THE DOUBLE-EDGED SWORD:
BENEFITS AND CHALLENGES THAT ARTIFICIAL INTELLIGENCE TOOLS CAN BRING TO EFL TEACHING AND LEARNING
THE DOUBLE-EDGED SWORD:

BENEFITS AND CHALLENGES THAT ARTIFICIAL INTELLIGENCE TOOLS CAN BRING TO EFL TEACHING AND LEARNING

ABSTRACT

AI tools are applications or programs that simulate human intelligence to assist people in automating tasks or solving problems. Thereby, they can be beneficial in EFL classes. Nonetheless, they could also present limitations that should be known in order to use them wisely. Considering this, the present qualitative review aims at reporting on the impacts of AI tools on EFL teaching and learning. More specifically, it seeks to shed light on the benefits and potential challenges associated with their use, according to the literature. The results of benefits of AI for teachers indicated that the highest volume of supporting articles was for the subcategory Teachers’ Experience and that Improved Skills and Accessibility had the lowest number of works. For learners, Students’ Experience and Improved Skills received more support and Digital Knowledge and Openness to Technology had the fewest studies. Regarding challenges of AI for educators, Implementation Challenges had the highest number of articles and Ethical Concerns and Quality received the lowest research coverage. In the case of learners, Quality and Implementation Challenges were principally supported and Lack of Variety and Personalized Learning had the fewest works. Lastly, what these results suggest for teaching and learning is further discussed.

Keywords:
Benefits, challenges, AI tools, EFL, teaching, learning.

RESUMEN

Las herramientas de IA son aplicaciones o programas que simulan la inteligencia humana para ayudar a las personas a automatizar tareas o resolver problemas. Por lo tanto, pueden ser beneficiosas en las clases de inglés como lengua extranjera. Sin embargo, también podrían presentar limitaciones que deberían ser conocidas para utilizarlas sabiamente. Considerando esto, la presente revisión cualitativa tiene como objetivo informar sobre los impactos de las herramientas de IA en la enseñanza y aprendizaje del inglés como lengua extranjera. Más específicamente, busca aclarar los beneficios y desafíos potenciales asociados con su uso, según la literatura. Los resultados sobre los beneficios de la IA para los profesores indicaron que el mayor volumen de artículos de apoyo fue para la subcategoría Experiencia de los Docentes y que Mejora de Habilidades y Accesibilidad tuvieron el número más bajo de trabajos. Para los aprendices, Experiencia de los Estudiantes y Mejora de Habilidades recibieron más respaldo y Conocimiento Digital y Apertura a la Tecnología tuvieron la menor cantidad de estudios. Con respecto a desafíos de la IA para los educadores, Desafíos de Implementación tuvieron el mayor número de artículos y Preocupaciones Éticas y Calidad recibieron la menor cobertura de investigación. En el caso de los aprendices, Calidad y Desafíos de Implementación fueron principalmente respaldados y Falta de Variedad y Aprendizaje Personalizado tuvieron la menor cantidad de trabajos. Por último, se discute lo que estos resultados sugieren para la enseñanza y el aprendizaje.

Palabras clave:
Beneficios, desafíos, herramientas de IA, inglés como lengua extranjera, enseñanza, aprendizaje.
INTRODUCTION

Curiosity inherent in humans has led them to wonder about the possibility of making machines think since ancient times (Boden, 2008). In other words, AI is quite far from being new. In actual fact, it has a long history that can be tracked. Thus, in 1950, the mathematician Alan Turing, developed his famous Turing Test to evaluate how capable a machine was to show intelligent behavior that could not be distinguishable from that of a human being (Toosi et al., 2021; Wooldridge, 2021). Some years later, in 1956, John McCarthy popularized the term “Artificial Intelligence” in his workshop The Dartmouth Summer Research Project on Artificial Intelligence at Dartmouth College (Abdelatif & Siddiqui, 2021). In this way, the word was officially born.

In the upcoming years, AI underwent some ups and downs. Nonetheless, nothing stopped it from evolving in different fields such as Linguistics, Psychology, Neuroscience, and of course Education (Toosi et al., 2021). As Abdelatif and Siddiqui (2021) claimed, in the last four decades, there has been too much study and debate about language education and technology. As a matter of fact, the significance of technological advancements increased the capabilities of the traditional CALL (Computer-Assisted Language Learning) and led to the ICALL framework (Intelligent Computer-Assisted Language Learning) which leverages AI technologies to enhance language education (Rebolledo & González, 2023).

As can be noted, AI has been evolving at a fast rate since the very first moment when it was conceived as the idea of making machines think (Boden, 2008; Mohammadkarimi, 2023). At present, it is still advancing, always trying to meet the constant demands of productivity, competitiveness, and creativity of a world undergoing rapid change (Rao, 2022). In light of this, it becomes paramount for both educators and students to develop a sound understanding of the benefits and challenges that AI tools might bring. In this way, they will be capable of using them wisely.

Considering this, the proposed literature review The Double-Edged Sword: Benefits and Challenges that Artificial Intelligence (AI) Tools can Bring to EFL Teaching and Learning intends to report on the impacts of AI on EFL education. Specifically, it seeks to shed light on how AI tools can support the teaching and learning process and to identify the potential challenges or limitations associated with their use in EFL teaching and learning, according to different sources of literature.

Lastly, it is worth mentioning that working on the creation of a literature review on the aforementioned topic could definitely be advantageous since according to previous search, the number of reviews on AI tools in the field of EFL education is scarce. This is definitely worrying considering the crucial role of AI in today’s world. Conversely, primary sources have become more numerous since 2020. It could be a possible consequence of the COVID-19 pandemic which forced people to isolate and therefore turn their attention to technology to keep working and/or studying (Matosina, 2023; Vera, 2023). This emerging amount of research is extremely significant since it could be taken as valuable raw material to develop a future literature review.

METHODOLOGY

To begin with, this literature review makes use of a qualitative research design. It is so because it is ideal to conduct an in-depth exploration of rich data so as to come up with a comprehensive picture of different research works (Ma, 2015). Besides, it is a semi-systematic review considering the available time for the project.

Regarding data collection, the academic search engine, Google Scholar, was employed in view of the previous experience that the author has with it and the fact that it is free. Moreover, the digital base DOAJ (Directory of Open Access Journals) was utilized due to its focus on promoting quality in research (Holder, 2022). The gathered sources of literature came from different journals such as English Language Teaching, The Journal of Language Teaching and Research, IEEE Transactions on Learning Technologies, Cogent Education, Trends in Higher Education, The Journal of English Language Teaching, The Journal of Language and Science, Cogent Education, International Journal of Technology and Educational Innovation, just to mention a few. Besides, the key terms that were taken into account to search for sources of literature were: Benefits and Challenges of AI Tools in EFL Education, Pros and Cons of AI in the EFL Classroom, The Impact of AI on EFL Teaching and Learning, and AI Tools in EFL Education. Furthermore, the inclusive and exclusive criteria that was considered for selecting appropriate articles is detailed in Table 1.

Table 1. Inclusive and Exclusive Criteria.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus of the article</td>
<td>Research works with only benefits, only challenges, or both of them about AI tools</td>
<td>Articles with neither benefits nor challenges of AI tools</td>
</tr>
<tr>
<td>EFL/ESL</td>
<td>Academic sources related to EFL teaching and/or learning</td>
<td>Articles about ESL teaching and/or learning</td>
</tr>
<tr>
<td>Ages</td>
<td>Works encompassing learners who are kids, teenagers, or adults</td>
<td>No exclusion criterion</td>
</tr>
<tr>
<td>Language</td>
<td>Research articles written in English</td>
<td>Works written in other languages</td>
</tr>
<tr>
<td>Research Design</td>
<td>Sources with a Quantitative, Qualitative, or Mixed-Methods design</td>
<td>No exclusion criteria</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Types of Sources</td>
<td>Published</td>
<td>Unpublished</td>
</tr>
<tr>
<td>Rosenthal (1994, as cited in Norris &amp; Ortega, 2006) argued that both types of sources should be considered to prevent bias in research. Notwithstanding, the review only used published works in view of the challenge of accessing unpublished sources.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary works</td>
<td>Secondary sources</td>
<td></td>
</tr>
<tr>
<td>Publication date</td>
<td>Studies published in the timeframe 2020-2024</td>
<td>Studies published before 2020 or after 2024</td>
</tr>
<tr>
<td>Articles whose research methods are clear and appropriate to meet their research objectives and whose data is analyzed rigorously.</td>
<td>Sources whose research methods are unclear and inappropriate to meet their research objectives and whose data is not analyzed rigorously.</td>
<td></td>
</tr>
</tbody>
</table>

It is worth mentioning that initially, 59 primary sources were selected for the work. Then, after considering the aforementioned criteria, those articles were reduced to 50. Eight works were excluded since they were secondary sources and one article was eliminated because it was not part of the timeframe 2020-2024. Afterwards, information of the 50 sources related to AI tools, benefits, and challenges of AI was gathered to be organized and analyzed. Lastly, conclusions are presented.

DEVELOPMENT

The content areas: **AI Tools in EFL Education, Benefits of AI Tools in EFL Teaching and Learning**, and **Challenges of AI Tools in EFL Teaching and Learning** will be addressed in this part. It is worth mentioning that the last two areas seek to respond the general research question: What is the impact of AI tools on EFL teaching and learning? And the specific questions: How can AI tools support the teaching and learning process, according to the collected literature? And What are the potential challenges and limitations associated with the use of AI tools in EFL teaching and learning, according to the gathered sources?

To start with, this first cluster intends to help readers understand the landscape of AI technologies through information about the most and the least commonly used tools in EFL contexts before tackling benefits and challenges. To accomplish this, AI Tools that were used by teachers and students in the 50 articles were selected. Next, the author registered them.

The most commonly used tool in EFL contexts by educators is ChatGPT (see annex 1). It was found in six works (Alenizi et al., 2023; Algaraady & Mahyoob, 2023; Chan & Hu, 2023; Denecke et al., 2023; Gültekin Talayhan & Babayiğit, 2023; Ulla et al., 2023) out of the initial 50 which corresponds to 12% of the total. On the other hand, the least commonly used AI technologies by teachers were Perplexity, British Council Website, Breaking News Website, Randall Cyber, PowerPoint, Facebook, Google Search, Google Scholar, and Google Translate. Each of them was encountered in only one work or 2% of the total. Based on this information, two things are clear. First, even though six articles for ChatGPT might not seem numerically significant, they position the tool as a frontrunner for teacher adoption in the dataset of 50 sources, possibly because it supports specific teaching needs as Ahmed (2023) and Ulla et al. (2023) stated. Nonetheless, more investigation is crucial to corroborate its importance among the other tools. Second, the variety observed could suggest educators’ interest in exploring distinct AI functionalities to support their work. In this regard, further research is also necessary to confirm this idea.

Regarding students, their most commonly used tool in EFL contexts is ChatGPT (see annex 2). It was found in 13 sources (Ahmed, 2023; Aljabr, 2023; Alrasheedi, 2023; Bok & Cho, 2023; Chan & Hu, 2023; Denecke et al., 2023; Han et al., 2023; Harunasari, 2023; Malik et al., 2023; Sol et al., 2024; Song & Song, 2023; Sotelo Muñoz et al., 2023; Xiao & Zhi, 2023) out of the 50 which is 26% of the initial works. Besides, Grammarly occupied the second place. It was encountered in nine studies (Alotaibi, 2023; Chang et al., 2021; Ginting et al., 2023; Malik et al., 2023; Marghany, 2023; Nazari et al., 2021; Phan, 2023; Sol et al., 2024; Wei et al., 2023) or 18% of the total. Conversely, the least commonly used tools by learners were AI Based Program, Falou, DeepL, Plot Generator, Lyra, Siri, Netflix, Fitbit, Google Maps, Loora, SmallTalk2Me, Copy.ai, Essay Writer, AI KAKU, Turnitin, Peppertype, Surfer SEO, Jasper, Gemini (formerly Bard), Vocabulary Learning Chatterbot, Wordtune, Write and Improve Tool, and Paraphrasing Tool. Each of them was present in only one article or 2% of the total. Considering this, it is thought-provoking to see that ChatGPT is also the most commonly used tool by students. Perhaps, it indicates that it is more versatile and user-friendly than other technologies (Chan & Hu, 2023) or simply that it is more well-known in AI education. In fact, ChatGPT achieved 1 million users in only five days, which took Facebook 10 months and Netflix three and one-half years according to OpenAI (Duarte, 2024). Whatever the reason could be, what is clear is that more research is needed to accurately trace a possible cause. Furthermore, it should be noticed that ChatGPT was spotted in more articles where it was used by students rather than by teachers. It might suggest a higher interest on it by learners which should be confirmed with more studies. Moreover, even though Grammarly appeared in less sources than ChatGPT, the results underscore its significance for learners, possibly due to its effectiveness.
to give explanations, detect grammatical errors, and enhance writer confidence (Alotaibi, 2023). Lastly, students use a wider variety of AI tools compared to teachers (see annexes 1 and 2). It can indicate that learners are exploring a broader range of functionalities of AI. In relation to this, additional research could also be beneficial in this direction.

The present cluster intends to answer the general question: What is the impact of AI tools on EFL teaching and learning? And the first specific question: How can AI tools support the teaching and learning process, according to the collected literature? In order to achieve this, information of benefits on AI tools in EFL education was selected. Afterwards, codes were assigned to the different chunks of data. Then, those codes were included into subcategories that were created inside the overarching categories Benefits of AI tools in EFL Teaching and Learning (see annexes 3 and 4). Next, the number of articles with benefits in each subcategory is illustrated through Figures 1 and 2 for teaching and learning, respectively.

As can be observed, the subcategory with the highest volume of supporting articles is Teachers’ Experience. In actual fact, it contains codes found in 13 research articles or 26% of the 50 primary works. The benefits of this subcategory which were found in more sources were “satisfying” and “effective” in seven works each of them. This equals to 14% of the total. Thus, Hamied et al. (2022) investigated teachers’ perceptions of four educators on employing AI in the classroom and found that all of them were positive about AI tools due to their help in each of their classes. Similarly, Kohnke et al. (2023) explored university instructors’ attitudes towards these technologies and encountered that they perceived them as satisfying for their work. For example, Oliver, one of the participants, highlighted that AI is beneficial to remind him of important tasks. Besides, Mohammadkarimi (2023) examined educators’ perceptions on academic dishonesty in writing with regard of the use of AI. As a result, the author found that educators agreed on the harmful influence of AI tools for academic honesty. Nevertheless, they also presented positive attitudes towards AI tools based on their personal experience with them. In relation to the code “effectiveness,” Algaraady and Mahyoob (2023) investigated how capable ChatGPT was to spot students’ writing errors in comparison with instructors. The results showed that the AI tool could detect surface-level issues. Nonetheless, it could not identify more complicated ones which could be spotted by educators. In spite of this limitation, the researchers concluded that ChatGPT is an effective tool that instructors could use for error correction of students’ work. Likewise, Ulla et al. (2023) did an exploration of teachers’ perspectives of ChatGPT in which they found its effectiveness for lesson preparation and activity creation. Lastly, Zitouni (2022) carried out a study to unveil the significance of AI tools in online learning and its different pros and cons. The author’s findings revealed that AI technologies are effective to facilitate and improve the teaching process.

In addition, due to the number of articles of Teachers’ Experience, it is clear that this subcategory is a relevant area of research about AI in education. Besides, the focus on the codes “satisfying” and “effective” suggests a positive instructor perception on AI technologies. Nevertheless, the mention of the limitation of ChatGPT to detect complex errors underscores the need of further research in regard of AI capabilities.

Finally, Figure 1 illustrates that the subcategories with the lowest numbers of sources are Improved Skills and Accessibility. Each of them has benefits encountered in two articles or 4% of the primary works and in three sources or 6%, respectively. This information shows that these subcategories are not necessarily less significant, but that they have not been extensively explored. Therefore, they might be valuable topics of research.
Figure 2 shows the 10 subcategories of Benefits of AI Tools in EFL Learning: Digital Knowledge, Students’ Experience, Openness to Technology, Usefulness, Learning Support, Improved Skills, Accessibility, Personalized Learning, Emotional Impact, and Quality. Digital Knowledge refers to the understanding and ability to utilize AI tools effectively. Students’ Experience involves benefits of AI that learners have encountered during their learning process. Openness to Technology is the willingness to learn about and adopt new tools in learning. Usefulness is the degree to which AI technologies are helpful or beneficial in achieving goals. Learning Support refers to how AI tools can assist students in their learning. The subcategory Improved Skills is about the enhancement of existing abilities or the development of new skills with AI tools. Accessibility refers to the extent to which AI technologies can be used and benefit all students. Personalized Learning refers to tailored content and educational experiences that satisfy needs and styles of every learner. Emotional Impact is how AI tools may influence a student’s emotional state while learning. Finally, Quality makes reference to the overall performance and functionality of an AI tool.

According to Figure 2, the subcategories predominantly supported by literature are Students’ Experience and Improved Skills. They have codes encountered in 41 research articles or 82% of the 50 sources and in 40 works, or 80% of the total, respectively. The most significant codes of the subcategory Students’ Experience are “satisfying,” found in 22 works or 44% of the 50 sources and “helpful” encountered in 33 works or 66%. In regard of the first code, Bok and Cho (2023) investigated strategies to incorporate ChatGPT into a writing class. As part of this work, the researcher discovered that 13 students, out of the 16 participants, saw ChatGPT as helpful for their writing process.

In view of the heavy number of sources for Students’ Experience, it is crystal clear that this subcategory is a key area of research of AI in education. Moreover, the emphasis on “satisfying” and “helpful” suggests that AI tools are also positively perceived by learners as they are by educators. What is thought-provoking up to this point is that two subcategories that deal with human experience, Teachers’ Experience and Students’ Experience, have been the ones with the largest numbers of supporting sources and that the code “satisfying” has appeared in both of them. This might highlight a growing emphasis on a human-centered design in education which could suggest future directions of research on human experience with AI.

Regarding the subcategory Improved Skills, its most relevant code was “improve writing” encountered in 27 sources which is 54% of the analyzed works. In relation to this, Song and Song (2023) conducted a mixed-methods study aiming at evaluating the impact of ChatGPT on Chinese students’ writing skill. The results showed that the learners who used the AI tool exhibited a superior improvement of their writing in comparison to the students who did not use it. Similarly, Wei et al. (2023) investigated the effectiveness of Grammarly on Chinese learners’ writing skill for 12 weeks. As a consequence, they found that the employment of the tool led to a higher performance of the experimental students in contrast with the control learners.

Considering the numerous articles that support Improved Skills, this subcategory also appears as paramount for research on AI tools in education. Besides, the significance of the code “improve writing” indicates that this skill could have priority in education, perhaps due to its relation with critical thinking and effective communication. Besides, the concentration of literature on writing underscores the need of broader research on the other skills.

Moreover, Figure 2 indicates that Digital Knowledge and Openness to Technology are the subcategories with the fewest articles. They have benefits found in 11 sources or 22% of the total and in eight works or 16%, respectively. These results could also indicate a need of more investigation in regard of those areas that possibly have not been comprehensively examined.

Lastly, it is worth mentioning that based on Figures 1 and 2, it is easy to notice that more research has been done for Benefits of AI Tools in EFL Learning rather than for Benefits of AI Tools in EFL Teaching. This clearly suggests that more studies should be conducted in that direction.
The present group wishes to answer the general question: What is the impact of AI tools on EFL teaching and learning? And the second specific question: What are the potential challenges or limitations associated with the use of AI tools in EFL teaching and learning, according to the gathered sources? To accomplish this, information of challenges on AI tools in EFL education was chosen. Subsequently, codes were allocated to data segments. Later, the codes were included into subcategories inside Challenges of AI tools in EFL Teaching and Learning (see annexes 5 and 6). Next, Figures 3 and 4 show the distribution of sources with challenges within each subcategory for teaching and learning, respectively. Finally, articles’ recommendations on how to address the main limitations of AI tools for teachers and learners are provided.

As can be seen, the subcategory Implementation Challenges presented a small portion of studies, it underscores a dual focus. First, the necessity of educators to receive training programs to deal with AI technologies (Ghoneim & Elghotmy, 2021). Second, that AI developers should try to create more user-friendly tools that could demand minimal teachers’ preparation. Whether receiving training, the creation of user-friendly technologies, or a combination of both could work better, should be determined by further research.

Lastly, Figure 4 illustrates that the subcategories with the least research coverage are Ethical Concerns and Quality. Each of them contains challenges found in three articles, which is 6% of the 50 works. This does not necessarily mean that they are less important, but it could suggest that they have not been deeply studied and therefore that they require more investigation.
Based on Figure 4, the subcategories mainly supported by studies are **Quality** and **Implementation Challenges**. They include codes found in 22 works or 44% of the analyzed articles and in 21 sources, or 42% of the total, respectively. The most relevant codes of **Quality** are “errors” present in 13 works or 26% of the 50 sources and “no human touch” in 12 primary sources or 24%. In relation to “errors”, Ahmed (2023) compared the levels of satisfaction of 64 students with ChatGPT assisted-writing against teacher-mediated writing. The results showed that learners felt more satisfaction with the educators’ help. One of the reasons for this was the limitation of the AI tool to generate false information. In the same vein, Alotaibi (2023) conducted a survey in order to gauge learners’ experiences and perceptions of the grammar-checking capabilities of Grammarly. The findings revealed that one of the challenges of the tool is that its suggestions are occasionally incorrect. Regarding the code “no human touch,” Darwin et al. (2024) carried out a study to shed light on master’s degree learners’ perceptions of benefits and limitations of AI tools in relation to critical thinking. As a result, one of the identified challenges of AI was a lack of human touch. For instance, participant six highlighted that these technologies could miss out human elements. Similarly, Sol et al. (2024) conducted a survey study to examine perceptions, attitudes, and experiences of Cambodian learners toward the utilization of AI. Consequently, they identified that students perceived that lack of human touch was a significant limitation of AI in English education.

Considering the strong number of works for **Quality**, this subcategory could also be crucial for further research on AI tools. Besides, the focus on the codes “errors” and “no human touch” suggests that there is a need of improvement of AI tools in education, particularly for accuracy and the human element. Moreover, it might indicate the significance of human teachers even if AI technologies are being used.

Concerning **Implementation Challenges**, its most relevant code was “overreliance” which was identified in 13 works or 26% of the 50 sources. In this regard, Gültekin Talayhan and Babayiğit (2023) investigated the impact of AI tools on organization and content of learners’ writing from the view of 12 university teachers. As a consequence, the researchers found that AI technologies positively affected writing. However, instructors showed concerns about students’ overreliance on AI. Likewise, Marzuki et al. (2023) examined available AI tools for writing and evaluated their influence on this learners’ skill according to the perceptions of four teachers. As a result, they found that the educators acknowledged the positive impact of AI tools on improving students’ writing. Nonetheless, some of them also expressed their worriedness about an overreliance on them.

Taking into consideration the various articles that support **Implementation Challenges**, it could be implied that this subcategory might also be a fertile area of research on AI. Additionally, the number of studies with the code “overreliance” positions it as a relevant hurdle to implement AI tools effectively. It suggests the need to find balance between employing AI for its benefits and avoiding trusting too much on it. In this regard, exploring distinct strategies to mitigate this limitation might be valuable for future research.

Besides, Figure 4 shows that the subcategories with the lowest numbers of works are **Lack of Variety** and **Personalized Learning**. Each of them has challenges identified in three research articles or 6% of the total and in two sources or 4%, respectively. It suggests that possibly they have not been intensively examined. Thereby, they could be significant topics of future investigation.

Furthermore, it is striking to see that both, **Challenges of AI Tools in EFL Teaching** and **Learning**, contained **Implementation Challenges** as crucial subcategories. This underscores the importance of addressing these issues for the successful utilization of AI tools in education. Additionally, it is worth mentioning that according to Figures 3 and 4, more research has been done for **Challenges of AI Tools in EFL Learning** rather than for **Challenges of AI Tools in EFL Teaching**. This indicates that further investigation could be needed in that direction. Finally, in general, the present work identified a focus on benefits of **AI Tools in EFL Teaching** and Learning rather than on challenges. Thereby, it suggests that limitations could merit further examination.

**Implementation Challenges** was the subcategory with more supporting articles and inside it the code “require training” was particularly relevant. In view of the significance of this challenge for teachers, the study of Abdelatif and Siddiqui (2021) recommends providing staff members with some training and with a manual containing practical activities to incorporate AI tools. Similarly, Denecke et al. (2023) suggest teaching lecturers to use AI-based technologies more effectively. In the same vein, Ghoneim and Elghotmy (2021) state that educators should be equipped with training programs to better utilize AI interactive features. Another suggestion to overcome this limitation is given by Kohnke et al. (2023). They advise to promote the creation of teams with colleagues, so teachers can help each other to deal with this challenge. Likewise, Vera (2023) recommends to collaborate with other educators who have experience using AI in education. Finally, Han et al. (2023) suggest to use instructor-friendly platforms since they could reduce the need of training and thus mitigate this issue.

In accordance with subsection 3.2, **Quality** and **Implementation Challenges** were subcategories with the strongest backing from the 50 sources. **Quality** contained the codes “errors” and “no human touch” and **Implementation Challenges** encompassed “overreliance” as the most significant. In light of the relevance of
the subcategory *Quality* and inside it the challenge “errors” of AI, Will, a participant of the study of Kohnke et al. (2023), suggests that educators might teach students how to evaluate data so as to avoid wrong information. Besides, Harunasari (2023) points out that instructors should encourage learners to always verify output that AI provides. Furthermore, Ulla et al. (2023) corroborate the importance of confirming data with additional sources. In relation to “no human touch,” the second code of *Quality*, Aljabr (2023) advises to explore hybrid learning models which take advantage of technology and human intervention. This is in accordance with Moybeka et al. (2023) who highlight that combining AI with human interaction could create the best learning experience. Lastly, regarding *Implementation Challenges* and as part of it the code “overreliance,” Sotelo Muñoz et al. (2023) suggest learners to filter the output provided by AI instead of blindly relying on it due to possible inaccuracy.

The results indicated that the most commonly used AI tool by teachers was ChatGPT and some of the least commonly employed AI technologies by educators were Perplexity, British Council Website, and Breaking News Website (see annex 1). This information suggests that ChatGPT appears as a frontrunner for teacher adoption. Despite this, further research is recommended to confirm its significance in relation to the other tools. Besides, the diversity of AI technologies that was observed might signify instructors’ interest to explore AI functionalities for supporting their work. This should also be confirmed with more investigation.

Regarding students, their most commonly utilized tool was ChatGPT and in the second place was Grammarly. On the other hand, some of the least frequently used technologies by learners were AI Based Program, Falou, DeepL, and Plot Generator (see annex 2). Based on this, it calls attention to see that ChatGPT was the most commonly employed tool by students, too. It could suggest that it is more versatile and user-friendly than others (Chan & Hu, 2023). Another possibility is that it is more popular in AI education. What is clear is that more investigation is required to have a clear answer. Besides, it is worth mentioning that ChatGPT was spotted in more sources where it was utilized by learners rather than by instructors. This could indicate that students are more interested on it. However, further studies should be done to confirm it. Additionally, the results highlight the relevance of Grammarly perhaps because of its effectiveness to identify grammatical errors or enhance confidence (Alotaibi, 2023). Finally, it was found that students employ a greater variety of AI technologies in comparison with teachers. It might suggest that learners are currently exploring more AI functionalities. In this regard, future investigation could be advantageous, too.

The results for teachers indicated that the subcategory with the highest support of primary sources was *Teachers’ Experience* and that the main benefits of this subcategory were “satisfying” and “effective” (see Figure 1). Instructors regarded these tools as satisfying since they can help them with their classes (Hamied et al., 2022; Mohammadkarimi, 2023) or remind them of important tasks (Kohnke et al., 2023). Besides, educators considered them effective because they are capable of helping them to detect surface-level errors of students’ work (Algarad & Mahyoob, 2023). Moreover, they can be used for lesson preparation and activity creation (Ulla et al., 2023) and in general, as Zitouni (2022) mentioned, they may facilitate and improve the teaching process. It should be stated that the number of articles that support *Teachers’ Experience* make it a relevant area of investigation on AI in education and the significance of “satisfying” and “effective” indicates a positive teacher perception on AI tools. In addition, it was found that the subcategories *Improved Skills* and *Accessibility* have the lowest numbers of sources (see Figure 1). It shows that they are not necessarily less significant, but that they have not been extensively examined. Thereby, it indicates that they might be sound topics of research.

According to the results for students, the subcategories with more support of primary sources were *Students’ Experience* and *Improved Skills*. The most important codes of *Students’ Experience* were “satisfying” and “helpful” (see Figure 2). Learners considered these technologies as satisfying because they can assist them in checking paragraphs (Bok & Cho, 2023) or they can provide useful corrections (Chang et al., 2021). Moreover, students saw them as helpful since they can aid in improving vocabulary or in correcting grammar (Aljabr, 2023) and because they can help them with their writing process (Harunasari, 2023). Furthermore, it should be said that considering the strong number of articles for this subcategory, it could be a crucial area of research on AI. Besides, the focus on “satisfying” and “helpful” indicates that AI technologies are also positively seen by students. Moreover, it is thought-provoking to notice that two subcategories related to human experience, *Teachers’ Experience* and *Students’ Experience*, contained the heaviest number of supporting works and that both had the code “satisfying.” This could be a sign of a growing emphasis on a human-centered design for education which might suggest further directions of investigation on human experience with AI.

In relation to the subcategory *Improved Skills*, its most important code was “improve writing.” Taking into account the various works that support this subcategory, it could also be relevant for investigation on AI tools in education. Additionally, the significance of the code suggests that this skill can hold priority in education. Furthermore, the amount of literature for writing indicates the need of more examination on the other skills.

Additionally, it was found that the subcategories *Digital Knowledge* and *Openness to Technology* had the fewest
articles (see Figure 2). These results might suggest a need of more research for those areas which perhaps have not been extensively explored. Finally, more studies have been conducted for Benefits of AI Tools in EFL Learning rather than for Benefits of AI Tools in EFL Teaching (see Figures 1 and 2). It signifies that further research should be done in that direction.

CONCLUSIONS

Considering the findings for teachers, the subcategory with more supporting works was Implementation Challenges. Its principal limitation was “require training” (see Figure 3). Even though this subcategory had few studies, it indicates the necessity of training programs for instructors in order to deal with AI tools (Ghoneim & Elghotmy, 2021) and that AI developers could consider creating user-friendly technologies that might require minimal training.

Finally, it was observed that the subcategories Ethical Concerns and Quality had the lowest research coverage (see Figure 3). It signifies that they are not necessarily less relevant, but that they have not been deeply examined. Thereby, they might require more investigation.

Taking into account the results for learners, the subcategories principally supported by research works were Quality and Implementation Challenges. The most important codes of Quality were “errors” and “no human touch” (see Figure 4). Considering the heavy number of articles for this subcategory, it might also be key for further research on AI tools. Moreover, the emphasis on the codes “errors” and “no human touch” indicates a need of improvement of AI technologies in education. Furthermore, it could suggest the relevance of human educators even if AI tools are being used. With regard of Implementation Challenges, its most significant code was “overreliance.” Taking into account the several works that support the subcategory, it might also be a fertile area of investigation on AI. Furthermore, the number of articles with “overreliance” make it a significant challenge. It suggests a need to find balance between using AI for its benefits and avoiding relying too much on it. In relation to this, exploring distinct strategies to mitigate this limitation could be significant for future research.

Moreover, the subcategories with the lowest numbers of works were Lack of Variety and Personalized Learning (see Figure 4). It indicates that perhaps they have not been comprehensively examined. Hence, they might be relevant for future research.

Besides, it is interesting to notice that both, Challenges of AI Tools in EFL Teaching and Learning, contained Implementation Challenges as key subcategories. This highlights the relevance of addressing these limitations for the successful use of AI technologies in education. In addition, more investigation has been done for Challenges of AI Tools in EFL Learning rather than for Challenges of AI Tools in EFL Teaching (see Figures 3 and 4). It could suggest that more research should be conducted in that direction. Lastly, in general, this work encountered more benefits of AI Tools in EFL Teaching and Learning rather than challenges. Therefore, the last area undoubtedly requires future studies.

In relation to the limitations of this work, it is worth mentioning that it only considered published articles due to the challenge of finding unpublished material. Second, it was limited to the analysis of 50 sources in view of the timeframe to carry it out. Finally, it is suggested for future reviewers of AI to consider published and unpublished works since working with both of them could foster a more comprehensive understanding of AI tools in EFL education. Besides, it is recommended for them to analyze more than 50 studies to increase the generalizability of findings. Moreover, it is advised to conduct further research on benefits and challenges of AI tools in EFL teaching since according to the results, not much investigation has been done for educators in comparison than for learners. Finally, based on the emphasis on benefits identified in this work, further examination on challenges is suggested.

REFERENCES


Aljabr, F. (2023). Gauging the Saudi EFL learners level of awareness and attitudes towards the use of ChatGPT. Information Sciences Letters (ISL), 12(11), 3100-3109. https://doi.org/10.18576/isl/121120


Mahmoud, R. (2022). Implementing AI-based conversational chatbots in EFL speaking classes: An evolutionary perspective. *In Research Square, 1*-21. [https://doi.org/10.21203/rs.3.rs-1911791/v1](https://doi.org/10.21203/rs.3.rs-1911791/v1)


Rebolledo, R., & Gonzalez, F. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention, 10*(1), 7569-7576 [https://doi.org/10.18535/ijsshi/v10i01.02](https://doi.org/10.18535/ijsshi/v10i01.02)


