditions without automatic feedback tools.

RESEARCH QUESTION 2 (RQ2): *What are the main goals in using automatic feedback generation techniques in online learning environments?*

In particular, this research question aimed to explore whether the objective of the studies was to: (i) help students with specific content, (ii) support students to improve self-regulation, and (iii) assist instructors in the creation of feedback.

RESEARCH QUESTION 3 (RQ3): *Is there any evidence suggesting that automatic feedback helps instructors?*

This question examines whether approaches proposed by existing studies provide evidence that the use of automatic feedback tools/resources enhances the capability of instructors to develop better feedback.

RESEARCH QUESTION 4 (RQ4): *What techniques are used to generate automatic feedback?*

This question aimed to investigate which techniques and resources had been used to generate automatic feedback. The techniques could be, for example, machine learning, Natural Language Processing (NLP), and ontologies.

2.2. StArt tool

All the steps of the systematic review were performed using a systematic review management tool called StArt (State of the Art through Systematic Review).[1] StArt assists the researcher in the development of a systematic literature review (LAPES, 2014), i.e., the steps presented previously (planning, execution, and reporting steps). The StArt tool was evaluated empirically and it was demonstrated that this tool had positive results in the execution of systematic literature reviews (Hernandes et al., 2012; Tenório et al., 2016).

2.3. Search strategy

According to Kitchenham (2004), in a systematic literature review, it is necessary to determine and follow a search strategy. The first stage is to define the keywords and their possible combinations. In this step, we followed the same approach used by Tenório et al. (2016). The following keywords (and their synonyms) were used:

- feedback;

[1] http://lapes.dc.ufscar.br/tools/start_tool.

- online learning environment (virtual learning environments, massive open online courses, MOOC, intelligent tutoring system, e-learning, online courses, distance education, educational environment, learning management system);
- student (learner);
- instructor (tutor, teacher).

After defining the keywords and their synonyms, we built a search string using the logical operators (OR) and (AND). The operators (OR) and (AND) were used between the synonyms and keywords, respectively. Therefore, the following search strings were generated:

1. "feedback";
2. "online learning environment" OR "virtual learning environments" OR "educational environment";
3. "massive open online courses" OR "MOOC";
4. "intelligent tutoring system";
5. "e-learning" OR "online courses" OR "distance education" OR "educational environment" OR "learning management system";
6. "student" OR "learner";
7. "teacher" OR "tutor" OR "instructor".

The ultimate combination of the search string used was:
((1) AND (2 OR 3 OR 4 OR 5) AND (6 OR 7))

We employed the proposed search string in the following databases that are prominent in publishing research in the field of educational technology (Tenório et al., 2016):

- ACM – (https://dl.acm.org/)
- IEEE xplorer – (https://ieeexplore.ieee.org/)
- Engineering Village (https://www.engineeringvillage.com/)
- Science Direct – (https://www.sciencedirect.com)
- Scopus – (http://www.scopus.com)
- SpringerLink – (https://link.springer.com/)

2.4. Selection criteria

In this step, the studies have to meet the selection criteria (inclusion and exclusion) to be included in the systematic review (Keele et al., 2007). The inclusion criteria are primary studies that propose an automatic feedback approach in online learning environments and published from January 2009 to December 2018. On the other hand, the exclusion criteria are secondary and tertiary studies, short papers, duplicated studies, non-English written papers, grey literature (e.g., books, theses, dissertations, and so on), and incomplete studies. Table 1 summarizes the step-by-step of our selection criteria.

2.5. Selection process

In step one, the reviewers only read the title and abstract and decided to include or exclude the study based on inclusion and exclusion criteria.

Table 1
Selection criteria.

Number	Type	Description
1	Inclusion	Primary study
2	Inclusion	Study that proposes an automatic feedback approach in Online Learning Environments
3	Inclusion	Study published from January 2009 to December 2018 (10 years)
4	Exclusion	Secondary and tertiary studies
5	Exclusion	Short papers (<5 pages)
6	Exclusion	Duplicated studies
7	Exclusion	Non-English written papers
8	Exclusion	Grey literature
9	Exclusion	Incomplete Studies

If the reviewers did not have enough information to exclude, the study went to the next step, where the reviewers read the introduction and final considerations in order to define the relevance of the paper to the review.

2.6. Extraction process

In step three, the reviewers read the full text of the articles to extract data relevant to answering our research questions. Table 2 shows all the fields that were extracted from the articles.

3. Execution of the systematic literature review

This section describes the execution of the systematic review of the literature. The first step was to use the search string in the digital libraries and download all returned articles in the .bibTex format. This step was performed manually for each digital library. Table 3 shows the number of articles obtained in each of the digital libraries.

The next step was to import the files of each digital library into the StArt tool. This step was divided into three phases: (1) Automatic removal of duplicate articles using the StArt tool; (2) The reviewers read the title and abstract of the article and applied the inclusion and exclusion criteria; (3) The reviewers read the introduction and conclusion sections of the article and applied the inclusion and exclusion criteria. Fig. 1 shows the number of articles selected in each phase.

In Phase 1, duplicate articles were automatically removed using the StArt tool. The tool can detect the same articles comparing texts between the articles. In this phase, 1079 articles were removed. In Phase 2, the reviewers excluded 1964 articles that did not satisfy the inclusion criteria. About 92% of articles were excluded because they were out of scope, 3% grey literature (e.g., books, theses, dissertations and so on), 3% short papers, 1% secondary and tertiary studies, 0.9% duplicated studies, and 0.1% incomplete studies. It is important to note that an article may have been removed by more than one exclusion criteria.

Some information, such as the number of pages of an article or keywords, sometimes did not appear in the StArt tool. Therefore, the reviewers did not have enough information to determine whether the article would be included or excluded in phase 2. As a result, researchers reviewed these articles manually in phase 3. In this phase, some articles such as short papers and articles not written in English were discovered

Table 2
Data extraction form fields.

#	Field	Description
1	ID	Unique identifier for the study
2	Title	Title of the paper.
3	Authors	Authors of the paper.
4	Year	Year in which the paper was published.
5	Country	Country of the first author of the paper.
6	Type	Conference, journal, and workshop.
7	Educational tool	Does it propose a new tool?
8	Is the tool available?	If yes, what is the URL?
9	Database	If the study uses or proposed a corpus for analyzing a feedback system.
10	Tools	Tools used in the study.
11	Type of evaluation	Experiment, case study, application in the real environment, and questionnaires, among others.
12	Subject area	Subject area of the course in which the system was applied.
13	Main results	What are the main results of the paper?
14	Educational level	Higher education, secondary education, primary education, N/A (i.e., no enough data to conclude).
15	Impact on student performance (RQ1)	Evidence of positive or negative impact
16	Main goal (RQ2)	What are the main goals of the paper?
17	Impact on teaching? (RQ3)	Evidence of positive or negative impact
18	Methods (RQ4)	What techniques were used to generate automatic feedback?

Table 3
Number of articles returned for the search string in each digital library.

Digital Library	Number of articles
ACM	354
IEEExplorer	361
Engineering Village	25
Science Direct	667
Scopus	1371
Springer Link	400
Total	3178

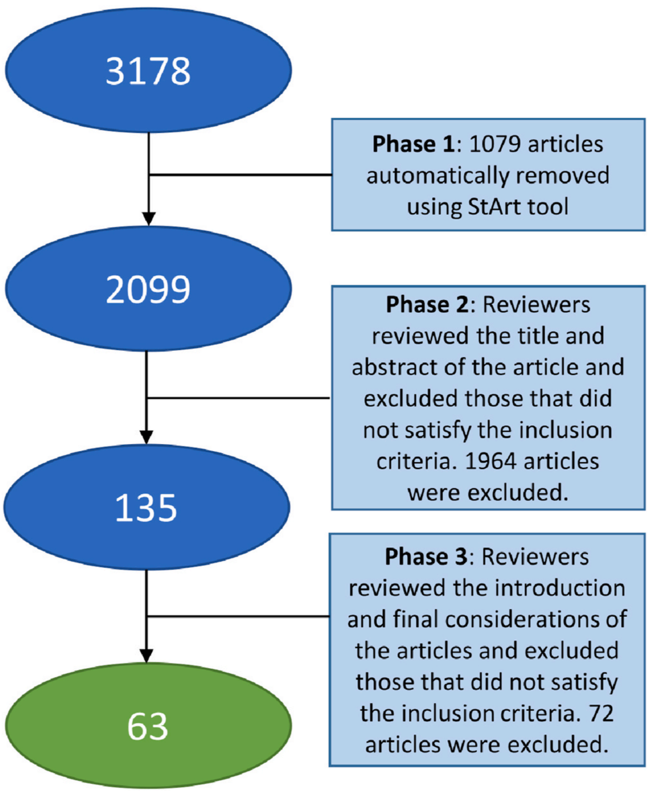

Fig. 1. Selection phases of articles.

and excluded. In summary, in Phase 3, the reviewers excluded 72 articles that did not satisfy the inclusion criteria. About 74% of articles were excluded because they were out of scope, 20% short papers, 2% incomplete studies, 2% duplicated studies, and 2% non-English written papers. The final number of included studies in this systematic literature review was 63.

4. Results

This section summarizes findings of the systematic literature review based on 63 selected studies. The attributes extracted from each study are shown in Table 2.

4.1. Year of publication

The first attribute to be analyzed is the year of publication. Fig. 2 shows the division of studies by year of publication. The figure shows an increase in publications in recent years on automatic feedback. The last four years (2015–2018) had more than 50% of the articles in comparison with the early years (2009–2014). During the period analyzed, the year with the lowest publications was 2010 (n = 3), and the years with the highest publications were 2017 and 2018 (n = 9).

4.2. Type of publication

The second attribute is the type of publication, that is, whether the source of the publication was conference proceedings, journal, workshop proceedings, or others. Table 4 shows the division of types of publications by digital library. Most of the publications were conference papers (71%), followed by journal articles (27%), and workshop papers (2%). Examples of the most common publication venues are the IEEE Transactions on Learning Technologies, Assessment & Evaluation in

Table 4
Type of publication by Digital Library of the selected studies.

Digital Library	Conference	Journal	Workshop
ACM	25 (39%)	1 (2%)	0 (0%)
IEEE	15 (24%)	7 (11%)	0 (0%)
Scopus	4 (6%)	1 (2%)	0 (0%)
Science Direct	0 (0%)	3 (4%)	0 (0%)
Springer	1 (2%)	5 (8%)	1 (2%)
TOTAL	**45 (71%)**	**17 (27%)**	**1 (2%)**

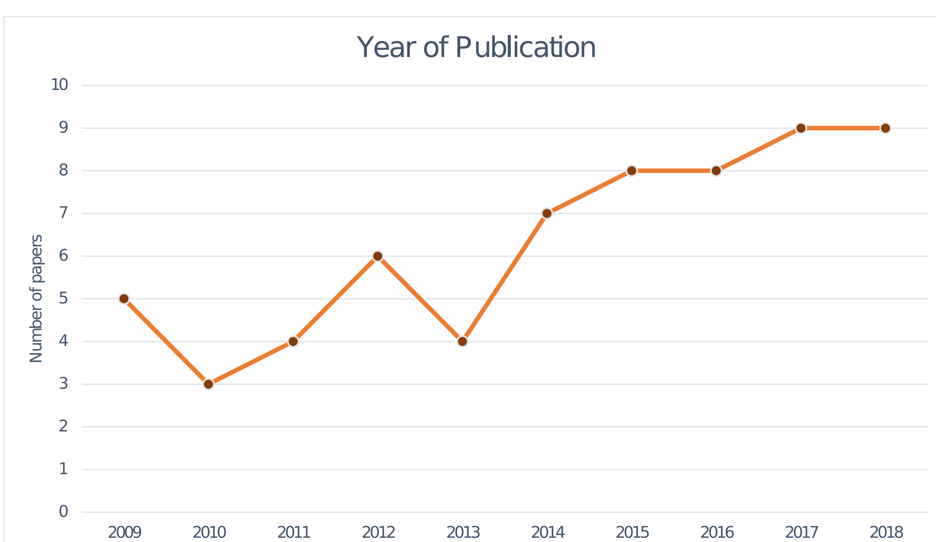

Fig. 2. Year of publication of the selected studies.

Higher Education, Computers in Human Behavior, International Conference on Advanced Learning Technologies, and International Conference on Computing Education Research.

4.3. Publication country

Table 5 shows the number of articles by country, which is derived from the address of the first author of the articles included in the review. The country with the most publications was the USA (n = 10), followed by the United Kingdom (UK) (n = 6), and The Netherlands and China (n = 5 each).

4.4. Subject area

As the articles aimed to propose an automatic feedback system in educational environments, we analyzed subject areas in which automatic feedback was applied and categorized them as shown in Table 6.

The most popular subject area (n = 19) was computer programming, that is, courses that teach some programming language to computer science students. We did not take into account which programming languages were covered in those courses. The second most common subject area (n = 8) was related to different aspects of computer science. It is important to mention that an article may have had more than one course involved. Some more specific subject areas were categorized as "Other subject areas". Among them are courses such as Chemistry, Data Networking, Biotechnology, Biochemistry, Data Structures, SQL Programming, Handwriting, and Electrical Engineering. Some articles did not present any course (n = 5).

4.5. Research questions

In addition to the above attributes, we extracted information relevant to the four research questions, as described in Table 2 in section 2.6.

4.5.1. Research question 1

The first research question, "**Does automated feedback in online learning environments improve students' performance in activities?**" investigated if automatic feedback helped student performance. We coded the articles based on: (i) the result, if the feedback had a positive or negative influence, and (ii) the evaluation, if the study presented an empirical evaluation of an automatic feedback system or not. We also coded those papers with "no evidence" in cases where a feedback practice/tool was described without a consistent evaluation. Table 7 shows the results.

As already reported in other studies, manual instructor feedback helps student performance (Hattie & Timperley, 2007; Parikh et al., 2001; Nicol & Macfarlane-Dick, 2006). This has also been reflected in automatic feedback, with 65.07% of positive results (50.79% with and 14.28% without empirical evaluation), proving that feedback is an important factor in the teaching/learning process, whether it is manual or automatic. For instance, Krusche and Seitz, (2018) administered a survey in a programming language course to the students of a computer science major program to analyze the impact of feedback. Krusche and Seitz, (2018) concluded that automatic feedback increased the students' participation in the exercises and submission of solutions. They also reported that more than 60% of the students successfully completed the course tasks.

Some papers showed positive results with empirical evaluation used a methodology where they first proposed an activity without the use of an automatic feedback tool and then another activity with the aid of the automatic feedback tool. Thus, they showed the change in the students' behavior and their increase in performance, since feedback informed student learning by showing mistakes and successes. For instance, in the study by Krusche and Seitz (2018), students stated that the test results and feedback helped solve the exercises in an introductory programming course, and they enjoyed working with the ArTEMiS tool as it provided instant feedback. In an online questionnaire, the authors found that over 90% of students found interactive instructions useful to improve exercise performance. Kebodeaux et al. (2011) presented a sketch recognition-based tutoring system (Mechanix) that provides immediate feedback on problems in statistics for engineering. The system was evaluated in an introduction to engineering course for 2 semesters on 2 different tasks. The results showed that students who used the tool to answer the task had a significantly higher score (p-value < 0.001), with an average difference of 2.5 out of 10 points, than those who took the course before Mechanix was introduced. These results were attributed to the fact that the tool provides immediate feedback to students before sending the final answer.

Several other studies compared conditions with and without the use of automatic feedback. Al-Hamad and Mohieldin, 2013 proposed an E-assessment tool that supports the design of the assessment of a chemistry course. The tool showed the learning outcomes to the student, where the learning outcomes were augmented with qualitative feedback written by the instructor. Their results showed that the quality of teaching and learning could be improved by using technology to increase faculty efficiency and provide students with real-time feedback mechanisms to help them develop a culture of self-monitoring and self-assessment. In the study by Wong et al. (2012), a tool was proposed for facilitating an efficient and transparent coursework assessment and feedback process. The tool was evaluated in a computer science degree program. Most students who participated in the survey reported that

Table 5
Number of articles by country.

Country	Number of articles
USA	10
UK	6
Netherlands, China	5
Spain	4
Finland, Japan	3
Belgium, Tunisia, India, Germany, Taiwan	2
Australia, Indonesia, Bahrain, Turkey, Korea, Romania, Cyprus, Singapore, Serbia, Malaysia, New Zealand, Brazil, Croatia, Ireland, South Korea, Colombia and Canada	1
Total	**63**

Table 6
Course that the system was applied.

Course	Number of articles
Programming	19
Introduction to Computer science	8
Game Exercises	4
Mathematics	3
Circuits	3
Engineering course	2
Different courses	2
Foreign Language Learning	2
Software Engineering	2
No details	5
Other subject area	16

Table 7
Statistics about the papers related to student performance.

Evidence	Number of articles (%)
No evidence	22 (34.92%)
Positive with empirical evaluation	32 (50.79%)
Positive without empirical evaluation	9 (14.28%)
Negative with empirical evaluation	0 (0%)
Negative without empirical evaluation	0 (0%)
Total	**63 (100%)**

they preferred to receive feedback through the proposed system because the feedback was easy to read and it highlighted mistakes that students made. Results from a comparative study between an experimental group and a control group of students from a programming course, showed that using the proposed tool (an online multiple-choice questions system integrated with a neural network) improved the learning outcomes (Alemán et al., 2010).

Studies that did not perform empirical evaluation but indicated positive results (14.28%) generally focused on assessing student satisfaction with tools. For example, the results of the study by Wang et al. (2018), which proposed a data-driven program repair framework to automate feedback generation for introductory programming exercises. The study showed that the system was effective and could generate concise and useful feedback for 89.7% of the incorrect student submissions, in just 2 s on average. The study by Keuning et al. (2014) presents a prototype of a programming tutor to help students with feedback and hints. The authors found that they could recognize between 33% and 75% of the exercise solutions collected during two programming courses. Zhou et al. (2018) analyzed the design of existing online judge systems (Wasik et al., 2018) and their advantages and problems in applying to programming education. The authors state that after applying the system in a course on the C programming language, the students' performance and satisfied grades increased. However, the article does not show details about this assessment.

Papers that did not show any evidence (34.92%) are more descriptive, showing details about the tools and how they work. It means that they did not present evidence on how feedback potentially enhanced student performance. For example, the study by Lan et al. (2015) presents the development of a framework for mathematical language processing (MLP). As a result, the authors stated that the structure could substantially reduce the human effort required for classification in large-scale courses and also allows instructors to visualize solution groups to help them identify groups of students with the same misconceptions. The study by Lodder et al. (2017) describes a system that is part of a set of tools that help students study logic by providing automatic feedback. The authors state that the system's performance for resolution logic proofs reached a quality comparable to that of a group of experts. The work proposed by Ying and Hong (2011) presents a SQL (Structured Query Language) teaching system with an automatic feedback mechanism. The system helps the student in the construction of SQL queries. This system provides tips to assist the students in understanding a specific concept of SQL more quickly and then verifies the effectiveness of the student solution to the exercise.

4.5.2. Research question 2

The answer for the second research question, "**What are the main goals in using automatic feedback generation techniques in online learning environments?**" is shown in Table 8. The articles had very specific objectives. We grouped the articles included into this systematic literature review into four categories based on the objectives of the feedback approaches.

The articles that help students in a particular content/course (52.38%) are systems developed to assist programming courses (Karavirta et al., 2012; Arends et al., 2017), the teaching of circuit analysis (Baneres et al., 2014; Weyten et al., 2010), and teaching of a foreign language (Ono et al., 2013; Murad et al., 2018), among others. These systems often provide feedback to show students what went wrong (showing where the error is and giving tips on how to get the answer right) or well (showing a congratulations message) (Krusche & Seitz, 2018; Marin et al., 2017). In the article by Weyten et al. (2008), a new web-based system for training students in electrical and electronic circuit theory is presented. The system can be used to gain valuable information from students and thereby bring improvements in instructor teaching. The article by Bryfczynski et al. (2013) describes a system called beSocratic, which assists students who study data structures; the students can be evaluated and the results of their task completions are analyzed automatically to help instructors refine their activities and improve future performance. The study by Ono et al. (2013) proposed a new type of feedback system based on a text mining method. The system encourages students to reflect on their own presentation and has shown positive results in the use of foreign language teaching in Japan. The tool proposed by D'antoni et al. (2015) aims to provide feedback for the construction of a deterministic finite automaton that accepts chains corresponding to a described pattern. The system provides automatic feedback with counterexamples or tips so that students can complete the activity.

In contrast to the first main goal, the second goal (41.26%) of the included studies was to provide more general feedback to promote self-regulated learning. These articles generally are focused on providing personalized feedback (Demaidi et al., 2018), gamification (Utomo & Santoso, 2015; Ying et al., 2012), or dashboards (Davis et al., 2017; Yu, 2016) in an online environment to motivate students, detect poor performance and reduce dropouts (Khan & Pardo, 2016). The paper by Jin (2017) presents a visualization tool to motivate students to participate in collaborative online learning communities actively. The work of Smithies et al. (2010) presents a tool called CONSPECT, which aims to provide formative feedback and monitor students' conceptual development. It uses an NLP method, based on latent semantic analysis, to compare student answers to generated reference models. The article by Alencar and Netto (2014) introduces TUCUMÃ, an intelligent 3D virtual agent integrated with Moodle for virtual learning. The tool automatically simulates a distance course tutor, monitors student activities, and answers student questions through dialogue.

As Table 8 shows, only 3 studies aimed to assist instructors. Several studies that introduce approaches to help students in online learning operate under the assumption that automatic feedback can also benefit instructors in terms of teaching efficiency (Martin et al., 2009; Xie & Li, 2018). We hypothesize that if the student can learn from automatically generated feedback, these systems have great potential to reduce the effort of the instructors in answering questions or giving feedback to the students. Our third research question explored this assumption.

The study by Akçapınar (2015) is the only article that presented an automatic feedback system with the goal of reducing student plagiarism behavior. This study aimed to reduce students' plagiarism in written tasks by providing automated feedback based on text mining analysis.

4.5.3. Research question 3

The third research question, "**Are there indications that automatic feedback helps instructors?**" aimed to understand if the approaches

Table 8
Main goals for using automatic feedback generation.

Goal	Number of articles (%)
Use feedback to help students on a specific content/course	33 (52.38%)
Use feedback to support self-regulation	26 (41.26%)
Use feedback to help instructors	3 (4.76%)
Use feedback to reduce plagiarism behavior	1 (1.60%)
Total	**63 (100%)**

Table 9
Numbers of papers that show the support of automatic feedback system to instructors.

Evidence	Number of articles (%)
No evidence	29 (46.03%)
Positive with empirical evaluation	6 (9.52%)
Positive without empirical evaluation	28 (44.44%)
Negative with empirical evaluation	0 (0%)
Negative without empirical evaluation	0 (0%)
Total	**63 (100%)**

proposed in the literature provided insights and assisted instructors during preparation and teaching phases. Table 9 shows the information found in the literature as the response to this question.

Most articles (46.03%) have not shown any evidence in their findings of whether automatic feedback helps the instructor. This result corresponds to the objectives of existing studies (Section 4.5.2), where 93.64% of studies (articles that use feedback to help students on specific content or in specific disciplines and articles that use feedback to support learning) aimed to assist student learning using automated feedback. Only 3 studies (4.76%) aimed to help instructors (see Table 8). Among these studies, Martin et al. (2009) were able to support instructors' needs when they tried to integrate various learning systems to improve students' learning process. The authors proposed a system called MAGADI that helps instructors with visualizations of relevant information about students. Trausan-Matu et al. (2014) proposed the PolyCAFe system, which provides tools that support a polyphonic analysis of chat conversations and discussions of small student groups on online forums. The system uses NLP to identify topics, semantic similarities, and links between utterances. A statement chart is created with the detected links, which is the central element for polyphonic analysis and for providing automatic feedback and support for instructors and students. The study by Xie and Li, (2018) proposes a system that combines a recommendation model based on big data content and a clustering model to personalize exercises and feedback in online education.

The majority of the papers presented positive results, with and without empirical evaluation, (53.96%). The main goal of the studies selected in this review, as shown in Table 8, was to assist online learning in specific disciplines. Furthermore, these studies also indicate success in reducing the instructor's workload, since the amount of questions and problems from students is reduced using the automatic feedback system (Krusche & Seitz, 2018). Most articles (44.44%) claim that automatic feedback helps instructors but offer no empirical evaluation to support such claims. These studies demonstrate in their results the satisfaction reported by instructors regarding the reduction of students' difficulties (Krusche & Seitz, 2018; Wong et al., 2012) or the reduction of instructors' workload (Fast et al., 2013; Marin et al., 2017). There were no negative results in this analysis.

4.5.4. Research question 4

The fourth research question aimed to analyze which methods and techniques are used in the automatic generation of feedback. Table 10 shows the main methods and techniques found in the articles.

The main technique used was the comparison between student answers and the desired solution (15 articles). Among these articles are those that aim to propose feedback to help students solve specific exercises in subjects, such as programming, circuit analysis, automation, among others. In this way, the proposed systems provided instant feedback comparing a student's response with a possible response already registered in the system. The research by Lodder et al. (2017) used this method to determine the quality of LOGAX (a tutoring tool that helps students to build an axiomatic proof), comparing the proofs generated by experts and student solutions.

Many articles (n = 14) have not detailed the methods or techniques used to generate automatic feedback. Most of these articles propose prototype systems that are still in the development phase (Efstathiou et al., 2018; Jeremić et al., 2012; Riofrío-Luzcando et al., 2017). Other articles do not describe how the proposed systems were developed, but only describe how they were applied in a real environment and the results of the implementation (Al-Hamad and Mohieldin, 2013; Kebodeaux et al., 2011; Wong et al., 2012; Ying et al., 2012).

The second most used technique was dashboard (n = 7). These studies used graphical elements to motivate the students or the class to carry out activities in the online environment. For example, the article by Khan and Pardo (2016) presents a study that categorized students based on how they interact with the dashboard, taking into account time, number and timing of hits.

The third most used technique was NLP, which is a field of computer science applied to manipulate text or speech in natural language. It can process and analyze large amounts of textual data using algorithms for semantic and syntactic analysis (Ferreira-Mello et al., 2019). For example, Trausan-Matu et al. (2014) proposed a system which provides tools that support the polyphonic analysis of chat conversations and online discussion forums, and NLP is used to identify topics, semantic similarities and links between utterances. In the study by Ono et al. (2013) text mining technology was used to produce instant feedback in a foreign language presentation course.

Other specific methods found in the articles are: longest common subsequences (LCS), feature extraction with clustering, cybernetic principles, linguistic analysis engine, tree edit distance, abstract syntax trees (ASTs), knowledge databases, predictive analytics, mobile sensors, and data mining. It is important to note that 1 article can have more than 1 method.

Table 10
Main methods and techniques used to generate automatic feedback.

Method	Number of articles
Comparison with desired solution	15
No details	14
Dashboard	7
NLP	7
Ontology	4
Graphs	3
Neural Network	2

5. Discussion

Based on the insights obtained from this systematic literature review, we highlight three factors that should be considered when researching and developing systems to provide feedback. These factors include methods and goals, relevance for instructors, and techniques adopted. Based on our results, these factors were considered critical in the process of sending feedback. Each of the three factors is discussed in the remainder of the section.

5.1. Feedback impact and educational goals

The first research question aimed to assess the impact of the automatically provided feedback on students' performance. In this case, performance could be related to a specific activity or the final marks. Unsurprisingly, the majority of the papers retrieved in this review, about 65% (Table 7), concluded that the feedback had a positive impact on students' performance (Hattie & Gan, 2011). However, the papers do not provide enough information to determine if the positive impact was caused by the use of the tool or the final feedback product (dashboard/message). For instance, several papers proposed tools to evaluate programming activities automatically (Gulwani et al., 2014; D'antoni et al., 2015; Birch et al., 2016). In this case, the authors do not analyze if the improvement in the students' abilities was due to the usage of the entire system or just because of the feedback. Price et al. (2010) suggests that the perceived value of feedback and the students' final performance should be analyzed separately.

Additionally, a few papers reported an increase in the students' performance, but some degrees of dissatisfaction with the feedback message. In this direction, Burke (2009) presented several factors that led to poor evaluation of the feedback, even with the improvement of final marks. The students listed the feedback length (brief), polarity (always negative), and complexity (difficult to interpret) as the main drawbacks (Burke, 2009). Possible reasons for this can be the lack of training related to good feedback practices. Weaver (2006) showed that more than 50% of university students never received any guidance on "how to understand and use feedback", and three-quarters of students received no advice on how to understand and use feedback before university, and Mutch (2003) highlighted the need for more research on how students "receive and respond" to feedback. This is in line with what Carless and Boud (2018) refer to as the importance of student feedback literacy to enhance the feedback impact. Carless and Boud (2018) also state that the instructors have a key role to enable students to appreciate, make judgments, manage affect, and take action on the feedback messages. Moreover, the current literature also offers recommendations for good feedback practices. Nicol and Macfarlane-Dick (2006) proposed a conceptual model of self-regulation based on a review of the research literature on formative assessment and feedback. The main idea of the work is to identify how the training processes of evaluation and feedback can help promote self-regulation. Based on the conceptual model, seven principles of good feedback practices to enhance teaching feedback were proposed.

In our review, we also analyzed the educational goals of the feedback systems. Table 8 revealed that more than half of the systems (52.38%) aimed to provide feedback about a specific content/course. More specifically, the majority of these papers were applied to student performance (Table 7) and more procedural and specific activities, such as Programming and circuit analysis (Table 6). This result could explain the possible reasons for the weaknesses such as length (brief), polarity (always negative), and complexity (difficult to interpret).

A total of 41.26% of the articles in this review used feedback as a method to support self-regulation. This goal is more aligned with the literature on good practices of feedback (Nicol & Macfarlane-Dick, 2006; Hattie & Timperley, 2007). However, these works did not present an analysis to support the effectiveness of the feedback in terms of improving the students' performance and self-regulation processes (41.27% listed as no evidence in Table 7). Moreover, the majority of these papers proposed the adoption of a dashboard to support students. Although, the literature shows that this kind of visualization does not guarantee effective feedback and does not offer sufficient support for self-regulated learning (Matcha et al., 2019).

The papers in this review have not considered several factors that are well-established in the literature to enhance the feedback process. They do not align the proposed feedback systems with educational research on the provision of feedback, which could be extremely helpful in order to improve the final result of the feedback process, in terms of learning outcomes, learning processes, and students' satisfaction.

There are several popular frameworks for good feedback practices that are proposed in educational research. For instance, Nicol and Macfarlane-Dick (2006) suggested incorporating more than just simple instructions in feedback messages. According to Nicol and Macfarlane-Dick (2006), good feedback practice is broadly defined as anything that might strengthen students' capacity to self-regulate their performance. Although, Nicol and Macfarlane-Dick (2006) suggested seven good practices to enhance feedback, the papers included in this literature review just focused on providing specific information related to the student activities. It is a limitation that could have influenced the students' satisfaction level reported in the studies.

Educational research has documented factors that should be considered when creating feedback. Hattie and Timperley (2007) investigate several conditions that could maximize the positive effects of feedback on student learning, including the increase in student awareness about the overall learning goal, the progress towards the goal, and the subsequent goals required to achieve the overall goal. Hattie and Timperley (2007) also propose a model that encapsulates four levels of information to be considered in feedback messages: (i) task level such as whether the activity is correct or incorrect, can include instructions for more or different information; (ii) process level includes suggestions about study methods to the student to create a product or complete a task and is more directed to information processing, or learning processes that require understanding or completing the task; (iii) self-regulation level which includes greater self-assessment or confidence skills, can have major influences on self-efficacy, self-regulatory proficiency, and students' personal beliefs as learners; (iv) self-level, feedback can be personal in the sense that it is directed to the self; self-level is often unrelated to task performance. Hattie and Timperley (2007) research showed that the most potent feedback is on process and self-regulation levels, while self-level is usually ineffective for learning. Task level is typically ineffective unless it is combined with either process or self-regulation levels. Feedback in the systems proposed in the papers included in this review is focused on the task level only, which reduces the potential of feedback to positively impact in student motivations and participation in class (Robison et al., 2009).

Finally, the papers in this review do not consider feedback as a dialogic process. Pardo (2018) and Pardo et al. (2019) proposed that feedback should be a process where students and instructors have a conversation about the course, assisting not only the students to understand the course content better, but raising the capability of the instructor to personalize the content and improve the course design and orchestration (Dillenbourg et al., 2013; Prieto et al., 2016). Furthermore, Pardo et al. (2018) also advised that timely feedback increases the chances to help students to reach the learning goals and improve their final performance. Among other things, this concept could reduce student dropout rates (Lee & Choi, 2011).

5.2. Feedback relevance for instructors

The results of this study suggest that the existing feedback systems do not consider the instructors' needs. Table 8 shows that only 4.76% of the papers initially intended to support instructors. However, none of the systems proposed a platform to assist instructors/teachers to write better feedback (Harvey, 2003; Mulliner & Tucker, 2017). The arguments for the importance of students in the feedback process are undeniable (Brookhart, 2017). However, recent literature also advises that the instructor role is crucial in the adoption of automatic tools for the provision of automated feedback (Lim et al., 2019, p. 101202; Pardo et al., 2018). More broadly, the importance of the instructor in the adoption of education technology has already been made by several researchers (Ali et al., 2013; Gašević et al., 2017, 2019; Rogers, 2000; Zhao & Cziko, 2001).

The results also point out that several feedback systems (53,96%) have impacted the instructor experience positively, as shown in Table 9. Possible reasons for this can be found in the capability of automation and personalization provided by the feedback systems, which can potentially decrease the instructors' workload (Hentea et al., 2003; Manoharan, 2016; Sheridan, 2006). The majority of the works retrieved in this review that reduce the activities performed by instructors are based on student dashboards, showing simple statistics (Kebodeaux et al., 2011; Yu, 2016), systems comparing the students' results with a predefined desired solution (Baneres et al., 2014; Lan et al., 2015), or feedback in a particular domain, e.g., programming language problems (Alemán et al., 2010; Arends et al., 2017; Helminen & Malmi, 2010; Karavirta et al., 2012; Keuning et al., 2014) and essays evaluation (Toshniwal et al., 2015; Usener, 2015; Whitelock et al., 2015). In a nutshell, this result reveals a preponderance of papers related to intelligent tutoring systems in the provision of automatic feedback which explains the decrease of instructors' workload (Polson & Richardson, 2013). Nevertheless, this approach fails to provide analytics to inform the instructor to support the feedback process and inform their teaching alongside an automatic feedback system.

Current educational research indicates that supplying instructors with relevant information about the students and the learning environment could enhance the capability of instructors to provide more informative feedback at scale and adjust the course content/methodology to reach better educational results (Dillenbourg et al., 2013; Pardo et al., 2019; Prieto et al., 2016). Learning dashboards focusing on the use of visualizations to support instructors are potentially a powerful instrument to understand student behavior supporting the provision of feedback (Charleer et al., 2014; Verbert et al., 2014). However, learning dashboards have to be carefully designed to support instructional decision-making. Wise and Jung (2019) concluded that a learning dashboard for instructors should contain informative content regarding the students' activities and learning context; otherwise, the instructors will not engage or take action based on the visualization.

Many authors define feedback as a dialogical process whereby learners obtain information on their performance and instructors better understand students' needs (Boud & Molloy, 2013; Pardo, 2018). In other words, feedback should not be unidirectional from instructors to learners, but it has to incorporate information for both actors. From the instructor's point of view, the dialogue enabled by feedback could aid the process of refining course design and the orchestration of activities (Dillenbourg et al., 2013; Prieto et al., 2016; Wise, 2014). Therefore, an essential improvement in the current feedback systems is to provide support for both instructors and students and the entire process of feedback instead of just supporting specific tasks in a course design, such as programming tasks.

Finally, Dawson et al. (2019) advise including only content-related information for students is not enough to provide a good feedback message, it is also important to include affective aspects in feedback that encourage positive motivational beliefs and provide information that can be used to help shape teaching (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). Thus, future research on systems that aim to assist instructors with feedback provision is the analysis of the message content to suggest improvement of quality of the overall feedback, including non-content aspects (Cavalcanti et al., 2019, 2020).

5.3. Techniques adopted to provide feedback

The last research question of this study aimed to identify which methods, tools, and techniques were applied to provide the feedback and discuss their alignment with the educational goals and student performance. Specifically, we analyzed how researchers develop approaches to create and send feedback messages in online environments. Our analysis suggests that commonly adopted methods are the direct comparison of students' answers with the desired solutions (predefined by instructors) (Birch et al., 2016; Dutchuk et al., 2009; Lodder et al., 2017; Mitrovic et al., 2011; Usener, 2015), dashboards/graph visualizations (Bodily et al., 2018; Jin, 2017; Khan & Pardo, 2016; Yu, 2016), and natural language processing/machine learning (Corrigan et al., 2015; Jugo et al., 2014; Ono et al., 2013; Trausan-Matu et al., 2014).

The majority of the papers that focused on comparing students' answers with the desired solutions were reported in programming or automation courses where the main goal is to evaluate programming activities automatically, providing information on the students' performance and possible improvements to enhance software programs (as shown in Table 6). The literature confirms the importance of answer comparison to provide feedback to students (Ihantola et al., 2010; Keuning et al., 2016), but it has two main limitations: (i) it is necessary for the instructor to register the answer in the system beforehand; and (ii) the student must respond the same as the answer given by the instructor. Unsurprisingly, it provides minimal information to students and is not connected with the good feedback practices found in the educational research literature (Nicol & Macfarlane-Dick, 2006). For instance, Nicol and Macfarlane-Dick (2006) suggest that for feedback to be effective, it is necessary to provide more valuable information such as helping to clarify good performance or encouraging students with positive motivational beliefs. More importantly, it should offer guidance to the students in terms of learning strategy (Hattie & Timperley, 2007) they can adopt to learn the concept they missed to answer correctly. To achieve this, automatic feedback should not only consider students' responses on assessments, but it should also include data about how is the students' learning process. Therefore, the recent literature recommendation for this kind of information is to inform instructors on students' progress systematically, so they could write effective feedback messages on student activity and performance (Blikstein et al., 2014; Pardo et al., 2019).

Dashboards and visualizations have also been widely used to provide student feedback on their learning process and progress (Bodily et al., 2018; Jin, 2017; Khan & Pardo, 2016; Schwendimann et al., 2016; Yu, 2016). Few studies demonstrated that these visualizations were effective in improving students' performance (Davis et al., 2017; Utomo & Santoso, 2015). However, a systematic review on learning analytics dashboards by Matcha et al. (2019) reveals the negative effects of dashboards on students and the need for improvement in dashboards to address the recommendations for effective feedback. None of the studies included in this review presented enough evidence of effectiveness in providing feedback for students by dashboards.

Some authors used NLP and machine learning techniques to provide or assist in the feedback process (Corrigan et al., 2015; Jugo et al., 2014; Ono et al., 2013; Trausan-Matu et al., 2014). The development of the fields of learning analytics and educational data mining could explain the increase in the adoption of these methods that should become a trend in the area of (semi-)automatic feedback systems (Cavalcanti et al., 2020; Er et al., 2020; Tempelaar et al., 2020; Tsiakmaki et al., 2020). However, this kind of application requires a substantial amount of data to build a consistent model that works for different contexts (Barbosa et al., 2020; Romero & Ventura, 2020). In sum, this line of work has a considerable potential to provide useful information, but problems such as data contamination should be carefully avoided (Farrow et al., 2019).

6. Threats to validity

Considering that a systematic review process has some steps that can be subjective, this section presents the threats to the validity of this review, which were classified in the categories Construction, Internal, External, and Conclusion (Wohlin et al., 2000).

6.1. Construction validity

This type of threat refers to when problems occur in the construction of the process to carry out a systematic review. This review aims to select papers related to automatic feedback in online learning environments. For this, a search string was created based on the main terms "feedback", "online learning environment", "student", and "instructor". As we know that there are several synonyms for these concepts and some synonyms are used in many educational literature papers, it was necessary to study the main synonyms of these words so that our search string could cover the largest number of papers related to automatic educational feedback. This string was later used to select articles in some bibliographic repositories. Because of the combinations used, it is possible that the created sequence does not cover all articles relevant to the topic of interest. On the other hand, it was considered the main vehicles of scientific dissemination that allowed the self-extraction of articles from research using a research string, which reduces the threats to validity. In this way, the threats to the validity of this type were reduced using synonyms for keywords and a search in the main bibliographic databases.

6.2. Internal validity

During the process of selecting and extracting the data from the works, subjective decisions had to be made due to the absence or imprecise description of relevant information in the studies analyzed, making it difficult to objectively apply the inclusion/exclusion criteria or the impartial extraction of the data. However, to minimize possible impacts, the review took place interactively and collaboratively, with the same study being analyzed by at least two collaborators. In addition, in cases of disagreement, a third collaborator was contacted. In this way, we tried to mitigate the threats due to the personal bias in understanding the study.

6.3. External validity

According to Dermeval et al. (2016) external validity is concerned with establishing the generalization of the SLR results, as well as the representative level of the primary studies concerning the reviewed topic. The authors validated the review protocol and automatically tested it in bibliographic repositories to minimize such a threat.

6.4. Conclusion validity

Another threat is that the proposed protocol allows the exclusion of relevant articles. That is, it is possible that some studies excluded in this review should have been included. To mitigate this threat, the authors carefully designed and discussed the inclusion/exclusion criteria to minimize the risk of exclusion of relevant studies.

7. Limitations

The primary limitation of this study is related to the search process in which we focused on papers that only contain feedback provided in online environments. This could potentially exclude papers that describe feedback systems, but that were not evaluated in a virtual context.

Second, a few papers had limited information about the methods and techniques used, which led to several categories such as “no details” and “no evidence” in the result tables. We decided to keep these papers nevertheless because they contain information relevant to at least one research question.

Finally, this review did not focus on systems that provide more information and algorithms different from the feedback process (i.e., Intelligent Tutoring Systems), and papers that do not focus on online education.

8. Conclusions

This article presents an overview of existing studies on automatic feedback in online learning environments from 2009 to 2018. It analyzed the benefits that automatic feedback generation can bring in relation to instructors and students. The systematic literature review showed the main techniques used and the main objectives in applying automatic feedback in online learning environments. The research questions were answered by analyzing the results of the articles and verifying whether an empirical or non-empirical evaluation was performed with positive or negative results. Research questions examined whether automatic feedback helps student performance, whether it helps the instructor, and whether it can override and be more efficient than manual feedback.

We concluded that there is evidence that automatic feedback increases student performance in activities (50.79% of articles). The main purpose of using automatic feedback systems was to help students on a specific content/discipline. Moreover, the majority of these articles have the same type of assessment: comparing students’ scores in a discipline before using the system and after using the system. In this case, the studies did not show an analysis of other factors besides the feedback that could influence these results.

This study also assessed if automatic feedback can also help the instructor. As described in many of the articles included in this review, the objective of the automatic feedback systems is precisely to decrease the instructor’s effort in correcting various student exercises. Our results confirm this statement showing that there is evidence that automatic feedback also helps reduce instructor effort (53.96% of articles). Finally, we found that the main methods and techniques used to generate automatic feedback were: comparison with desired solution, dashboards and NLP.

This systematic literature review highlighted that the main shortcoming in the research literature about the automatic provision of feedback are: (i) the insufficient use of educational research on feedback to inform development tools for automatic feedback; and (ii) the exclusive focus on students which neglect the role of teachers in feedback practice. Therefore, this study proposed the following recommendations for further research:

- **Develop tools focused on the instructors**: Providing tools for instructors would inform their teaching practice and even involve them in the improvement of feedback for students. Although there are recent initiatives in this direction (Lim et al., 2019, p. 101202; Tsai et al., 2021), in general, the feedback tools proposed focused only on the students.
- **Analysis of feedback quality**: Many studies included in this review aimed to provide automatic feedback. However, few papers have attempted to analyze the quality of feedback provided through forms applied to students and/or instructors. The recent paper by Cavalcanti et al. (2019) focused on the analysis of the feedback quality extracted from evaluations collected in an online course offered at a Brazilian higher education institution. It shows the potential of using machine learning to achieve this goal.
- **Automatic feedback generation**: Almost all work reviewed in this study aimed to provide feedback for a specific context, for instance, introduction to programming, circuit analysis, and foreign language essays evaluation. Yet, the papers reviewed did not present any evidence of the generalizability of their approaches. One possible solution is to use Natural Language Generation (NLG) techniques to produce automatic feedback (Perera & Nand, 2017). NLG is defined as a systematic approach to produce human-understandable natural language texts based on analytics or representations of meaning.

The results presented in this systematic review can be of great use to the community, as it gathers evidence from the primary studies included in the review, forming a body of knowledge about the use of automatic feedback in Online Learning Environments. Table 11 shows the details of the selected articles.

Declaration of competing interest

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

Appendix A. Summary of Selected Papers

Table 11
Summary of all papers retrieved in this review

Reference	RQ 1	RQ 2	RQ 3	RQ 4	Study Overview
(Karavirta et al., 2012)	No evidence	Use feedback to support programming learning	Positive without empirical evaluation	longest common subsequences (LCS)	This paper presents a tool that facilitates the learning of programming by providing a mobile application for Parsons problems.
(Arends et al., 2017)	No evidence	Use feedback to support programming learning		Domain reasoner	

(continued on next page)

Table 11 (*continued*)

Reference	RQ 1	RQ 2	RQ 3	RQ 4	Study Overview
			Positive without empirical evaluation		This paper presents a prototype intelligent tutoring system (MicK) for programming learning in embedded systems
(Krusche & Seitz, 2018)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	No details	This article presents an AuTomated Assessment Management System (ArTEMiS) for interactive learning, evaluating solutions to program exercises automatically, and provides instant feedback so students can solve iteratively.
(Marin et al., 2017)	No evidence	Use feedback to support programming learning	Positive without empirical evaluation	Graphs	This paper presents a semantic-aware technique to provide personalized feedback that aims to mimic an instructor looking for code snippets in student submissions.
(Fast et al., 2013)	No evidence	Use feedback to support formal logic learning	Positive without empirical evaluation	Comparision with database	This paper presents a system for creating, grading, and analyzing derivation assignments across arbitrary formal domains (DeduceIt), provides students with incremental feedback, and aggregates student paths through each proof to produce instructor analytics.
(Khan & Pardo, 2016)	No evidence	Use feedback to students engagement	No evidence	Dashboard with students interactions	This paper describes a case study of a flipped learning strategy to explore the behavior of students interacting with a dashboard updated in real time providing indicators of their engagement with the course activities.
(Kebodeaux et al., 2011)	Positive with empirical evaluation	Use feedback to support sketch recognition	Positive without empirical evaluation	No details	This paper presents a sketch recognition based tutoring system (Mechanix) that provides immediate feedback for engineering statics problems.
(Ying et al., 2012)	No evidence	Use feedback to support students engagement with games	No evidence	No details	This work has developed an online multiplayer based game learning system (MOGLS), which based on the ARCS motivation model that provides learning registration, classification record, end-of-test feedback to motivate students to learn.
(Utomo & Santoso, 2015)	Positive with empirical evaluation	Use feedback to motivate students	Positive without empirical evaluation	Dashboard	This paper presents the idea of developing a pedagogical agent is to assist the facilitators in providing automatic feedback to the students based on their behavior in e-Learning system.
(Al-Hamad and Mohieldin, 2013)	Positive with empirical evaluation	Use feedback to support students learning	Positive without empirical evaluation	No details	The present study is part of an ongoing quality assurance and enhancement framework to develop innovative computer-based assessment methods at University of Bahrain. The e-assessment tool supports the design of the assessment of the course intended learning outcomes augmented with instantaneous qualitative feedback to the student.
(Usener, 2015)	Positive with empirical evaluation	Use feedback to support binary tree data structure operations	Positive without empirical evaluation	Comparision with desired solution	This paper presents EASy-DSBuilder, an e-assessment tool for assessing fundamental concepts of binary tree data structure operations taught in Computer Science (CS) lectures.
(Wong et al., 2012)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	No details	This paper presents a system (eCAF) supports detailed marking scheme editing and enables tutors to use such schemes to pin-point errors in students' work so as to provide helpful feedback efficiently.
(Zhang & Jia, 2017)	Positive with empirical evaluation	Use feedback to support students learning	Positive without empirical evaluation	No details	This paper presents a system uses big data analysis techniques to analyze students' online learning behavior and provides students with personalized counseling for evaluate the teaching effects in schools, and put forward some suggestions for improvement based on the results of analysis.
(Alemán et al., 2010)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	Neural Network	This paper presents an experience of generating diagnostic feedback for guided learning in an introductory programming course. An on-line Multiple Choice Questions (MCQs) system is integrated with a neural network based data analysis.
(Baneres et al., 2014)	Positive with empirical evaluation	Use feedback to support students learning	Positive without empirical evaluation	Comparision with desired solution	This paper presents a VerilUOC system, an educational platform to support digital circuit design learning, allowing students to solve the exercises and verify if the solution is correct, receiving continuous feedback on the errors that can be used to review and correct the initial design.
(Gulwani et al., 2014)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	Comparision with desired solution	This paper propose a light-weight programming language extension that allows a teacher to define an algorithmic strategy by specifying certain key values that should occur during the execution of an implementation, using a dynamic analysis based approach to test whether a student's program matches a teacher's specification.
(Davis et al., 2017)	Positive with empirical evaluation	Use feedback to regulate their learning behavior	No evidence	Dashboard with feedback metrics	This paper presents a personalized feedback system that facilitates social comparison with previously

(*continued on next page*)

Table 11 (*continued*)

Reference	RQ 1	RQ 2	RQ 3	RQ 4	Study Overview
					successful learners based on an interactive visualization of multiple behavioral indicators.
(Lodder et al., 2017)	No evidence	Use feedback to support axiomatic proof learning	No evidence	Comparision with desired solution	This paper describe LOGAX, which is part of a set of tools that help students in the study of logic, as a tool to practice rewriting formulas in the normal disjunctive or conjunctive form, and to prove equivalence using standard equivalences, using an algorithm to generate axiomatic proofs, and the generation of suggestions and feedback based on this algorithm.
(Martin et al., 2009)	No evidence	Use feedback to help teachers	Positive with empirical evaluation	Dashboard with students information	This article presents an adaptive feedback generation tool (SigMa) for the three types of actors involved in a mixed learning process: teachers, students and the e-learning platform, which focuses on the teacher's tool and defines its desirable functionalities and components .
(Akçapınar, 2015)	Positive with empirical evaluation	Use feedback to reduce plagiaristic behavior	Positive with empirical evaluation	Text Mining	This study is intended to decrease the plagiarism behavior of students in online tasks by providing automated feedback based on the analysis of text mining based on documentary similarity analysis.
(D'antoni et al., 2015)	Positive with empirical evaluation	Use feedback to help students learn automata theory	No evidence	Comparision with desired solution	This article studies the effectiveness of feedback types in the learning process of a finite deterministic automaton that accepts strings that correspond to a described pattern.
(Helminen et al., 2012)	No evidence	Use feedback to support programming learning	No evidence	No details	This paper presents an use of several methods to extract meaningful information from Parsons' troubleshooting sessions based on a new data source for detailed, interactively recorded interaction traces.
(Helminen & Malmi, 2010)	No evidence	Use feedback to support programming learning	Positive without empirical evaluation	Comparision with desired solution	This paper presents a web-based tool (JYPE) that provides an environment for viewing line-by-line execution of programs in Python and for solving programming exercises with support for immediate automatic feedback and an integrated visual debugger that allows you to go back in the execution view as if it was running in reverse order.
(Lan et al., 2015)	No evidence	Use feedback to help students with Mathematical Questions	Positive with empirical evaluation	Feature Extraction with Clustering	This paper presents the development a framework for mathematical language processing (MLP), takes inspiration from the success of natural language processing for text data and comprises three main steps, for reduce the human effort required for grading in large-scale courses.
(Bhatia et al., 2018)	No evidence	Use feedback to support programming learning	No evidence	RNNs with constraint-based synthesis	This paper presents a technique to combine Recurrent Neural Networks (RNNs) with constraint-based synthesis can provide a basis for providing effective feedback on student programs with syntax errors.
(Westera, 2015)	No evidence	Use feedback to regulate their learning behavior	No evidence	Cybernetic principles	This paper have explored how cybernetics in principle can be used to generate learner feedback in complex learning environments, by constructing a (dual) cybernetic control loop, error-correcting performance feedback can be generated, which can in principle be used for supporting learners.
(Demaidi et al., 2018)	Positive with empirical evaluation	Use feedback to improve students learning	Positive without empirical evaluation	Ontology	This paper contributes to research carried out on personalized feedback frameworks by proposing a generic novel system which is called the Ontology-based Personalized Feedback Generator (OntoPeFeGe).
(Whitelock et al., 2015)	Positive with empirical evaluation	Use feedback to help students with essay writing	Positive without empirical evaluation	Linguistic analysis engine	This paper focuses on the use of a natural language analytics engine to provide feedback to students when preparing an essay for summative assessment.
(Lee & Kim, 2009)	No evidence	Use feedback to help students with simultaneous equations	No evidence	No details	This paper reports on the progress of our project, which explores the adoption of a pen-based interface for teaching intelligent math for simultaneous equations and efficient user interfaces for an intelligent math tutoring system.
(Belcadhi, 2016)	Positive with empirical evaluation	Use feedback to support students learning	No evidence	Ontology	This paper proposes a customized and intelligent feedback structure based on Semantic Web technologies, which provides personalized feedback for self-assessment and is appropriate for the Lifelong Learning environment.
(Trausan-Matu et al., 2014)	Positive with empirical evaluation	Use feedback to support to tutors and learners	Positive with empirical evaluation	Natural Language Processing	This paper presents in more detail the polyphonic model proposed and provides novel insights about the associated analysis method and the computer support provided by PolyCAFe.
(Riofrío-Luzcando et al., 2017)		Use feedback to support students learning.		No details	This paper presents a model integrated into an ITS architecture that how it can be used to improve the

(*continued on next page*)

Table 11 (*continued*)

Reference	RQ 1	RQ 2	RQ 3	RQ 4	Study Overview
	Positive with empirical evaluation		Positive without empirical evaluation		tutoring feedback by anticipating student errors as long as this is pedagogically convenient .
(Ono et al., 2013)	No Evidence	Use feedback to help students in foreign language learning	Positive without empirical evaluation	Text Mining	This paper outlined how the new qualitative instant feedback was created on Moodle suitable for Japanese settings and showed the positive results on the use in the foreign language teaching in Japan.
(Efstathiou et al., 2018)	Positive with empirical evaluation	Use feedback to support students learning.	No evidence	No details	This study employed a quasi-experimental design to assess a computer based tool, which was intended to scaffold the task of designing experiments when using a virtual lab for the process of experimentation.
(Remolina et al., 2009)	No evidence	Use feedback to support students learning.	No evidence	Natural Language Processing	This paper describes a deployed simulation-based Intelligent Tutoring System (ITS) for training of Tactical Action Officers (TAOs) using artificial intelligence (AI) techniques.
(Xie & Li, 2018)	No evidence	Use feedback to help teacher	Positive without empirical evaluation	TF-IDF, EM Algorithm	This paper published a content-based recommendation template in big data and a clustering model based on the EM algorithm, to solve the problem of lack of personalized exercises and accurate feedback of teaching in online education.
(Wang et al., 2018)	Positive without empirical evaluation	Use feedback to help students with introductory programming exercises	No evidence	Tree edit distance, Abstract Syntax Trees (ASTs)	This paper introduces the "Search, Align, and Repair" data driven program repair framework to automate feedback generation for introductory programming exercises, wich goal is to develop an efficient, fully automated, and problem-agnostic technique for large or MOOC-scale introductory programming courses.
(Kaleeswaran et al., 2016)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	Comparision with desired solution	This paper presents a semi-supervised verified feedback generation to deal with both scale and variations in student submissions, while minimizing the instructor's efforts and ensuring feedback quality.
(Murad et al., 2018)	Positive with empirical evaluation	Use feedback to help students in foreign language learning	No evidence	Comparision with desired solution	This paper describes a multi-language karaoke application called SLIONS: Singing and listening to improve our natural speech. The main feature of SLIONS is the use of Automatic Speech Recognition (ASR) to provide students with personalized and granular feedback based on their singing pronunciation.
(Ohtsuki et al., 2016)	Positive with empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	Comparision with desired solution	This paper proposes an education support system ALECSS to train software developers by integrating several DevOps tools widely used for software development.
(Keuning et al., 2014)	Positive without empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation	Graphs	This paper presents a development a prototype of a programming tutor to help students with feedback and hints to progress towards a solution for an introductory imperative programming problem.
(Kakeshita & Ohta, 2016)	Positive without empirical evaluation	Use feedback to support programming learning	No evidence	Comparision with desired solution	This paper proposes the student feedback function for web-based programming education support tool pgtracer. The feedback function provides various information to the students .
(Jeremić et al., 2012)	Positive with empirical evaluation	Use feedback to support students learning.	No evidence	No details	This paper presents the design, implementation, and evaluation of a student model in DEPTHS (Design Pattern Teaching Help System), an intelligent tutoring system for learning software design patterns.
(Bryfczynski et al., 2013)	Positive with empirical evaluation	Use feedback to support data structures learning	Positive without empirical evaluation	No details	This paper describes a novel intelligent tutoring system called beSocratic, which targets question types that allow students to respond with free-form input but are able to be automatically evaluated and analyzed to help teachers refine their activities for improved future performances.
(Yu, 2016)	Positive with empirical evaluation	Use feedback to support students learning.	Positive with empirical evaluation	Dashboard	This paper presents a beneficial two-channel mechanism for a classroom feedback system in a digital classroom environment.
(Rahman et al., 2016)	Positive without empirical evaluation	Use feedback to support students' learning.	No evidence	knowledge databases	In this paper, the architecture for an agent-based ITS has been proposed which applied the multi agent concept.
(Bodily et al., 2018Bodily et al., 2018)	Positive with empirical evaluation	Use feedback to support students learning.	No evidence	Dashboard	This paper presents the design, development, and implementation of a student-focused Learning Analysis Panels (LADs) provide real-time feedback, recommendations and/or visualizations to students in order to support student reflection and awareness of knowledge in online environments.
(Ying & Hong, 2011)	No evidence	Use feedback to support SQL learning.	Positive without empirical evaluation	Comparision with desired solution	In this study was developed the e-learning system that has detailed feedback whether the program of practice were correct or wrong.
(Zhou et al., 2018)				No details	

(continued on next page)

Table 11 (*continued*)

Reference	RQ 1	RQ 2	RQ 3	RQ 4	Study Overview
	Positive without empirical evaluation	Use feedback to support programming learning	Positive without empirical evaluation		This paper presents a development of a OJ (Online Judge) system, a web software for compilation, execution and evaluation of programs submitted by users; and applied in the C Programming language course helped to improve students' programming skills.
(Mitrovic et al., 2011Mitrovic et al., 2011)	Positive with empirical evaluation	Use feedback to support thermodynamics learning	No evidence	Comparision with desired solution	This paper presents the project and an evaluation of the ThermoTutor, an Intelligent Tutoring System (ITS) that teaches thermodynamic cycles in closed systems, analyzes and provides appropriate feedback, and can progress through the material at their own pace.
(Hammadi et al., 2012)	Positive without empirical evaluation	Use feedback to support arabic handwriting learning.	No evidence	graph matching algorithm	In this paper, was proposed an Arabic handwriting education system with automatic errors detection.
(Alencar & Netto, 2014)	No evidence	Use feedback to support students learning.	No evidence	No details	This paper presented the development of a TUCUMÃ intelligent agent that is integrated with Moodle Virtual Learning Environment in order to monitor the activities of students in distance learning courses, as well as answer questions about the course.
(Jugo et al., 2014Jugo et al., 2014)	No evidence	Use feedback to support students learning.	No evidence	Data Mining	This paper presents a model of an adaptive intelligent tutoring systems (ITS) paired with an integrated a educational data mining (EDM) tool designed for educators and non-experts in Data Mining.
(Corrigan et al., 2015)	Positive with empirical evaluation	Use feedback to support students learning.	Positive without empirical evaluation	Predictive analytics	In this paper was described a method of predicting students outcomes in a first year University module or subject based on feature extraction from VLE access patterns, and the impact of feeding these predictions directly back to students on a weekly basis during semester.
(Birch et al., 2016)	Positive without empirical evaluation	Use feedback to support students learning.	No evidence	Comparision with desired solution	This paper describes a fast and accurate fully automated fault localisation tool for C programs and demonstrate its application to a corpus of student programs.
(Smithies et al., 2010)	Positive without empirical evaluation	Use feedback to support students learning.	No evidence	Natural Language Processing	This study shows the development of Version 1 of a Web-based service, which is designed to help learners monitor their conceptual development.
(Jin, 2017)	Positive without empirical evaluation	Use feedback to support students learning.	No evidence	Dashboard	This paper presents a development a visualization tool to motivate learners to participate actively in collaborative online learning communities and examine its effects on online participation, perceived learning, perceived satisfaction, team project performance, and usability.
(Weyten et al., 2010)	Positive with empirical evaluation	Use feedback to support Circuit Analysis learning.	No evidence	No details	In this paper, a new Web-based system for training students in the framework of teaching electric circuit theory and electronics was presented.
(Toshniwal et al., 2015)	Positive with empirical evaluation	Use feedback to support students learning.	No evidence	Mobile Sensors	This paper proposes VibRein to enrich the student interaction with multimedia learning content by making use of different sensors that are available on a mobile device to provides an assistive mechanism that keeps track of the user attention using the device camera, and uses haptic feedback to recapture attention.
(Weyten et al., 2008)	Positive with empirical evaluation	Use feedback to support Circuit Analysis learning.	Positive with empirical evaluation	No details	In this paper, a new Web-based system for training students in the framework of teaching electric circuit theory and electronics was presented. The main benefit is that the system approximates a well-established method of providing practical (homework) assignments enhanced by immediate individualized feedback provided by a private tutor.
(Dzikovska et al., 2014)	No evidence	Use feedback to support Basic Electricity and Electronics learning	No evidence	Domain reasoner, Natural Language Processing	In this paper the problem of developing an effective curriculum based on the method of conceptual change within the context of an Intelligent Tutor System (ITS) for a sub-topic of physics, electricity and electrical circuits was addressed.
(Salazar et al., 2017)	Positive without empirical evaluation	Use feedback to support students learning	Positive without empirical evaluation	Ontology	This paper presents the design and development of a multi-agent model for the assessment and diagnosis of failures which seeks to discover the shortcomings in learning from the virtual assessment process.
(Dutchuk et al., 2009)	No evidence	Use feedback to support students learning	No evidence	Comparision with desired solution	This paper describes a research project in progress of developing a Multi-Agent System-based educational game QuizMASter for e-learning that would help students learn their course material through friendly competition.

References

Akçapınar, G. (2015). How automated feedback through text mining changes plagiaristic behavior in online assignments. *Computers & Education, 87*, 123–130.

Alemán, J. L. F., Palmer-Brown, D., & Draganova, C. (2010). Evaluating student response driven feedback in a programming course. In *2010 10th IEEE international conference on advanced learning technologies* (pp. 279–283). IEEE.

Alencar, M., & Netto, J. F. (2014). Tutor collaborator using multi-agent system. In *International conference on collaboration technologies* (pp. 153–159). Springer.

Al-Hamad, B., & Mohieldin, T. (2013). E-assessment as a tool to augment face-to-face teaching and learning environment. In *2013 fourth international conference on e-learning" best practices in management, design and development of e-courses: Standards of excellence and creativity"* (pp. 348–359). IEEE.

Ali, L., Asadi, M., Gašević, D., Jovanović, J., & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education, 62*, 130–148.

Arends, H., Keuning, H., Heeren, B., & Jeuring, J. (2017). An intelligent tutor to learn the evaluation of microcontroller i/o programming expressions. In *Proceedings of the 17th Koli calling international conference on computing education research* (pp. 2–9). ACM.

Baneres, D., Clarisó, R., Jorba, J., & Serra, M. (2014). Experiences in digital circuit design courses: A self-study platform for learning support. *IEEE Transactions on Learning Technologies, 7*, 360–374.

Barbosa, G., Camelo, R., Cavalcanti, A. P., Miranda, P., Mello, R. F., Kovanović, V., & Gašević, D. (2020). Towards automatic cross-language classification of cognitive presence in online discussions. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 605–614).

Becheikh, N., Landry, R., & Amara, N. (2006). Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993–2003. *Technovation, 26*, 644–664.

Belcadhi, L. C. (2016). Personalized feedback for self assessment in lifelong learning environments based on semantic web. *Computers in Human Behavior, 55*, 562–570.

Bhatia, S., Kohli, P., & Singh, R. (2018). Neuro-symbolic program corrector for introductory programming assignments. In *2018 IEEE/ACM 40th international conference on software engineering (ICSE)* (pp. 60–70). IEEE.

Birch, G., Fischer, B., & Poppleton, M. (2016). Using fast model-based fault localisation to aid students in self-guided program repair and to improve assessment. In *Proceedings of the 2016 ACM conference on innovation and technology in computer science education* (pp. 168–173). ACM.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*, 7–74.

Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. *The Journal of the Learning Sciences, 23*, 561–599.

Bodily, R., Ikahihifo, T. K., Mackley, B., & Graham, C. R. (2018). The design, development, and implementation of student-facing learning analytics dashboards. *Journal of Computing in Higher Education, 30*, 572–598.

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education, 38*, 698–712.

Brookhart, S. M. (2017). *How to give effective feedback to your students*. ASCD.

Bryfczynski, S., Pargas, R. P., Cooper, M. M., Klymkowsky, M., & Dean, B. C. (2013). Teaching data structures with BeSocratic. In *Proceedings of the 18th ACM conference on Innovation and technology in computer science education* (pp. 105–110). ACM.

Burke, D. (2009). Strategies for using feedback students bring to higher education. *Assessment & Evaluation in Higher Education, 34*, 41–50.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245–281. https://doi.org/10.3102/00346543065003245.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*, 1315–1325.

Cavalcanti, A. P., de Mello, R. F. L., Rolim, V., André, M., Freitas, F., & Gaševic, D. (2019). An analysis of the use of good feedback practices in online learning courses. In *2019 IEEE 19th international conference on advanced learning technologies (ICALT)* (Vol. 2161, pp. 153–157). IEEE.

Cavalcanti, A. P., Diego, A., Mello, R. F., Mangaroska, K., Nascimento, A., Freitas, F., & Gašević, D. (2020). How good is my feedback? A content analysis of written feedback. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 428–437).

Charleer, S., Klerkx, J., & Duval, E. (2014). Learning dashboards. *Journal of Learning Analytics, 1*, 199–202.

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of computers & education. *Computers & Education, 151*, 103855.

Chen, X., Zou, D., & Xie, H. (2020b). Fifty years of british journal of educational technology: A topic modeling based bibliometric perspective. *British Journal of Educational Technology, 51*, 692–708.

Corrigan, O., Smeaton, A. F., Glynn, M., & Smyth, S. (2015). Using educational analytics to improve test performance. In *Design for teaching and learning in a networked world* (pp. 42–55). Springer.

D'antoni, L., Kini, D., Alur, R., Gulwani, S., Viswanathan, M., & Hartmann, B. (2015). How can automatic feedback help students construct automata? *ACM Transactions on Computer-Human Interaction (TOCHI), 22*, 9.

Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G.-J. (2017). Follow the successful crowd: Raising mooc completion rates through social comparison at scale. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 454–463). ACM.

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2019). What makes for effective feedback: Staff and student perspectives. *Assessment & Evaluation in Higher Education, 44*, 25–36.

Demaidi, M. N., Gaber, M. M., & Filer, N. (2018). Ontopefege: Ontology-based personalized feedback generator. *IEEE Access, 6*, 31644–31664.

Dermeval, D., Vilela, J., Bittencourt, I. I., Castro, J., Isotani, S., Brito, P., & Silva, A. (2016). Applications of ontologies in requirements engineering: A systematic review of the literature. *Requirements Engineering, 21*, 405–437.

Dillenbourg, P., Nussbaum, M., Dimitriadis, Y., & Roschelle, J. (2013). Design for classroom orchestration. *Computers & Education, 69*, 485–492.

Dutchuk, M., Muhammadi, K. A., & Lin, F. (2009). Quizmaster-a multi-agent game-style learning activity. In *International conference on technologies for E-learning and digital entertainment* (pp. 263–272). Springer.

Dzikovska, M., Steinhauser, N., Farrow, E., Moore, J., & Campbell, G. (2014). Beetle ii: Deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *International Journal of Artificial Intelligence in Education, 24*, 284–332.

Efstathiou, C., Hovardas, T., Xenofontos, N. A., Zacharia, Z. C., deJong, T., Anjewierden, A., & van Riesen, S. A. (2018). Providing guidance in virtual lab experimentation: The case of an experiment design tool. *Educational Technology Research and Development, 66*, 767–791.

Er, E., Dimitriadis, Y., & Gašević, D. (2020). Collaborative peer feedback and learning analytics: Theory-oriented design for supporting class-wide interventions. *Assessment & Evaluation in Higher Education*, 1–22.

Farrow, E., Moore, J., & Gašević, D. (2019). Analysing discussion forum data: A replication study avoiding data contamination. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 170–179).

Fast, E., Lee, C., Aiken, A., Bernstein, M. S., Koller, D., & Smith, E. (2013). Crowd-scale interactive formal reasoning and analytics. In *Proceedings of the 26th annual ACM symposium on User interface software and technology* (pp. 363–372). ACM.

Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9*, Article e1332.

Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior, 67*, 207–220.

Gašević, D., Tsai, Y.-S., Dawson, S., & Pardo, A. (2019). How do we start? An approach to learning analytics adoption in higher education. *The International Journal of Information and Learning Technology, 36*, 342–353. https://doi.org/10.1108/IJILT-02-2019-0024.

Gulwani, S., Radiček, I., & Zuleger, F. (2014). Feedback generation for performance problems in introductory programming assignments. In *Proceedings of the 22nd ACM SIGSOFT international symposium on foundations of software engineering* (pp. 41–51). ACM.

Hammadi, M., Bezine, H., Njah, S., & Alimi, A. M. (2012). Towards an educational tool for Arabic handwriting learning. In *International conference on education and e-learning innovations* (pp. 1–6). IEEE.

Harvey, L. (2003). Student feedback. *Quality in Higher Education, 9*, 3–20.

Hattie, J., & Gan, M. (2011). Instruction based on feedback. In *Handbook of research on learning and instruction* (pp. 263–285). Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112.

Helminen, J., Ihantola, P., Karavirta, V., & Malmi, L. (2012). How do students solve parsons programming problems?: An analysis of interaction traces. In *Proceedings of the ninth annual international conference on International computing education research* (pp. 119–126). ACM.

Helminen, J., & Malmi, L. (2010). Jype-a program visualization and programming exercise tool for Python. In *Proceedings of the 5th international symposium on Software visualization* (pp. 153–162). ACM.

Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E., & Mahoney, P. (2019). Conditions that enable effective feedback. *Higher Education Research and Development, 38*, 1401–1416.

Hentea, M., Shea, M. J., & Pennington, L. (2003). A perspective on fulfilling the expectations of distance education. In *Proceedings of the 4th conference on Information technology curriculum* (pp. 160–167).

Hernandes, E., Zamboni, A., Fabbri, S., & Thommazo, A. D. (2012). Using gqm and tam to evaluate start-a tool that supports systematic review. *CLEI Electronic Journal, 15*, 3–3.

Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., & Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: A case study. *Computers in Human Behavior, 47*, 68–80.

Ihantola, P., Ahoniemi, T., Karavirta, V., & Seppälä, O. (2010). Review of recent systems for automatic assessment of programming assignments. In *Proceedings of the 10th Koli calling international conference on computing education research* (pp. 86–93).

Jeremić, Z., Jovanović, J., & Gašević, D. (2012). Student modeling and assessment in intelligent tutoring of software patterns. *Expert Systems with Applications, 39*, 210–222.

Jin, S.-H. (2017). Using visualization to motivate student participation in collaborative online learning environments. *Journal of Educational Technology & Society, 20*, 51–62.

Joksimović, S., Gašević, D., Loughin, T. M., Kovanović, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers & Education, 87*, 204–217. https://doi.org/10.1016/j.compedu.2015.07.002.

Jones, C., Ramanau, R., Cross, S., & Healing, G. (2010). Net generation or digital natives: Is there a distinct new generation entering university? *Computers in Education, 54*, 722–732.

Jugo, I., Kovačić, B., & Slavuj, V. (2014). Using data mining for learning path recommendation and visualization in an intelligent tutoring system. In *2014 37th international convention on Information and communication technology, electronics and microelectronics (MIPRO)* (pp. 924–928). IEEE.

Kakeshita, T., & Ohta, K. (2016). Student feedback function for web-based programming education support tool pgtracer. In *2016 5th IIAI international congress on advanced applied informatics (IIAI-AAI)* (pp. 322–327). IEEE.

Kaleeswaran, S., Santhiar, A., Kanade, A., & Gulwani, S. (2016). Semi-supervised verified feedback generation. In *Proceedings of the 2016 24th ACM SIGSOFT international symposium on foundations of software engineering* (pp. 739–750). ACM.

Karavirta, V., Helminen, J., & Ihantola, P. (2012). A mobile learning application for parsons problems with automatic feedback. In *Proceedings of the 12th Koli calling international conference on computing education research* (pp. 11–18). ACM.

Kebodeaux, K., Field, M., & Hammond, T. (2011). Defining precise measurements with sketched annotations. In *Proceedings of the eighth eurographics symposium on sketch-based interfaces and modeling* (pp. 79–86). ACM.

Keele, S., et al. (2007). *Guidelines for performing systematic literature reviews in software engineering. Technical Report Technical report, Ver. 2.3 EBSE Technical Report*. EBSE.

Keuning, H., Heeren, B., & Jeuring, J. (2014). Strategy-based feedback in a programming tutor. In *Proceedings of the computer science education research conference* (pp. 43–54). ACM.

Keuning, H., Jeuring, J., & Heeren, B. (2016). Towards a systematic review of automated feedback generation for programming exercises. In *Proceedings of the 2016 ACM conference on innovation and technology in computer science education* (pp. 41–46).

Keuning, H., Jeuring, J., & Heeren, B. (2018). A systematic literature review of automated feedback generation for programming exercises. *ACM Transactions on Computing Education (TOCE), 19*, 1–43.

Khan, I., & Pardo, A. (2016). Data2u: Scalable real time student feedback in active learning environments. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 249–253). ACM.

Kitchenham, B. (2004). *Procedures for performing systematic reviews* (Vol. 33, pp. 1–26). Keele, UK: Keele University.

Krusche, S., & Seitz, A. (2018). Artemis: An automatic assessment management system for interactive learning. In *Proceedings of the 49th ACM technical symposium on computer science education* (pp. 284–289). ACM.

Lan, A. S., Vats, D., Waters, A. E., & Baraniuk, R. G. (2015). Mathematical language processing: Automatic grading and feedback for open response mathematical questions. In *Proceedings of the second (2015) ACM conference on learning@ scale* (pp. 167–176). ACM.

LAPES. (2014). Start-state of the art through systematic review tool. Disponível em http://lapes.dc.ufscar.br/tools/start_tool. Acesso em: 02 abril 2019.

Lee, Y., & Choi, J. (2011). A review of online course dropout research: Implications for practice and future research. *Educational Technology Research & Development, 59*, 593–618.

Lee, J. S., & Kim, J. H. (2009). Pen-based intelligent tutor for simultaneous equations. In *2009 Chinese conference on pattern recognition* (pp. 1–5). IEEE.

Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). *What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course*. Learning and Instruction.

Lodder, J., Heeren, B., & Jeuring, J. (2017). Generating hints and feedback for hilbert-style axiomatic proofs. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 387–392). ACM.

Manoharan, S. (2016). Personalized assessment as a means to mitigate plagiarism. *IEEE Transactions on Education, 60*, 112–119.

Marin, V. J., Pereira, T., Sridharan, S., & Rivero, C. R. (2017). Automated personalized feedback in introductory java programming moocs. In *2017 IEEE 33rd international conference on data engineering (ICDE)* (pp. 1259–1270). IEEE.

Martin, M., Alvarez, A., Ruiz, S., Fernandez-Castro, I., & Urretavizcaya, M. (2009). Helping teachers to track student evolution in a b-learning environment. In *2009 ninth IEEE international conference on advanced learning technologies* (pp. 342–346). IEEE.

Matcha, W., Gasevic, D., Pardo, A., et al. (2019). *A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective*. IEEE Transactions on Learning Technologies.

Mitrovic, A., Williamson, C., Bebbington, A., Mathews, M., Suraweera, P., Martin, B., Thomson, D., & Holland, J. (2011). Thermo-tutor: An intelligent tutoring system for thermodynamics. In *2011 IEEE global engineering education conference (EDUCON)* (pp. 378–385). IEEE.

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education, 42*, 266–288.

Murad, D., Wang, R., Turnbull, D., & Wang, Y. (2018). Slions: A karaoke application to enhance foreign language learning. In *2018 ACM multimedia conference on multimedia conference* (pp. 1679–1687). ACM.

Mutch, A. (2003). Exploring the practice of feedback to students. *Active Learning in Higher Education, 4*, 24–38.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*, 199–218.

Ohtsuki, M., Ohta, K., & Kakeshita, T. (2016). Software engineer education support system ALECSS utilizing DevOps tools. In *Proceedings of the 18th international conference on information integration and web-based applications and services* (pp. 209–213). ACM.

Ono, Y., Ishihara, M., & Yamashiro, M. (2013). Preliminary construction of instant qualitative feedback system in foreign language teaching. In *2013 second IIAI international conference on advanced applied informatics* (pp. 178–182). IEEE.

Ouadoud, M., Nejjari, A., Chkouri, M. Y., & El-Kadiri, K. E. (2017). Learning management system and the underlying learning theories. In *Proceedings of the mediterranean symposium on smart city applications* (pp. 732–744). Springer.

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education, 43*, 428–438.

Pardo, A., Bartimote, K., Shum, S. B., Dawson, S., Gao, J., Gašević, D., Leichtweis, S., Liu, D., Martínez-Maldonado, R., Mirriahi, N., et al. (2018). Ontask: Delivering data-informed, personalized learning support actions. *Journal of Learning Analytics, 5*, 235–249.

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*, 128–138.

Parikh, A., McReelis, K., & Hodges, B. (2001). Student feedback in problem based learning: A survey of 103 final year students across five Ontario medical schools. *Medical Education, 35*, 632–636.

Perera, R., & Nand, P. (2017). Recent advances in Natural Language generation: A survey and classification of the empirical literature. *Computing and Informatics, 36*, 1–32.

Polson, M. C., & Richardson, J. J. (2013). *Foundations of intelligent tutoring systems*. Psychology Press.

Price, M., Handley, K., Millar, J., & O'donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*, 277–289.

Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 148–157).

Rahman, A. A., Abdullah, M., & Alias, S. H. (2016). The architecture of agent-based intelligent tutoring system for the learning of software engineering function point metrics. In *2016 2nd international symposium on agent, multi-agent systems and robotics (ISAMSR)* (pp. 139–144). IEEE.

Reiser, R. A., & Dempsey, J. V. (2012). *Trends and issues in instructional design and technology*. MA: Pearson Boston.

Remolina, E., Ramachandran, S., Stottler, R., & Davis, A. (2009). Rehearsing naval tactical situations using simulated teammates and an automated tutor. *IEEE Transactions on Learning Technologies, 2*, 148–156.

Riofrío-Luzcando, D., Ramirez, J., & Berrocal-Lobo, M. (2017). Predicting student actions in a procedural training environment. *IEEE Transactions on Learning Technologies, 10*, 463–474.

Robison, J., McQuiggan, S., & Lester, J. (2009). Evaluating the consequences of affective feedback in intelligent tutoring systems. In *2009 3rd international conference on affective computing and intelligent interaction and workshops* (pp. 1–6). IEEE.

Rogers, P. L. (2000). Barriers to adopting emerging technologies in education. *Journal of Educational Computing Research, 22*, 455–472.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*, e1355.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.

Salazar, O., Álvarez, S., & Ovalle, D. (2017). Multi-agent model for failure assessment and diagnosis in teaching-learning processes. In *International conference on practical applications of agents and multi-agent systems* (pp. 398–408). Springer.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*, 30–41.

Sheridan, R. (2006). Reducing the online instructor's workload. *Educause Quarterly, 29*, 65.

Smithies, A., Braidman, I., Berlanga, A., Haley, D., & Wild, F. (2010). Using language technologies to support individual formative feedback. In *The 9th European Conference on e-Learning (ECEL 2010)*.

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior, 28*, 1738–1747.

Tempelaar, D., Nguyen, Q., & Rienties, B. (2020). Learning feedback based on dispositional learning analytics. In *Machine learning paradigms* (pp. 69–89). Springer.

Tenório, T., Bittencourt, I. I., Isotani, S., & Silva, A. P. (2016). Does peer assessment in on-line learning environments work? A systematic review of the literature. *Computers in Human Behavior, 64*, 94–107.

Toshniwal, S., Dey, P., Rajput, N., & Srivastava, S. (2015). Vibrein: An engaging and assistive mobile learning companion for students with intellectual disabilities. In *Proceedings of the annual meeting of the Australian special interest group for computer human interaction* (pp. 20–28). ACM.

Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). Polycafe—automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning, 9*, 127–156.

Tsai, Y.-S., Mello, R. F., Jovanović, J., & Gašević, D. (2021). Student appreciation of data-driven feedback: A pilot study on ontask. In *LAK21: 11th international learning analytics and knowledge conference* (pp. 511–517).

Tseng, S.-C., & Tsai, C.-C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education, 49*, 1161–1174.

Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S., & Ragos, O. (2020). Implementing AutoML in educational data mining for prediction tasks. *Applied Sciences, 10*, 90.

Usener, C. A. (2015). EASy-DSBuilder: Automated assessment of tree data structures in computer science teaching. In *Proceedings of the 30th annual ACM symposium on applied computing* (pp. 220–226). ACM.

Utomo, A. Y., & Santoso, H. B. (2015). Development of gamification-enriched pedagogical agent for e-learning system based on community of inquiry. In *Proceedings of the international HCI and UX conference in Indonesia* (pp. 1–9). ACM.

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: An overview and future research opportunities. *Personal and Ubiquitous Computing, 18*, 1499–1514.

Wang, K., Singh, R., & Su, Z. (2018). Search, align, and repair: Data-driven feedback generation for introductory programming exercises. In *Proceedings of the 39th ACM SIGPLAN conference on programming language design and implementation* (pp. 481–495). ACM.

Wasik, S., Antczak, M., Badura, J., Laskowski, A., & Sternal, T. (2018). A survey on online judge systems and their applications. *ACM Computing Surveys (CSUR), 51*, 1–34.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education, 31*, 379–394.

Westera, W. (2015). On the cybernetic arrangement of feedback in serious games: A systems-theoretical perspective. *Education and Information Technologies, 20*, 57–73.

Weyten, L., Rombouts, P., Catteau, B., & De Bock, M. (2010). Validation of symbolic expressions in circuit analysis e-learning. *IEEE Transactions on Education, 54*, 564–568.

Weyten, L., Rombouts, P., & De Maeyer, J. (2008). Web-based trainer for electrical circuit analysis. *IEEE Transactions on Education, 52*, 185–189.

Whitelock, D., Twiner, A., Richardson, J. T., Field, D., & Pulman, S. (2015). Openessayist: A supply and demand learning analytics tool for drafting academic essays. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 208–212). ACM.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 203–211).

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics, 6*, 53–69.

Wohlin, C., Runeson, P., Host, M., Ohlsson, M., & Regnell, B. (2000). *Experimentation in software engineering-an introduction*. Doedrecht the Netherlands: kluwer academic publishers.

Wong, S. H. S., Taylor, J. E., & Beaumont, A. J. (2012). Enhancing student learning experience through a novel electronic coursework assessment and feedback management system. In *Proceedings for EE2012 - innovation, practice and research in engineering education*. GBR: Higher Education Academy.

Xie, X., & Li, X. (2018). Research on personalized exercises and teaching feedback based on big data. In *Proceedings of the 3rd international conference on intelligent information processing* (pp. 166–171). ACM.

Ying, M.-H., & Hong, Y. (2011). The development of an online SQL learning system with automatic checking mechanism. In *The 7th international conference on networked computing and advanced information management* (pp. 346–351). IEEE.

Ying, M.-H., Yang, K.-T., & Deng, G.-H. (2012). Development of a multiplayer online game-based learning system based on ARCS motivation model. In *2012 sixth international conference on genetic and evolutionary computing* (pp. 589–594). IEEE.

Ypsilandis, G. (2002). Feedback in distance education. *Computer Assisted Language Learning, 15*, 167–181.

Yu, Y.-C. (2016). Teaching with a dual-channel classroom feedback system in the digital classroom environment. *IEEE Transactions on Learning Technologies, 10*, 391–402.

Zhang, B., & Jia, J. (2017). Evaluating an intelligent tutoring system for personalized math teaching. In *2017 international symposium on educational technology (ISET)* (pp. 126–130). IEEE.

Zhao, Y., & Cziko, G. A. (2001). Teacher adoption of technology: A perceptual control theory perspective. *Journal of Technology and Teacher Education, 9*, 5–30.

Zhou, W., Pan, Y., Zhou, Y., & Sun, G. (2018). The framework of a new online judge system for programming education. In *Proceedings of ACM turing celebration conference-China* (pp. 9–14). ACM.