

Educational Assessment: Its Types and Outcomes: A Bibliometric Analytical Study

Abdul Ghani bin Md Din¹, Mohd Azizul Rahman bin Zabidin²,
Abdel Rahman Moawad Ali Tahawi³, Associate Professor Dr.
Omar bin Md Din⁴ & Assistant Professor Dr. Fareed Aawi⁵

³Sultan Abdul Halim Mua'dzam Shah International Islamic University (UniSHAMS), ⁴Al-Madinah International University (MEDIU) & ⁵Universiti Malaya

To Link this Article: <http://dx.doi.org/10.6007/IJARBSS/v13-i11/19516> DOI:10.6007/IJARBSS/v13-i11/19516

Published Date: 18-11-2023

Abstract

Educational assessment plays a significant role in enhancing students' learning. It serves as a fundamental systematic approach to draw conclusions about students' progress and development. Assessment helps determine their level of achievement in their studies and serves as a valuable method for maintaining education quality within institutions. The aim of this study is to correlate different types of educational assessment with their impact on the educational process from 2008 to 2022. The research includes all relevant publications during this period, using the VOS viewer program to identify similarities. In January 2023, an analysis of 3,058 publications in the Scopus database revealed the prominent topics covered by these journals. The results indicated that various types of educational assessment have accomplished numerous beneficial goals for students and educators. These assessments have significantly improved quality across various educational institutions and fields of study, including medicine, engineering, language, and more. The study also highlighted that students gain new insights and knowledge, especially through formative assessment, which enhances their critical thinking and performance. They have demonstrated an improved grasp of new and valuable concepts. Through bibliometric analysis, this research demonstrates that using various types of educational assessment significantly contributes to achieving the desired objectives of the educational process. This study included all relevant studies from 2008 to 2022 and visualized similarity patterns using the VOS viewer program. According to the Scopus database in August 2022, a total of 1,201 articles were analyzed to identify the most interesting topics covered in the journals. The results underscore the vital importance of using diverse educational assessment methods in the educational process. The research has shown the effectiveness of formative, diagnostic, and summative assessments in advancing the educational process. This highlights the positive impact of educational assessment on learning, boosting students' confidence in their ability to improve their skills at advanced

stages after assessment. Therefore, it is clear that various assessment types are essential for comprehending paths to educational development. Educational institutions can achieve desired quality, improve the educational process, accomplish objectives, and inspire learners to engage in and enjoy the educational journey.

Keywords: Assessment, Students, Educators, Educational Institutions, Benefits of Assessment, Educational Process Enhancement, Quality Achievement.

Introduction

Assessment is a vital tool for improvement and development across various fields, including practical, managerial, industrial, commercial, and more. In the realm of education, assessment is a fundamental process with substantial importance and numerous benefits. It serves to assess learners' comprehension levels and the extent to which they have benefited from lessons or activities during their studies. Various types of assessments assist teachers in measuring learners' progress and comprehension. It holds great significance for teachers, learners, and educational institutions alike.

Evaluating educators is crucial for refining teaching methods to help learners make strides toward achieving their desired goals. Moreover, assessment aids educational policymakers in reevaluating previously established educational objectives, aligning them more effectively with the challenges faced by educational institutions (Djedidi, 2018). Therefore, it is of utmost importance that educational institutions at all levels prioritize assessment processes.

Assessment is defined as the "process of gathering and analyzing information from various and varied sources to gain a profound insight into students' knowledge and comprehension, as well as their ability to apply their knowledge as a result of their educational experiences. This process concludes with the utilization of assessment outcomes to improve future learning" (Huba, 2000)

Assessment is an ongoing process that involves identifying clear and measurable learning outcomes and providing ample opportunities for students to achieve these outcomes. It encompasses the systematic collection, analysis, and interpretation of evidence to assess how well students' learning aligns with expectations. The information gathered is subsequently utilized to enhance student learning (Suskie, 2009).

Certain researchers define assessment as a process involving planning and gathering information to evaluate decision alternatives. Despite the varied definitions and concepts, as well as the differences surrounding educational assessment, there is a consensus that assessment is a human process used to determine the educational value of knowledge, behaviors, competencies, and curricula in general. The goal is to improve essential factors and expected processes to attain the best outcomes. Consequently, it represents the most effective means of investing in the field of education (Wasilah, 2010).

There are three general types of assessments: diagnostic assessment, formative assessment, and summative assessment. These assessments are conducted throughout the learning process, aiding both students and teachers in measuring learning.

Firstly, diagnostic assessment involves using various measures and reports to identify students' strengths and needs in specific skill areas. With this information, teachers can

provide guidance to address learning needs appropriately. It is also a fundamental component of a Multi-Tiered System of Support (MTSS). Within the MTSS framework, diagnostic assessment is used to gain a better understanding of students' academic or socio-emotional behavioral needs, which have been broadly identified through comprehensive screening.

The purpose of diagnostic assessment is to measure students' prior knowledge, skills, strengths, and weaknesses (McTighe, 2005)

The diagnostic assessment may also be referred to as placement assessment, as it is used to place students—based on previous achievements or personal characteristics—at the most suitable point in an instructional sequence, unique teaching strategy, or with an appropriate teacher (Madaus, 1970).

Secondly, formative assessment: generally conducted throughout a course or project—is also known as educational assessment. It is used to aid in learning. In an educational setting, formative assessment can be teacher-led (or peer-led) and provides feedback on student work without necessarily being used for grading purposes. Formative assessments can take the form of diagnostic or formative tests, quizzes, oral questions, or draft work. Formative assessments are conducted concurrently with instruction and may yield a score.

The formative assessments aim to determine whether students understand the instructions before conducting a summative assessment (McTighe, 2005).

Thirdly, summative assessment: Generally conducted at the end of a course or project, in an educational environment, summative assessments are usually used to assign a final grade to students. Summative assessments are conducted to summarize what students have learned, evaluating whether they understand the subject well. Typically, this type of assessment is categorized (e.g., pass/fail, 0-100) and can take the form of tests or projects. Summative assessments are often used to determine whether a student has passed or failed a course. A criticism of summative assessments is that they are reductive, revealing the extent of learners' knowledge acquisition too late to be useful (McTighe, 2005).

Summative and formative assessments are often referred to in the context of learning as assessment of learning and assessment for learning, respectively. Assessment of learning is generally summative in nature and aims to measure learning outcomes, reporting these results to students, parents, and administrators. Assessment of learning occurs at the end of a class, course, semester, or academic year. Assessment of learning is inherently formative and is used by teachers to reflect on curriculum and next steps for individual learners and the class (Earl, 2003).

It's also referred to as final assessment, as mentioned, final assessment is similar in nature to summative assessment, including end-of-unit tests, standardized tests, and cumulative work like maintaining a portfolio throughout the duration of the course. The final assessment refers to the task that evaluates students' progress in the course. This is a way for students to identify their weaknesses and for teachers to understand where they excelled and where improvement is needed in their teaching. The final assessment provides teachers with data about their teaching and their students' learning. (Top hat – [tophat.com/glossary/ final assessment](https://tophat.com/glossary/final-assessment))

In this study, we aim to achieve the following objectives

1. Identify the most effective assessments in the field of education to achieve desired goals and enhance the level of teachers, learners, and educational institutions.
2. Discover appropriate solutions for educational problems, overcoming difficulties, and improving the educational and pedagogical process.
3. Learn from the assessment procedures employed by researchers and leverage their findings to avoid educational issues, paving the way for guiding the educational and pedagogical process in the right direction.

For this purpose, the objectives of this study are to analyse online learning publications indexed within the domain using bibliometric measurements and visualization analysis. Additionally, in the current study, all data were collected from Scopus, a leading global database, and citation from research for peer review. Consequently, these research data included numerous prominent journals in assessment, detailing its types and diverse outcomes.

This analysis allowed us to observe how research interests in assessment, its types, and outcomes have evolved over time. Furthermore, this paper envisions the scientific collaboration among key contributors in assessment and its relationship with education, which was not available in previous studies.

Exclusively, we intended to address the following research questions:

- 1-What is the distribution of assessment, its types, and outcomes from educational curricula over the past decades?
- 2-Which journals and authors are most relevant in assessment, its types, and outcomes?
- 3-Which countries are the most productive in assessment, its types, and outcomes in educational research?
- 4-What are the key research keywords for the last decade in assessment, its types, and outcomes?

Materials and Methods

The aim of this review is to uncover the most productive studies conducted on educational assessment in the field of educational research. To achieve this bibliometric analysis, this research was conducted on January 17, 2023. The Scopus database was used in this study. The initial search revealed 4570 articles on educational assessment in the field of educational research. The keywords used were key in narrowing down the scope.

This study covered the last 15 years from 2008 to 2022. The research covered either the subject "community" or limited to the subject "companies" or further limited to the subject "arts". Lastly, the search language included only articles published in English, excluding any other languages. Moreover, the final articles used in this review were 1201 articles in the field of educational assessment, its types, and outcomes.

Bibliometric Analysis

Bibliometric analysis was employed in the study. Using Bibliometric Analysis, the most frequently used keywords, most cited journals, most published journals, journals that published the most studies on the subject, countries conducting the most studies on the

subject, international collaboration, relationships between keywords, most cited authors, author relationships, co-citation of journals, and most published fields were examined.

VOS viewer, a widely used software for bibliometric network visualization, was used to uncover the network visualization in the analysis. This review is conducted for the following purposes:

Firstly, the evolution of educational assessment, its types, and outcomes has developed into a compelling research domain with an increasing number of studies. Therefore, investigating the structural organization of this study area is needed through a precise mechanism-learning method that can spontaneously examine large literary data. Then, the current research is conducted to provide insights into the discussions and trends in the modern era of educational assessment, its types, and outcomes.

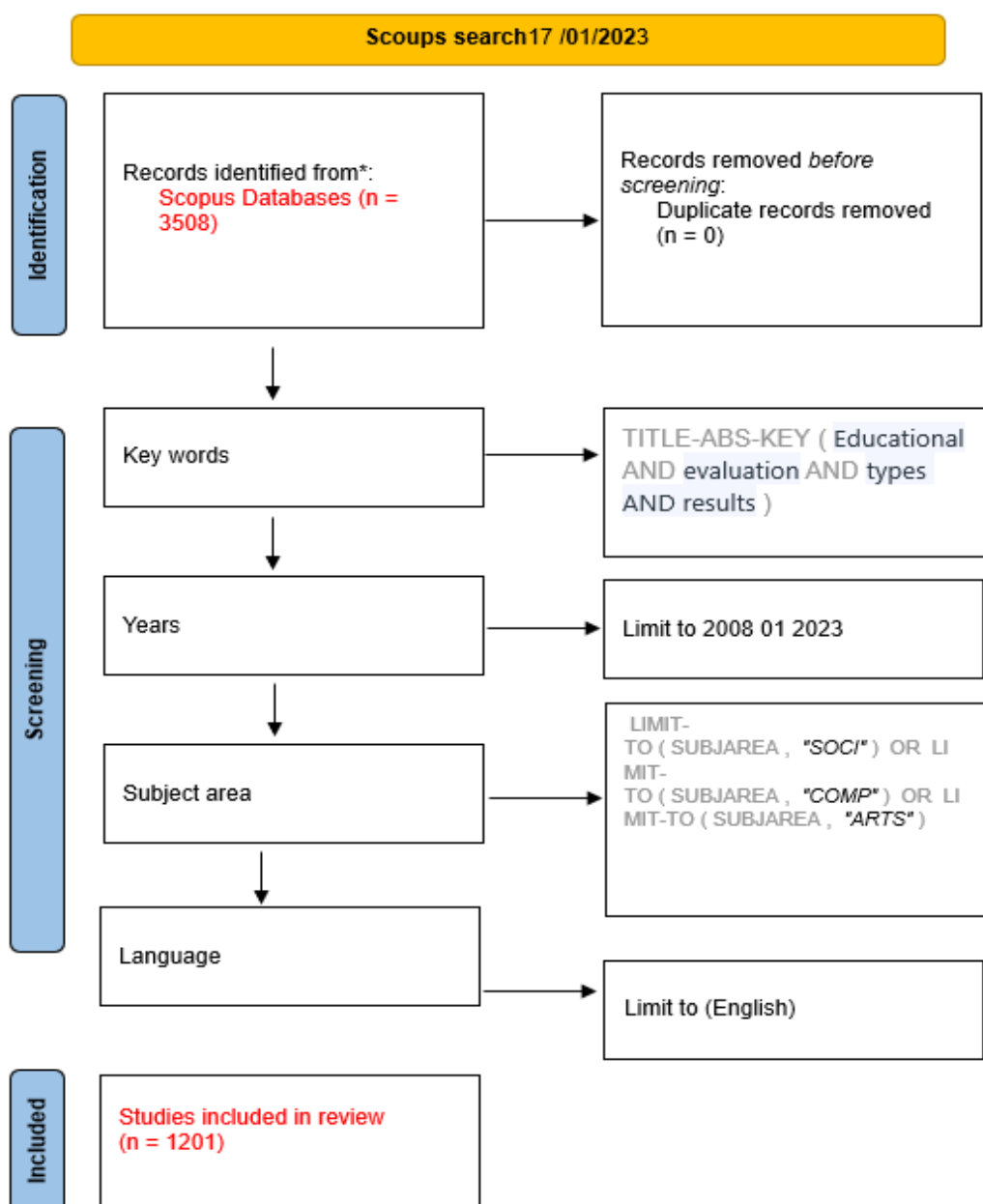


Table 1

Inclusion and exclusion criteria

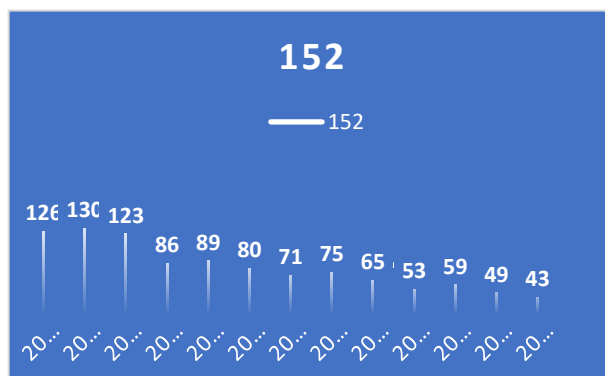
inclusion criteria	Keywords + any related keywords
exclusion criteria	Medical fields Engineering fields Conference papers Nonindexed papers

Findings

The first question: What is the distribution of educational assessment, its types, and outcomes according to years for the past decade?

To address the first question, an analysis was conducted for the publication years of articles over the past decade. It was observed that journals primarily published articles in the last two years. In the year 2022, a total of 152 articles related to educational assessment, its types, and outcomes were published. In 2021, a total of 126 articles on educational assessment, its types, and outcomes were published. It was noted that the number of articles in 2020 exceeded that of 2021, with 130 articles published on the same topic. The publication counts slightly decreased in 2019, with 123 articles published. In 2018, the publication count dropped further to 86 articles. The number of articles increased slightly in 2017, with 89 articles published. However, in 2016, the publication count decreased to only 80 articles. The decreasing trend continued in the years 2015, with 71 articles published, while there was a slight increase in 2014, with 75 articles published. The publication counts then decreased from 2013 to 2012, with 65 articles in 2013 and 53 articles in 2012. There was a small increase in 2011, with 59 articles published. The publication count decreased in 2010 to 49 articles. No articles were published in 2009, and in 2008, 43 articles were published.

This comparison reveals significant variation, where the number of published papers was 152 in the year 2022, while it was 43 papers in the year 2007, and no articles were published in 2009. It is also evident that the number of papers was elevated in the year 2020, reaching 132 papers.



The second question: What are the most relevant journals and authors in the field of educational assessment, its types, and outcomes?

In the content analysis conducted for the most cited journals, the criteria for analysis included "Total Publications," "Total Citations," "Journal Citation Score," "Most Cited Article," "Times Mentioned," and "Publisher," as indicated in Table 2.

Table 2

Top 10 High-Productivity Journals in Educational Assessment, its Types, and Outcomes for the Years (2008-2022)

Journal	Tp	Tc	Cite score	The most cited article	Times cited	Publisher
Lecture Notes in Computer Science	81183	172674	2.1	<u>UlyssesNER-Br: A Corpus of Brazilian Legislative Documents for Named Entity Recognition</u>	92	Springer Nature
ACM International Conference Proceeding Series	44179	44929	1.0	<u>Pandemic, Higher Education, and a Developing Country: How Teachers and Students Adapt to Emergency Remote Education</u>	17	Conference Proceeding

Sustainability	36485	181699	5.0	<u>The Impact of Financial Development and FDI on Renewable Energy in the UAE: A Path towards Sustainable Development</u>	58	Multidisciplinary Digital Publishing Institute (MDPI)
Communications in Computer and Information Science	21465	18992	0.9	<u>KERMITviz: Visualizing Neural Network Activations on Syntactic Trees</u>	6	Springer Nature
BMC Medical Education	1901	7103	3.7	<u>Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review</u> <i>Open Access</i>	11	Springer Nature
Journal of Surgical Education	978	4296	4.4	<u>The Effect of COVID-19 on the Medical School Experience, Specialty Selection, and Career Choice: A Qualitative Study</u> <i>Open Access</i>	5	Elsevier

Journal of Dental Education	849	1893	2.1	<u>Exploring predoctoral dental student use of YouTube as a learning tool for clinical endodontic procedures</u>	7	Wiley-Blackwell
Medical Teacher	838	4186	5.0	<u>A practical guide to reflexivity in qualitative research: AMEE Guide No. 149</u>	16	Taylor & Francis
Computers and Education	822	16299	19.8	<u>From dual digitalization to digital learning space: Exploring the digital transformation of higher education</u>	28	Elsevier
American Journal of Pharmaceutical Education	643	1893		<u>Pharmacists Leading the Way to Precision Medicine: Updates to the Core Pharmacist Competencies in Genomics</u>	6	American Association of Colleges of Pharmacy
Medical Education	573	3563	6.2	<u>Quality assurance in health professions education: Role of accreditation and licensure</u>	4	Wiley-Blackwell

Table 2 illustrates that the most productive journal in educational assessment, its types, and outcomes was "Lecture Notes in Computer Science" with a total of 81,183 publications and

172,674 citations. Following that is the "ACM International Conference Proceeding Series" with a total of 44,179 publications and 44,929 citations. Next on the list is "Sustainability" has a total of 36,485 publications and 181,699 citations.

Following that is "Communications in Computer and Information Science" with a total of 21,465 publications and 18,992 citations. Then comes "BMC Medical Education" with 1,901 publications and 7,103 citations, followed by "Journal of Surgical Education" with 978 publications and 4,296 citations. Subsequently, "Journal of Dental Education" has 849 publications and 1,893 citations. Next is "Medical Teacher" with 838 publications and 4,186 citations, followed by "Computers and Education" with 822 publications and 16,299 citations. Then "American Journal of Pharmaceutical Education" with 643 publications and 1,893 citations, and finally "Medical Education" with 573 publications and 3,563 citations.

In addition, the distribution of the most prolific authors in the field of educational assessment, types, and outcomes is also presented in Table 2. On another note, RQ2 also investigated the most prolific authors in the research area of assessment, types, and outcomes. In the content analysis conducted for highly productive authors in this field, the following criteria were selected for analysis: "Author," "Total Publications," "h-index," "Total Citations," "Current Affiliation," and "Country."

The third question: Which countries are the most productive in the field of educational assessment research?

This comparison reveals that the United States has participated in the largest number of research studies, totaling 159 research papers. On the other hand, there is a significant drop in the case of Turkey with 38 research papers and Australia with 39 research papers. Figure 3 illustrates the distribution of publications by country and research field.

It's notable that there is also a substantial difference between Spain and the United States, where Spain came in second with only 114 papers. The third position was slightly lower with the United Kingdom publishing only 81 papers. China occupied the fourth position with 64 papers, followed by Canada with 50 papers. Netherlands, with 47 papers, ranked sixth. The downward trend continues slightly with Germany at the seventh position, having published 43 papers. Italy follows with 40 papers.

Figure 3 depicts the countries and their primary fields of publication. The literature includes three groups of countries with the largest number of published papers. The top group is composed of the United States, followed by Spain with 114 articles. Then comes the United Kingdom, China, Canada, and Netherlands. The figure also includes other countries like Italy, Australia, and others, as depicted in Figure 3.

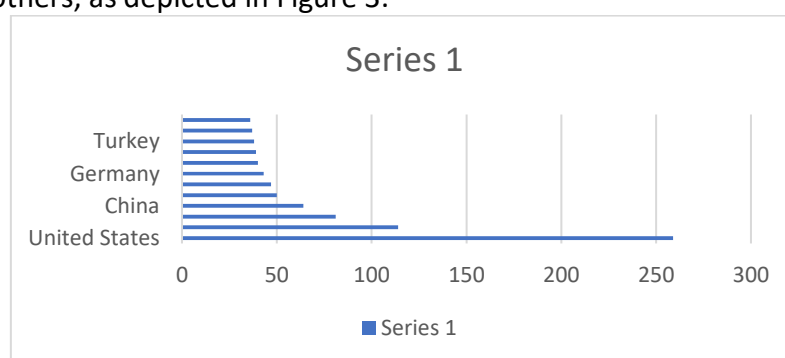


Fig. 3. Country and research field base Publication

Who are the most productive authors in the field of educational assessment research, its types, and outcomes?

As shown in Figure 4, the most productive authors have each produced three articles. The top ten authors include: Barbosa, E.F., Chen, A.M.H., Dickson, A., Garcia, F.J., Hotaling, L., Maine, A., Moreira, T., Rego, H., Satratzemi, M., and Truesdale, M. Following closely are the authors ranked eleventh and twelfth, each with two articles. This indicates a significant number of researchers who have written and contributed to this topic.

The figure also presents the distribution of articles across different types. Over the past five years, there were 885 articles categorized as "Article," 240 articles as "Conference Paper," 49 articles as "Review," 23 articles as "Book Chapter," 2 articles as "Book," and 2 articles as "Conference Review," as illustrated in Figure 5.

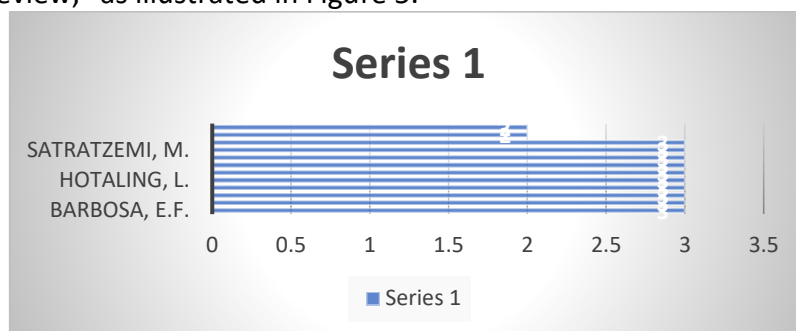


Fig 5: The fourth question: What are the keywords for research in the last decade on educational assessment: its types and outcomes?

For the bibliometric analysis of the most commonly used keywords, "co-occurrence" was selected as the analysis type, and "author keywords" were designated as the unit. In this context, a total of 400 keywords were identified from the dataset, as illustrated in Figure 4. Figure 4 shows the analysis of publication results through keywords.

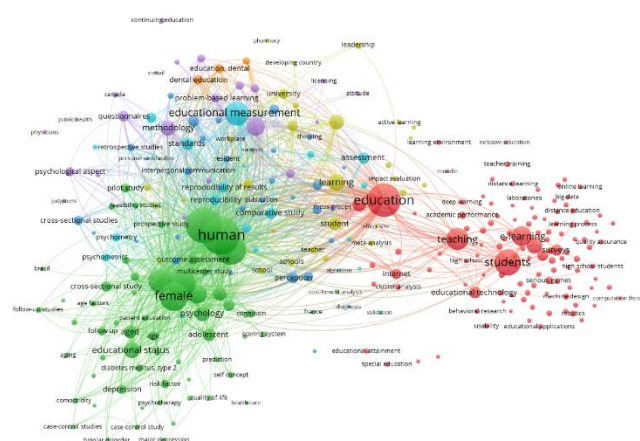


Fig. 4: Analysis results of publications by keywords

	Keyword	
1	Human	388
2	Humans	347
3	Article	254
4	Education	302
5	Female	220
6	Male	210
7	Students	183
8	Education measurement	153
9	Sdult	145
10	Teaching	132

When Figure 4 is examined, the keywords used in the studies listed as “online learning readiness” (Human ccurences “Oc” = 388), “Humans” (Oc = 347), “Education” (Oc = 302), “Article” (Oc = 254), “Female” (Oc = 220) and “Male” (Oc = 210). These were followed by Students, Education measurement, Sdult and Teaching.

Discussions

Discussion Of The Results Of Educational Assessment Research

The aim of this study was to establish the relationship between educational assessment, its types, and outcomes in improving the educational process between 2008 and 2022. It sought to determine the extent to which assessment benefits learners, educators, and educational institutions in achieving quality in education and realizing the desired goals of education. This involves achieving the maximum benefit for students who are the primary targets, requiring the evaluation of teachers and educational institutions to ensure that students' outcomes meet the desired levels as they progress through various educational stages.

It was observed that publications were predominantly concentrated in the last two years. In 2020, a total of 905 publications related to educational assessment types and outcomes were published, and in 2022, a total of 152 publications were centered around educational assessment types and outcomes. Based on 1201 collected research publications from the Scopus database, this research review provides an overview of the positive relationship between educational assessment types and outcomes to enhance and develop the educational process through content analysis and bibliometric measurements.

This analysis indicates that educational assessment has a significant positive impact on the improvement and development of the educational process. The study draws attention to the importance of various positive aspects of educational assessment types and outcomes, highlighting their significant role in advancing education and encouraging learners and educators to engage more effectively in the educational process.

This study is of utmost importance as it constitutes a fundamental step towards achieving quality in the educational process. Such studies have captured the attention of researchers and education experts, leading to the creation of numerous valuable studies and positive experiments that contribute to the development of teaching methods and the realization of desired educational goals. These efforts also yield significant benefits for learners across various academic levels.

Effective educational assessment provides invaluable feedback to higher education institutions. It also furnishes essential information to external stakeholders such as prospective students, parents, governmental and local regulatory bodies, professional and regional accrediting organizations, and representatives of the workforce. However, the selection of appropriate indicators for the educational effectiveness of programs and institutions is a challenging task, especially when clear effectiveness criteria are not well-defined.

Despite the focus of this article being on higher education, the principles and fundamental ideas will be applicable to various types and levels of educational programs (Praslova, 2010).

When we followed the research and articles we obtained from Scopus data, we found that the majority of them were formative assessments followed by summative or summary assessments. These evaluations encompassed teaching methods to enhance the educational process in different educational institutions. We also found assessments of writing and thinking skills, all of which converge under student formative preparation to reach a higher level or the desired level. Among these assessments is the critical evaluation through argumentative articles by Sato (2022), where this study aimed to identify computerized cross-sectional visualization (CT scan) skills for Level 2 students. The sample consisted of university students from Japan attending integrated content and language courses. The results indicated that evaluators' judgments of students' (CT scan) were closely correlated with task achievement. Their assessments were influenced by the appropriateness of the article to the question, content development, logic, and the quality of ideas. They concluded that these study results help to visualize (CT scan) as a structure and suggest its integration into assessment criteria for various educational contexts for Level 2 (Sato, 2022).

A researcher in his study, titled [Methodological Diversity in Language Assessment Research: The Role of Mixed Methods in Classroom-Based Language Assessment Studies], arrived at the conclusion that the formative assessment has attracted increasing attention from language assessment researchers. A considerable number of researchers have attempted to link formative assessment with contemporary psychological theories of learning, in order to explore appropriate ways to incorporate language assessment into everyday education for the purpose of enhancing learning. They found that formative assessment had positive effects on students' written performance (Tsushima, 2015). The researcher affirmed that the use of technology in education has gained acceptance among students at various levels, especially at the university level.

A study titled 'Exploring Students' Acceptance of E-Learning through Developing a Comprehensive Model of Technology Acceptance' indicated that computer self-efficacy, subjective/social norm, perceived enjoyment, system quality, information quality, content quality, accessibility, and computer playfulness were the most common external factors of TAM. (Technology Acceptance Model) (Sato, 2022).

There are many other positive outcomes of this study that confirm the benefits of educational assessment, and this falls under formative assessment. The researcher also pointed out that the results of this study have some effects on classroom-based assessment activities, especially regarding self-assessment and peer assessment in classes EEP and CLIL. These findings emphasize the development of writing skills. Self and peer assessment are recognized

as useful activities for improving student task performance. The features explored can be integrated into the development of rubrics, which guide students in a simple and suitable way for use in self-assessment and peer assessment. Additionally, students can comprehend the elusive concept of (CT scan) and recognize the characteristics of highly rated articles without deeply understanding the definition of (CT scan) itself (Sato, 2022).

In another positive study titled (Text Authenticity in Listening Assessment: Can Item Writers Be Trained to Produce Authentic-sounding Texts?)

In this study, we explored the effectiveness of an existing item-writing training course to produce authentic-sounding listening texts within the constraints of test specifications. Twenty-five trainees took an online item-writing course including training on creating authentic-sounding listening texts. Prior to and after the course, they developed a listening task. The resulting listening texts were judged on authenticity by three professional item reviewers and analysed linguistically by the researchers. Additionally, we interviewed the trainees following each item writing event and analysed their online discussions from during the course. Statistical comparison of the pre-and post-course authenticity scores revealed a positive effect of the training on item-writers' ability to produce authentic-sounding listening texts, while the linguistic analysis demonstrated that the texts produced after the training contained more instances of spoken language. The interviews and discussions revealed that item writers' awareness of spoken language features and their text production techniques influenced their ability to develop authentic-sounding texts (Rossi, 2022)

An evaluative study was also conducted in second language education on the sample selected by the researcher to assess the participants' learning abilities. This study, titled "The Interplay of Text and Image on the Meaning-making Processes of Adult L2 Learners with Emerging Literacy: Implications for Test Design and Evaluation Frameworks".

This study investigates meaning-making processes in language and literacy assessments. Using a social semiotic perspective, it examines how adult second language learners with emerging literacy self-articulate their understanding of multimodal elements and components utilized in low-stakes assessments, and the strategies they use to make meaning in such tests. The study examines how the intended meaning of assessment prompts aligns with the perceived meaning on the part of participants. To do so, we used an existing assessment tool developed for refugee-background adults in the United States as well as two experimental assessment texts created as part of the research. Data was collected via semi-structured interviews with 14 participants. Interviews were analysed using a critical multimodal social semiotic approach, informed by systemic functional linguistics. Data was coded according to meta function and theme, with results organized by test genre element and multimodal component. Results exposed tensions between participants' responses to textual and visual prompts and the expectations of test designers. Findings reveal that textual composition and assessment practices may be inadvertently biased against individuals with diverging literacy profiles. Results have implications for assessment of this population such that language, literacy, and test socialization should be considered during test design and evaluation framework development (Flores, 2021)

And this assessment serves as the summative evaluation to determine the extent of participants' progress in using American English. Among the studies conducted on the summative assessment is a study titled "The effect of testing on skills learning".

This study was conducted to determine whether testing as the final activity in a skills course increases learning outcome compared with an equal amount of time spent practising the skill.

The aim of this study is to examine the impact of the skills learning assessment in the context of a hospital-based resuscitation skills course for medical students, using simulation. The research question was: Does the final test in the hospital-based resuscitation skills course lead to improved learning outcomes compared to an equivalent amount of time spent in practice?

A total of 140 medical students participating in a mandatory 4-hour in-hospital resuscitation course in the seventh semester were randomised to either the intervention or control group and were invited to participate in an assessment of learning outcome.

In a recent review of studies on the intrinsic effect of testing, Roediger and Karpicke provided evidence that testing students on studied material results in improved retention of that material compared with spending an equivalent amount of time restudying the material. This so-called testing effect has, in both laboratory and classroom studies, been demonstrated to be a robust and independent phenomenon that applies to a variety of test formats and levels of knowledge learning. (Wheeler et al, 2002).

The researchers also mentioned that, simulation-based skills courses are fairly expensive and are confined to a limited amount of time. Thus, it is important to determine whether learning outcomes are better if some of this time is spent on testing rather than on more training. Although it has been speculated that testing might have only limited effect on skills learning, we found no systematic studies enquiring into whether the impact of the intrinsic testing effect on knowledge retention also pertains to skills learning. A major review by Rosenbaum et al. found that 'intellectual and perceptual motor skills are acquired in fundamentally similar ways' in multiple areas like feedback, massed or blocked learning.

Learning outcomes were significantly higher in the intervention group (n = 41; mean score 82.8%, 95% confidence interval [CI] 79.4-86.2) compared with the control group (n = 40; mean score 73.3%, 95% CI 70.5-76.1) (P < 0.001). Effect size was 0.93. This indicates significant real estate properties.

This study suggests that testing as a final activity in a resuscitation skills course increases learning outcome compared with an equal amount of time spent in practice. The mean score of the intervention group was superior to that of the control group, at 82.8% compared with 73.3%, yielding an ES of 0.93.

This study confirmed the significance of summative or formative assessment, as it results in positive outcomes that contribute to enhancing the learning capability and effective implementation of required procedures (Kromann et al., 2009).

One of the remarkable and great studies in the field of prominent e-learning is evaluating the success of electronic education systems.

Evaluating E-Learning Systems Success: An Empirical Study

this study aims to fill this void and address these problems by investigating the factors that influence the success of e-learning, and proposing a model that incorporates the determinants and aspects for e-learning success that are of recent concern and interest to e-

learning users, and sharing practical experiences of e-learning success measurements in developed countries such as, the UK.

This study has expanded to propose a model that integrates the determinants and aspects of e-learning success. It is clear that there is a need for a comprehensive success model for various levels of achievement (Eom & Ashill, 2018).

The model has been empirically validated by fitting the model to data collected from 563 students engaged with an e-learning system in one of the UK universities through a quantitative method of Partial Least Squares - Structural Equation Modelling (PLS-SEM). The determinants of e-learning perceived satisfaction are technical system quality, information quality, service quality, support system quality, learner quality, instructor quality, and perceived usefulness, which together explain 71.4% of the variance of perceived satisfaction.

The Researchers Have Reached Conclusions

The study revealed several findings, including

1. The study revealed that the presence of communication features, interactive activities, assessment and evaluation resources, and diverse learning methods have a positive impact on the utilization of the e-learning system. They help students engage more effectively in their learning. Therefore, more efforts should be directed towards the effective use of these tools to harness the full potential of the e-learning system.
2. The results also indicate that supportive issues within the e-learning system have a significant and positive impact on the following aspects: system usage, perceived usefulness, and noticeable satisfaction with the e-learning system.
3. The study's results highlighted the need for universities to focus considerable efforts on providing students with sufficient, concise, clear, organized, and understandable information. Regular content updates are also important. These efforts will enhance perceptions of usefulness and satisfaction with the system, ultimately leading to the realization of the benefits of using the e-learning system.
4. The study's findings can assist universities and other institutions in recognizing that system characteristics such as system usability, reliability, customization, and integration between system components need to be improved to make the system more reliable, user-friendly, customized, attractive, and easier to navigate. These aspects should be positively associated with perceived usefulness and satisfaction with the system.

This study was based on students' perspectives. Different stakeholder groups in e-learning, such as teachers and administrators, can enrich the research by providing diverse viewpoints and a better understanding of the challenges facing the success of e-learning systems. (Al-Fraihat et al., 2020).

And we perceive that the evaluation of educational effectiveness is crucial for higher education institutions. Therefore, we will present a brief overview of this study titled: Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education (Praslova, 2010)

The researcher mentioned in this study that Assessment of educational effectiveness provides vitally important feedback to Institutions of Higher Education. It also provides important

information to external stakeholders, such as prospective students, parents, governmental and local regulatory entities, professional and regional accrediting organizations, and representatives of the workforce. However, selecting appropriate indicators of educational effectiveness of programs and institutions is a difficult task, especially when criteria of effectiveness are not well defined. This article proposes a comprehensive and systematic approach to aligning criteria for educational effectiveness with specific indicators of achievement of these criteria by adapting a popular organizational training evaluation framework, the Kirkpatrick's four level model of training criteria (Praslova, 2010)

And it becomes evident to us that assessment in its various forms achieves high quality and maintains a high level of reliability and validity. Reliable assessment consistently produces the same results with the same group of students (Brown, Douglas; Abeywickrama, Priyanvada (2010)). In order for assessments to fulfil their role and effectiveness, they must rely on a set of criteria and be tailored for use in a variety of educational environments. These criteria provide guiding principles for designing, implementing, evaluating, and improving specific assessment methods. Each criterion is placed within one of four core categories to enhance educational assessments that are appropriate, useful, feasible, and accurate. In these sets of criteria, considerations of validity and reliability are covered under the topic of accuracy.

Advanced countries have shown a keen interest in assessment, as exemplified by the United Kingdom, where the Training, Assessment, and Quality Assurance (TAQA) award is provided to assist employees in learning and developing good practices related to educational assessment in various contexts such as adult education, additional education, and work-based learning.

It has become clear to us that the field of assessment, its types, and outcomes are of significant importance in identifying the strengths and weaknesses of education. Utilizing these insights to rectify deficiencies and finding appropriate solutions for the development of the educational process is essential. As a result, many educational institutions have funded and supported these assessments to leverage their results for enhancing teaching strategies. This support aids educators in mapping student learning and acquiring more information to improve teaching methods and classroom management strategies.

Graphs have revealed that the institution of Computer Science funded 389 articles, securing the first position. The Arts and Humanities category ranked second, with support for 96 articles. The third position was claimed by the Decision Science category, which endorsed and funded 21 articles.

Conclusion

1. Assessment constitutes an effective and integral component of the educational process. It achieves this by evaluating the performance of all facets within the educational framework that contribute to the educational institution. It entails a comprehensive evaluation of both the teacher's and the student's performance. Periodic assessment leads to increased productivity, development of strengths, identification and rectification of weaknesses. Additionally, it results in precise calibration of success criteria and academic achievements for students, while concurrently providing sustained and effective support to enhance the teacher's performance.

2. Evaluation achieves several goals in educational systems as it plays an effective role in posing questions and setting objectives, assessing the curriculum and tools used in the educational process, and providing tangible and meaningful results at the level of the curriculum. It leads to understanding what needs to be developed or continued within the same curriculum that the teacher is using, or selecting a new and more effective approach.
3. Evaluation is a fundamental element in various fields, and educational evaluation, in particular, is indispensable in the educational process. It is a vital tool for improvement and development. Evaluation is an essential component in achieving the goals of the educational process. Just as educational assessment helps measure the level of progress and understanding for the learner, it benefits the teacher in enhancing teaching methods and selecting the most suitable approaches for different student levels. Furthermore, it aids in choosing appropriate methods and means, achieving desired objectives, and developing teaching skills and student interaction. Evaluation also assists educational institution leaders in making informed decisions regarding educational policies and procedures. They can revisit educational objectives to make them more aligned with the educational reality experienced at all stages of educational institutions.
4. We see that the assessment process in teaching supports the effective performance of teachers in their tasks. Teaching is not limited to imparting subject alone but extends to the ability of teachers to monitor students' development and knowledge growth. Assessment is an integral part of the educational process and requires attention to key aspects: the quality of assessment, its objectives, and its primary purpose, as well as the method of formulation.
5. Assessment plays a significant and crucial role in educational processes, aiding educators and students in refining and developing teaching and learning activities. Assessment is a continuous and periodic process, offering assistance in forming judgments about educational values, the educational situation, or student achievement. Assessment, in one form or another, is inevitable in teaching and learning processes, as it is the case in all fields of educational activities that require making judgments
6. Evaluation is a process in which planning and obtaining information occur to aid in judging decision alternatives. Despite the numerous definitions, concepts, and even variations surrounding educational evaluation, there is agreement that evaluation is a human process used to determine the educational value of knowledge, behaviors, competencies, and curricula in general. Its purpose is to enhance necessary factors and anticipated processes to achieve the best outcomes, thus enabling more effective investments in the field of education.
7. Evaluation saves time, effort, and money for educational institutions, while reducing obstacles they might face in education and correcting errors appropriately and directly. Therefore, educational institutions focus on educational evaluation, continuously developing assessments and benefiting from useful experiences conducted by researchers worldwide. They apply the positive results obtained from such efforts. Investing in successful applications and designs is essential. Neglecting evaluation processes may diminish the chances of success, possibly leading to failure of both the educational institution and the educator, and ultimately failing to achieve the desired educational objectives.

8. Regarding the most frequently used keywords in bibliometric analysis, "common presence" was identified as the type of analysis, and "author keywords" were recognized as units. In this context, 400 key terms were identified from the dataset.
9. It has become clear to us that the field of evaluation, its types, and its outcomes are of great importance in understanding the strengths and weaknesses of education, and leveraging those shortcomings to avoid them and find suitable solutions to enhance the educational process. As a result, many educational institutions have funded and supported these evaluations to benefit from their results in enhancing teaching strategies. This support aids teachers in mapping out student education, acquiring more information to improve teaching methods, and managing classroom strategies. The graphical representations revealed that the "Computer Science" institution funded 389 articles, followed by "Arts and Humanities" with 96 articles, and "Decision Science" came in third, supporting and funding 21 articles. Due to the significance of evaluation, its types, and its impact on the educational process across its various stages, educational institutions have prioritized funding and supporting evaluation research due to its paramount importance in the educational process.

References

- Al-Fraihat, D., Joy, M., Masa'deh, R., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior*, 102, 67-86. <https://doi.org/10.1016/j.chb.2019.08.004>.
- Flores, A. J. A. (2021). The Interplay of Text and Image on the Meaning-making Processes of Adult L2 Learners with Emerging Literacy: Implications for Test Design and Evaluation Frameworks. *Language Assessment Quarterly an International Journal*, 18(1), 1-22. <https://doi.org/10.1080/15434303.2021.1984491>
- Cook, D. A., Hamstra, S. J., Hatala, R., & Brydges, R. (2013). What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. *Advances in Health Sciences Education*, 19. <https://doi.org/10.1007/s10459-013-9458-4>
- Djedidi, A. (2018). Importance Of Edumetry and Evaluation and Their Role in The Educational Process, *Educational Journal. Algerian Scientific Journal Platform*, 5(13), 188-198. <https://www.asjp.cerist.dz/en/downArticle/184/8/1/44186>
- Earl, L. M. (2013). *Assessment as Learning: Using Classroom Assessment to Maximise Student Learning* (2nd ed., pp. 1-139). Corwin.
- Eom, S. B., & Ashill, N. J. (2018). A system's view of e-learning success model. *Decision Sciences Journal of Innovative Education*, 16(1), 42-76. <https://doi.org/10.1111/dsji.12144>
- Huba, M. E., Freed, J. E., (2000). *Learner-Centered Assessment on College Campuses: Shifting the focus from teaching to learning* (1st ed., pp. 1-304). Pearson.
- Kunnan, A. J., Qin, C. Y., & Zhao, C. G. (2022). Developing a Scenario-Based English Language Assessment in an Asian University. *Language Assessment Quarterly an International Journal*, 19(4), 368-393. <https://doi.org/DOI:10.1080/15434303.2022.2073886>
- Kromann, C. B., Jensen, M. L., & Ringsted, C. (2008). The effect of testing on skills learning. *Medical Education*, 43(1), 21-27. <https://doi.org/10.1111/j.1365-2923.2008.03245.x>
- Lee, S. (2020). Towards Improved Assessment of L2 Collocation Knowledge. *Language Assessment Quarterly*, 18(4). <https://doi.org/10.1080/15434303.2021.1908295>

- Madaus, G. F., Peter, A. W. (1970, May 3). *Placement, Formative, Diagnostic, and Summative Evaluation of Classroom Learning* [Paper Presented at Annual Meeting]. <https://eric.ed.gov/?id=ED041829>
- McTighe, J., O'Connor, K. (2005). Seven Practices for Effective Learning. *ASCD*, 63(3). <https://www.ascd.org/el/articles/seven-practices-for-effective-learning>
- Moscal, B. M., & Leydens, J. A. (2019). Scoring Rubric Development: Validity and Reliability. *Practical Assessment, Research, and Evaluation*, 7(10). <https://doi.org/10.7275/q7rm-gg74>
- Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. *Educational Assessment, Evaluation and Accountability*, 22(3), pages215–225. <https://doi.org/10.1007/s11092-010-9098-7>
- Rossi, O., & Brunfaut, T. (2022). Text Authenticity in Listening Assessment: Can Item Writers Be Trained to Produce Authentic-sounding Texts. *Language Assessment Quarterly*, 18(4), 398-418. <https://doi.org/10.1080/15434303.2021.1895162>
- Rossi, O., & Brunfaut, T. (2021). Text Authenticity in Listening Assessment: Can Item Writers Be Trained to Produce Authentic-sounding Texts. *Language Assessment Quarterly*, 18(4). <https://doi.org/10.1080/15434303.2021.1895162>
- Russeler, J., Arendt, D., Münte, T. F., Mohammadi, B., & Boltzmann, M. (2021). What Can We Learn for Language Assessment in Low Literates? *Language Assessment Quarterly*, 18(5), 492-507. <https://doi.org/DOI: 10.1080/15434303.2021.1931231>
- Sato, T. (2022). Assessing critical thinking through L2 argumentative essays: An investigation of relevant and salient criteria from raters' perspectives. *Language Testing in Asia*, 12(1). <https://doi.org/10.1186/s40468-022-00159-4>
- Suskie, L. A. (2009). *Assessing Student Learning: A Common Sense Guide* (2nd ed.). Jossey-Bass.
- Tsushima, R. (2015). Methodological Diversity in Language Assessment Research: The Role of Mixed Methods in Classroom-Based Language Assessment Studies. *International Journal of Qualitative Methods*, 14(2), 104-121. <https://doi.org/DOI: 10.1177/160940691501400202>
- Vogt, K., Tsagari, D., Csepes, I., Green, A., & Sifakis, N. (2020). Linking Learners' Perspectives on Language Assessment Practices to Teachers' Assessment Literacy Enhancement (TALE): Insights from Four European Countries. *Language Assessment Quarterly*, 17(4). <https://doi.org/10.1080/15434303.2020.1776714>
- Wasilah, Q. H. (2010). *Assessment of the extent to which a competent approach to the goals of the new curricula is achieved within the framework of educational reforms according to primary school teachers and inspectors* [Doctoral dissertation, University of Constantine]. <http://archives.umc.edu.dz/handle/123456789/861>
- Wheeler, M., Ewers, M., & Buonanno, J. (2002). Different rates of forgetting following study versus test trials. *Memory*, 11(6), 571-580. <https://doi.org/10.1080/09658210244000414>
- Yu, C. H. (n.d.). *Reliability and Validity*. Creative Wisdom. Retrieved September 12, 2023, from <https://www.creative-wisdom.com/>