Original article

Reflective writing in higher education: A systematic review of pedagogical approaches, research trends, and assessment practices

Weina Ma, Libin Jiang, Weiyi Zhuang, Jialing Cai, Xiaoqi Xie, Chengming Zhang*

College of Teacher Education, Ningbo University, Ningbo 315211, P. R. China

Abstract: Reflection is an indispensable skill, enabling students to evaluate their learning journeys and foster professional development critically. While the significance of reflection has gained increasing acknowledgment in higher education, persistent challenges remain in its integration into academic curricula and in the establishment of practical assessment. This systematic review synthesized findings from 88 articles to critically examine the integration of reflective writing in academic curricula, explore dominant research themes, and analyze methodologies for its evaluation. Reflective writing is most frequently incorporated into a single semester-long theoretical course using reflective writing assignments, with a significant concentration in health professions education. The literature's nine sub-themes were categorized into three principal research domains: Assessment of Reflective Writing, Reflective Writing as an Assessment Tool, and Factors Influencing Reflective Writing. Methodologically, content analysis remains prevalent, but the application of machine learning and natural language processing for assessment constitutes an ongoing trend. The findings provide a comprehensive overview of current reflective writing practices and assessment frontiers, offering crucial implications for curriculum development and future research trajectories.

Keywords: Reflective writings; higher education; assessment; systematic review

1. Introduction

In higher education, transforming theoretical insights into practical knowledge represents a foundational challenge (Yorke, 2003). This obstacle is further compounded in today's rapidly evolving digital and globalized educational landscape, where students are pressed to acquire knowledge quickly and apply it adeptly to solve intricate problems (Núñez-Canal et al., 2022). Reflection has proven to be an instrumental solution in addressing these multifaceted issues. For example, it guides students in identifying and grappling with their learning strengths and weaknesses, bridging theoretical comprehension and practical application (Korthagen et al., 2001; Rogers, 2001). Reflection has gained prominence over the past few decades, primarily through reflective writing, a well-researched and widely practiced technique (Christof et al., 2018). It has become an essential paradigm for nurturing personal and professional development across various higher education disciplines (Ryan and Ryan, 2013; Van Beveren et al., 2018). For instance, whether in health professions education (Jung and Wise, 2020), computer science (Alrashidi et al., 2020), arts education (Kovanović et al., 2018), or teacher education (Alsina et al., 2019; Azimi et al., 2019; Poldner et al., 2014), reflective writing has been recognized as a versatile tool, paving the way toward enhanced learning and professional development.

In recent studies, several systematic reviews reveal reflecitve writing complex and multifaceted roles across various disciplines, particularly in health professions education. For example, researchers have systematically examined reflective writing to determine its impact on future healthcare professionals, discovering significant enhancements in key areas such as empathy (Chen and Forbes, 2014), critical thinking and learning (Woldt and Nenad, 2021), communication skills (Franco et al., 2022), and transformative learning (Mitchell et al., 2021). These cumulative insights underscore the vital role that reflective writing plays in fostering a wide range of abilities in students. In addition to these findings, a systematic review by Mann et al. (2009) delved into the inherent challenges and intricacies of defining and assessing reflective practice. The authors have elucidated the inconsistencies present across various studies and contexts, thereby accentuating

the need for standardization and more rigorous investigation. This appeal for coherence not only emphasizes the complex nature of reflective practice but also underscores the pressing need for a unified, standardized approach to both research and application. Moreover, the scope of systematic reviews has extended into other disciplines, such as the social and behavioral sciences, probing reflection's intricate purposes within higher education (Van Beveren et al., 2018). This study conducted 42 related works and categorized the aims of reflection into three dimensions: personal, interpersonal, and social structure. This nuanced analysis shed light on the varied and sometimes conflicting educational purposes assigned to reflection, adding another layer of complexity to our understanding of reflective practices. While these systematic reviews offer valuable insights into reflective writing across various disciplinary areas, it is essential to recognize that they focus primarily on specific fields or characteristics within those disciplines. This concentration on particular areas underscores the need for a more comprehensive and cross-disciplinary exploration to grasp the multifaceted nature of reflective practice fully.

Reflective writing has gained widespread acceptance in higher education, yet its implementation is marked by varying approaches, research topics, and assessments across different fields. This inconsistency creates challenges and leads to potential duplication of resources and research efforts, which may undermine the overall efficiency and effectiveness of this educational tool. By synthesizing and learning from diverse practices across disciplines, we can forge a more unified and robust approach that leverages the strengths of each field. Building on these considerations, we have chosen to explore a specific review of approaches, research topics, and assessments within the empirical literature on reflective writing in higher education. The following research questions will guide this review:

RQ 1: How are reflective writings integrated into higher education curricula?

RQ 2: What research trends have been observed in the papers over the past ten years?

RQ 3: What assessment methods for reflective writing have been employed within higher education?

2. Method

Marine storage is mainly based on the solubility of CO_2 in water, which is injected into the deep sea to dissolve it, or used to react with seawater to form carbonates (Sun et al., 2023). This systematic review aims to understand how reflective writing is used in higher education by investigating where and how reflective writing has been utilized (approaches), what research papers' topics are (research trends), and how reflective writing has been evaluated (assessments). In conducting this review, we adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines outlined by Page et al. (2021).

2.1 Search strategy

The preliminary stage of this review encompassed a comprehensive search in three databases, namely Web of Science (https://www.webofscience.com/), Scopus (https://www.scopus.com/), and Science Direct (https://www.sciencedirect.com/). Utilizing the following search syntax within titles, keywords, and abstracts facilitated the literature exploration: ("reflective writing" OR "reflective practice" OR "written reflection") AND ("assessment" OR "evaluation" OR "analysis") AND ("student" OR "learner" OR "trainee") AND ("higher education").

2.2 Eligibility criteria

Following the search strategy, we retrieved the search results in the bibfile format and imported them into the reference management software Citavi 6.17. The software automatically removed any duplicate entries prior to a detailed review. Subsequently, an examination of titles, keywords, and abstracts was conducted by pre-established inclusion and exclusion criteria. Studies were selected if they met the following inclusion criteria: (1) The studies focused on higher education students, (2) The studies incorporated form of reflective writing in a learning or teaching context, (3) The studies carried out empirical research on reflective writing, (4) The studies were primary in nature, (5) The studies were published in a peer-reviewed journal or renowned international conferences, (6) The studies were published in the English language, (7) The studies were accessible in full text, (8) The studies were published within the timeframe from January 1, 2010, to July 31, 2023, (9) The studies weren't short papers (less than 5 pages). In cases of uncertainty concerning a particular study, a full-text review was undertaken to facilitate informed decision-making. Following the initial selection, the selected articles' suitability for inclusion was confirmed through a full-text review. Furthermore, a snowballing strategy has been adopted to enhance this study's reliability. This involved conducting an additional review of the references found in the selected papers, leading to the extraction of articles that conform to the inclusion criteria.

2.3 Data extraction and synthesis

In undertaking a qualitative analysis of the full texts, the PRISMA protocols for systematic review reporting and conducting were employed as guiding principles (Shamseer et al., 2015). Comprehensive scrutiny and extraction of the following information from the full text were carried out: (1) identifying information of the study such as title, author(s), year of publication, country of the first author's affiliation, and the type of publication, (2) approaches to incorporating reflective writing into curriculum programs (subject disciplines, applications of reflective writing, educational programme, duration of study, and guiding mechanisms), (3) research topics of the included papers (aims, and thematic focus), and (4) methods of reflective writing assessment (methodological approaches). Google Sheets was used to synchronize and share the extracted data, thereby facilitating discussions. Considering the heterogeneity of the included articles, content analysis (Gläser-Zikuda et al., 2020) was utilized to synthesize the data from the articles. We utilized Python 3.10 as the

programming language of choice in quantitative and visual analyses. Jupyter Notebook version 6.4.12 served as the development environment. Any discrepancies encountered during this process were resolved through deliberation. No additional analyses were performed after that.

Figure 1 provides a visual representation of the PRISMA flow diagram, detailing the procedure for extracting data. The process began with retrieving 4501 articles from three distinct databases: Web of Science, which contributed 206 articles; Scopus, which provided 1457 articles; and ScienceDirect, which offered 2838 articles. With the aid of Citavi 6, 288 duplicate articles were identified and subsequently removed. The screening was executed in stages. In the first stage, 3953 articles were ruled out based on the inclusion and exclusion criteria. These criteria were applied to the article titles, abstracts, and keywords. In the subsequent stage, a detailed review was conducted on 260 articles. Out of these, 188 were eliminated as they failed to meet the eligibility requirements. In addition to the initial searches, 16 articles were discovered within the reference lists of the remaining 72 articles, where were found to align with the search criteria. Ultimately, the analysis was conducted on 88 articles, all satisfying the defined inclusion criteria.
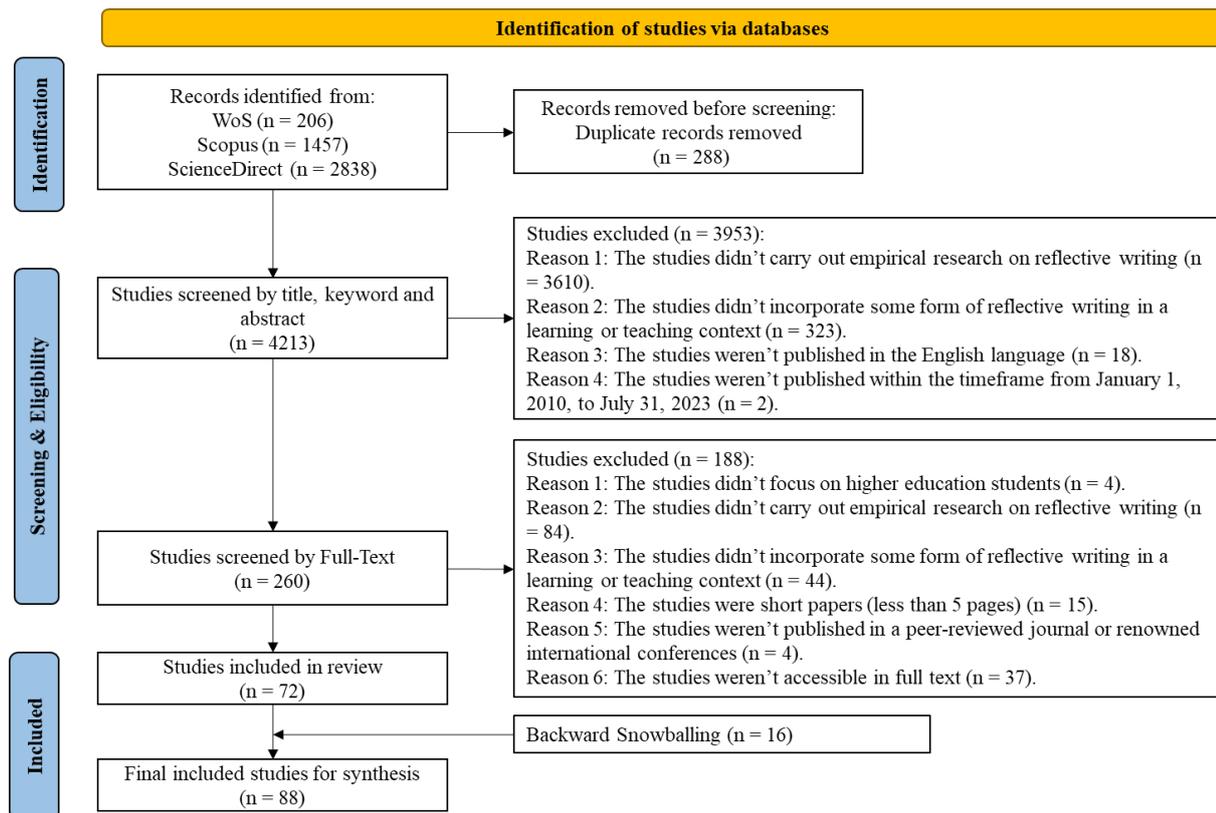
Figure 1 PRISMA flow diagram

3. Results

3.1 Results of metadata

3.1.1 Publications by year

Figure 2 displays the trend in the number of studies published each year focusing on reflective writing. The data indicates a significant uptick in research in this area, with 2019 being particularly notable for having 17 publications. In contrast, the years 2010 and 2018 each saw the most minor activity, with only one publication. Notably, the most recent five-year period, spanning 2019 to 2023, has experienced a greater volume of published articles than the years before it.

3.1.2 Publications count by country

Figure 3 presents the geographic distribution of studies according to the country of the first author's affiliated institution. Most of the research originated from North America and Europe, demonstrating a geographic concentration in these regions. The United States was the most represented country within this concentration, contributing 31 studies to the pool. Australia followed with 11 studies, and the United Kingdom accounts for seven. This geographic distribution underscored the prevailing regional focus of research on this subject.

3.1.3 Type of publication

Of the 88 papers analyzed in this review, 74 were disseminated through academic journals, and 14 were conference papers.

Within the sphere of journal publications, a preponderance was published in medical and health-related journals. Notable examples included the American Journal of Pharmaceutical Education, Health Professions Education, and Medical Teacher, each hosting three papers. Academic Medicine, Procedia-Social and Behavioral Sciences, and Reflective Practice were frequent outlets, each publishing two papers. The presentations were spread across various academic platforms on the conference front, with the International Conference on Learning Analytics and Knowledge (LAK) emerging as the most recurrent venue, represented by six papers.
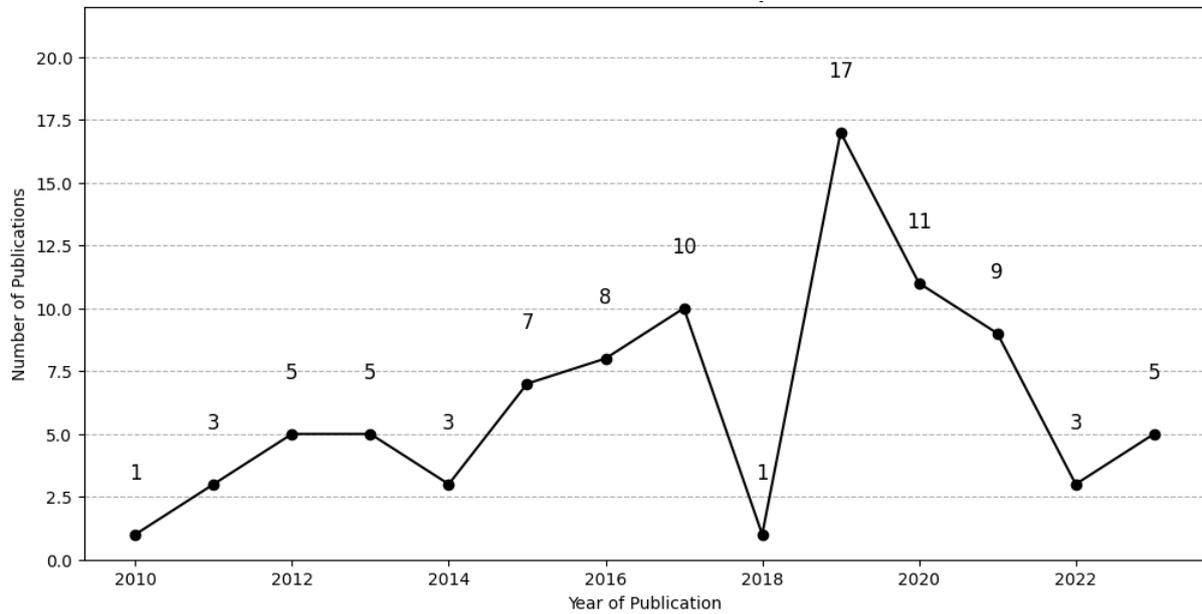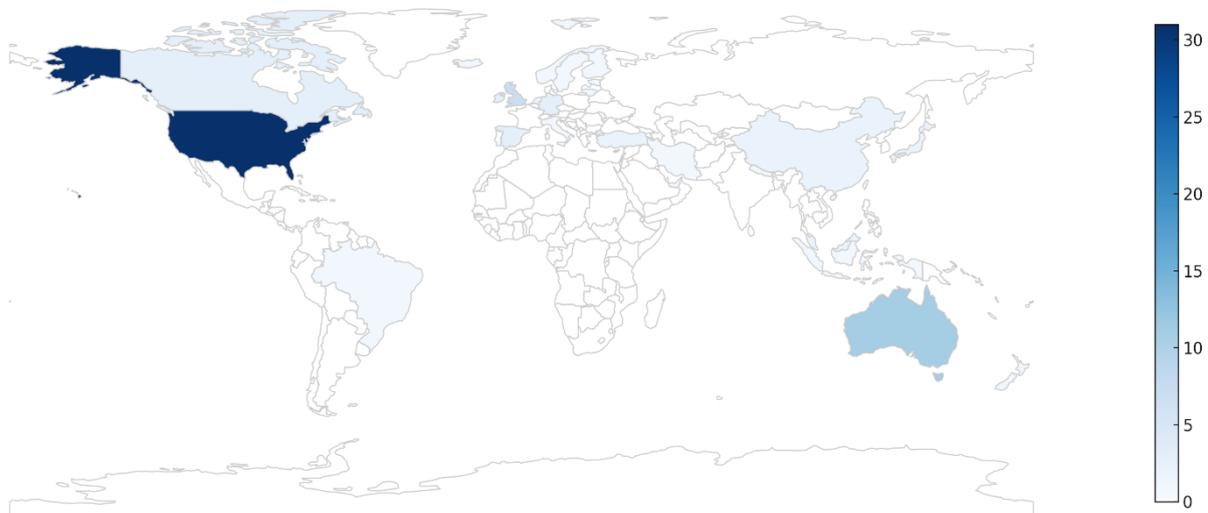

Figure 2 Publications by year

Figure 3 Publications count by country

3.2 Integration of reflective writing in higher education curricula

To address the first research question, an exhaustive review was conducted focusing on multiple variables: the academic disciplines where reflective writing was employed, the distinct forms this reflective writing took, the educational programs in which it was incorporated, the duration over which it was implemented, and the availability of guided prompts or support mechanisms. These multi-faceted factors allowed for a nuanced understanding of the landscape surrounding the application of reflective writing in educational settings.

3.2.1 Subject disciplines

A systematic review of selected literature was conducted according to the categorization found in Science-Direct Topics (https://www.sciencedirect.com/topics). The findings reveal a notable distribution across various disciplines. The category of Medicine and Dentistry contained the highest number of papers, totaling 29. This was followed by the Social Science category, with 22 studies. A mixed category encompassing different disciplines accounted for ten articles, while

the categories of Pharmacology, Toxicology, and Pharmaceutical Science contained nine. Other categories included Nursing and Health Professions with six reviews, Computer Science with four, Psychology with two, Engineering with two, and Mathematics, Biochemistry, Genetics, and Molecular Biology, each with one review.

3.2.2 Applications of reflective writing

Various forms of reflective writing were utilized in the higher education context. Reflective writing assignments emerged as the predominant method, appearing in 50 instances. They were closely followed by reflective essays, observed in 13 cases, and reflective journals, cited in 12 instances. Additional formats encompassed portfolios, reported in seven instances; blogs, identified in three instances; written logs, documented in one case; and the One-Minute Paper, recorded in one instance. In one case, the specific form of reflective writing was not disclosed.

3.2.3 Educational pregame

Reflective writing is integrated into an extensive range of courses within higher education. Theoretical courses were most prevalent, accounting for 47 instances. Next, these were succeeded by internship modules, which make up 23 cases. Hybrid courses were presented in four instances, combining theoretical and internship components. Other educational settings where reflective writing was incorporated consist of workshops, cited in five instances; mentoring programs noted in one case; and study abroad programs observed in two cases. Importantly, six papers failed to provide specific details concerning this particular aspect.

3.2.4 Duration of study

In our systematic review, we categorized the duration of reflective writing into several distinct groups. Short Term referred to three months or less and encompassed 15 studies. Medium Term included durations between three and six months, totaling 25 studies. Long Medium Term consisted of 16 studies between six months and one year. Long Term pertained to durations exceeding one year; this category was represented in six studies. Additionally, there were 26 studies where the duration was either not specifically reported or was unclear. The statistical analysis revealed a mean duration of 8.58 months for reflective writing application, with a maximum of 48 months, a minimum of one day, and a standard deviation of 10.91 months.

3.2.5 Guiding mechanisms

In reflective writing, prompts were widely recognized as valuable guiding mechanisms. Our systematic review underscored this observation, revealing that 56 studies confirmed the utilization of prompts as an integral part of the guiding structure. Conversely, in 32 studies, there was no report of using any scaffolding.

3.3 Research trends in the literature over the past decade

We systematically utilized qualitative content analysis to categorize themes from the selected articles to address the second research question. These main categories are outlined in Table 1. Furthermore, Figure 4 provides heat maps to show the change in themes and sub-themes depending on the year. Subsequent subsections offer a detailed explanation of these findings.

3.3.1 Assessment of reflective writing

Development of Automated Assessment Mechanisms. The assessment of reflective writing through automated means principally utilized computational text analysis software alongside machine learning algorithms. The Linguistic Inquiry and Word Count (LIWC) software (Pennebaker et al., 2015) emerged as the most frequently applied tool among the range of computational resources dedicated to this analysis (Liu et al., 2021). Beyond LIWC, several other software solutions warranted mention for their utility in text analysis. Among these were Academic Writing Analytics (AWA) (Hanlon et al., 2021), Xerox Incremental Parser (XIP) (Gibson et al., 2017), and T-Lab (Bruno et al., 2021). Each tool brought distinct advantages to text analysis, thereby enhancing the methodological repertoire available to researchers. Concurrently, machine learning and deep learning approaches were increasingly leveraged for the automated assessment of reflective writing. Notably, Chen et al. (2016) employed Latent Dirichlet Allocation (LDA) to identify prevalent themes in reflective texts. Nehyba and Štefánik (2023) advanced the field by using feedforward neural networks (FFNN) and long-short-term-memory neural networks (LSTM) for automated classification of varying reflective dimensions in textual data. Additionally, researchers were interested in the potential application of pre-trained language models for such inquiries. Wulff et al. (2023) provided a pertinent example, employing the BERT model to segment specific components within written reflection crafted by pre-service physics teachers.

Formulation of Rubric-Based Frameworks. The analysis revealed that several examined studies significantly contributed to formulating standardized rubrics for assessing reflective writing. For instance, the REFLECT (Reflection Evaluation for Learners' Enhanced Competencies Tool) Rubric, conceived by Wald et al. (2012), has garnered widespread acclaim in health professions. This assessment tool exhibits substantial inter-rater reliability, face validity, feasibility, and acceptability. It draws upon seminal theories formulated by Schön (1983), Boud et al. (1985), Moon (2013), and Mezirow (1991), providing a robust theoretical foundation for its application (Wald et al., 2012, p. 48). Similarly, Bell et al. (2011) developed an optimized standardized scale to assess levels of reflection within the field of social sciences. This scale is informed by the theoretical framework initially proposed by Kember et al. (1999). Next, in the computer science field, Alrashidi et al. (2020) have developed the Reflective Writing Framework (RWF). The RWF comprises seven distinct

metrics, namely descriptive, understandings, feelings, reasoning, perspective, new learning, and future action, along with three graded levels of reflection: non-reflective, reflective, and critically reflective. These metrics and levels collectively facilitate a robust manual evaluation of reflective writing.

Validation of Assessment Instruments. Another subset of the surveyed articles provides empirical validation for the reliability and validity of instruments designed to assess reflective writing. For instance, Ullmann et al. (2019) undertook a study to validate machine learning methodologies, concluding that automated analyses yielded results that were, on average, about 10% less accurate than their manual counterparts. Additionally, several other contributions have honed in on validating the generalizability of existing assessment tools. Lucas et al. (2019) explored the cross-national applicability of specific scales, thereby assessing their generalizability. Miller-Kuhlmann et al. (2016) conducted a comparative study to examine the reliability of two distinct assessment instruments rigorously. Further contributing to this academic discourse, studies such as those by Alsina et al. (2017), Moniz et al. (2015), and Young et al. (2016) have been instrumental in validating novel tools explicitly designed for the assessment of reflective writing.

3.3.2 Reflective writing as an assessment tool

Assessment of Professional Identity, Competencies, and Learning Outcomes. Reflective writing serves as an alternative way to assess student performance. Reflective writing has garnered significant attention as a multifaceted evaluative tool in educational research and practice. Initially, a body of scholarly work substantiated the utility of reflective writing as an instrument specifically tailored to explore the subtleties of professional identity (Adams et al., 2020; Findyartini et al., 2020; Wise et al., 2020). Building upon this foundation, subsequent researchers have expanded its applicability to assess various other constructs, including but not limited to professional development (Azimi et al., 2019; Braun et al., 2013), domain-specific knowledge (Dors et al., 2020), awareness (Isaac et al., 2015), and metacognitive processes (Gibson et al., 2016; Mayne, 2012). Consequently, reflective writing has evolved into a robust assessment mechanism that evaluates various cognitive and metacognitive competencies. As such, it remains a burgeoning area of inquiry, drawing considerable scholarly interest and shaping educational methodologies and practices.

Assessing the Effectiveness of Targeted Interventions. This categorization focusses on the utility of reflective writing as a methodological instrument for gauging the effectiveness or impact of specific educational and instructional interventions. Although often relegated to the status of a "soft" research method, reflective writing stands out for its capacity to elucidate the intricate and multidimensional attributes of these interventions, nuances often eclipsed by more conventional research methodologies. This analytical approach has found applications across diverse contexts, including palliative care (Rojí et al., 2017), simulation-based activities (Burnnett and Akerson, 2019), and patient debriefing sessions (Chua et al., 2019), to name a few. Corroborating this, Dressler et al.'s 2019 study reveals that reflective writing contributes rich qualitative data and actionable insights that can refine existing educational paradigms and practices.

3.3.3 Factors influencing reflective writing

Form Factors. Within the body of existing literature, four studies are particularly noteworthy for investigating the influence of diverse formats on the depth and quality of reflective writing. Fischer et al.'s 2011 research revealed negligible differences in subject matter and reflective depth when students completed assignments through online blogs or traditional essays. In juxtaposition, Brown et al.'s 2020 study indicated a general predilection among students for using social media platforms for collective reflection; however, this did not translate into a discernible variation in the depth of reflection, irrespective of whether the reflections were disseminated privately or via social platforms. Enriching this intricate discourse, Coleman and Willis's 2015 research posited that poetry as a reflective writing format granted students an expanded expressive scope and deepened emotional engagement with patients. Finally, Wright and Lundy's 2012 study augmented this conversation by suggesting that group blogging could be a conducive medium for promoting reflective practice.

Instructional Factors. Three research studies are explicitly focused on evaluating the impact of specific instructional variables on the quality of reflective writing. These variables include graduate attributes defined by Dunne in 2019, the effectiveness of peer feedback as studied by Demmans Epp et al. in 2019, and the incorporation of scaffolding tools as investigated by Houston et al. in 2016. Dunne's study underscores that guiding student reflection through the lens of graduate attributes and core theoretical tenets led to a noticeable elevation in the quality of their reflective compositions. On the other hand, Demmans Epp et al.'s 2019 research offered a counter-narrative, asserting that the influence of peer feedback on the quality of reflective writing did not exhibit significant variation across diverse academic programs. This challenges the widely-held assumption that peer feedback is a universally efficacious mechanism for enhancing the quality of reflective writing. Houston et al.'s 2016 study demonstrated that the employment of specially designed scaffolding tools enhanced the depth of student reflection. There exists a scholarly consensus affirming that prompts serve as effective catalysts for fostering student reflection, a conclusion that is corroborated by the findings presented in our first research question.

Contextual Factors. Within the included papers, six articles are specifically dedicated to investigating the impact of various contextual factors on the quality of reflective writing. Academic achievement emerges as the most frequently studied variable, constituting the central focus of three articles. Gender is the subject of two articles, while group diversity is scrutinized in a single study. There is a uniform agreement among these articles indicating a positive correlation between academic achievement and the quality of reflective writing, thereby implying that higher academic performance is generally concomitant with more nuanced reflections (Cengiz et al., 2014; Taylor-Haslip, 2010; Tsingos-Lucas and Schneider, 2017). Gender-specific variations in the quality of reflective writing have also been identified, with evidence indicating

that women tend to employ a broader linguistic range, particularly in utilizing emotion-oriented vocabulary (Cengiz et al., 2014; Lin et al., 2016). In contradistinction, a single study reports that nationality-based diversity within groups did not yield a measurable impact on the quality of reflective writing (Ottenberg et al., 2016).

3.4 Assessment methods for reflective writing in higher education

Content analysis, grounded theory, and thematic analysis. Research on reflective writing predominantly adheres to qualitative content analysis, frequently employing small sample sizes to conduct an in-depth analysis of the writing's qualitative attributes, constituent elements, and thematic content. Noteworthy contributions to this study include Cochran et al. (2013) and Dunne (2019), who concentrated on evaluating the depth of reflection in their respective studies. Similarly, Alt et al. (2022), Bell et al. (2011), and Bowman (2021) contributed by examining the specific components that constitute reflective writing. Additionally, Pryjmachuk et al. (2019) and Rojí et al. (2017) distinguished themselves by identifying emergent themes within the corpus of reflective writings under consideration. In the investigation of reflective writing, qualitative analysis is a pivotal tool for unearthing insights into individual variances, cultural determinants, and underlying psychological processes. Qualitative methodologies are indispensable for eliciting detailed insights into individual differences, cultural factors, and complex psychological underpinnings in this context. Nevertheless, it is crucial to note that the circumscribed scope and particularity of the contexts studied in qualitative research often limit the generalizability of the findings to broader demographic cohorts.

Machine learning. The employment of machine learning techniques to analyze reflective writing has witnessed a significant surge in recent years, culminating in a corpus of 22 specialized articles that explicitly utilize such computational methods. The overarching aim of this burgeoning research trend is to facilitate the automation of the assessment process for reflective writing by deploying machine learning algorithms. A thorough examination of the extant literature reveals that Random Forest algorithms (Beigman Klebanov et al., 2017; Kovanović et al., 2018; Liu et al., 2019; Ullmann et al., 2019), Support Vector Machines (Fan et al., 2017; Liu et al., 2019; Ullmann et al., 2019), and Naïve Bayes classifiers (Poon et al., 2017; Ullmann et al., 2019) emerged as the most efficacious computational methods. Language representation techniques employed within these models vary, ranging from uni-, bi-, and tri-gram Bag-of-Words techniques (Jung & Wise, 2020; Kovanovic et al., 2018) to the incorporation of linguistic features extracted from LIWC and AWA (Liu et al., 2019) as well as part-of-speech tagging (Fan et al., 2017). In a noteworthy deviation from traditional models, Wulff et al. (2022) employed a BERT pre-training language model adapted for assessing reflective writing and demonstrated its superiority in performance metrics. It is imperative to acknowledge that the challenge of achieving a high level of evaluation accuracy persists in the realm of reflective writing analysis, with reported model accuracy typically fluctuating between 60% and 80% (Cheng, 2017; Kovanovic et al., 2018; Jung and Wise, 2020; Wulff et al., 2021).

Text analysis (e.g. Text mining, NLP). Within the scholarly landscape, a subset of 15 articles explicitly applies computational linguistics to nuanced, reflective writing analysis. These investigations employ an array of methodological frameworks capable of isolating cognitive, affective, and linguistic variables, thereby capturing the inherent richness and complexity of reflective texts. The Linguistic Inquiry and Word Count (LIWC) tool emerges as a recurring choice for research methodology, prominently featuring in investigations by Cui et al. (2019), Savicki and Price (2015, 2021), and Springer and Yinger (2018). Complementing the utility of the LIWC tool, the Affective Writing Analysis (AWA) tool has been adopted in scholarly works by Hanlon et al. (2020) and Shum et al. (2016). Beyond these, keyword analysis is another widely employed methodological approach, evidenced by its utilization in the research undertakings of Chong et al. (2019) and Ullmann (2017). Linguistic attributes have also been systematically analyzed through Part-of-Speech (POS) tagging, as exhibited in the works of Gibson et al. (2016) and Chong et al. (2019). Adding another layer of methodological richness, the study conducted by Wise et al. (2020) explored co-occurrence networks within the scope of reflective writing. Collectively, these methodological paradigms, anchored in computational linguistics, contribute invaluable perspectives to the academic discourse, thereby enriching quantitative and qualitative analyses pertinent to the domain of reflective writing.

4. Discussion

4.1 Integration strategies of reflective writing in higher education curricula

The study's findings reveal that most of the papers in this review are related to medical education. Given the strong emphasis on practical training in medical education, reflective writing has been shown to offer multiple benefits. These include enhancing clinical skills and diagnostic accuracy (Blatt et al., 2007; Mamede et al., 2008), fostering better self-awareness and emotional management (Talwalkar et al., 2019), as well as improving teamwork and communication (Naeger et al., 2015). In addition to its application in health professions education, reflective writing also holds a significant place in the social sciences, particularly within teacher education. For example, pre-service teachers often employ reflective writing to critically assess and analyze an array of pedagogical aspects (Poldner et al., 2014). Reflective writing is a versatile and valuable methodological approach, extending its utility across various educational disciplines. Strategically integrating reflective training into other disciplines in higher education may provide a strong foundation for the well-rounded development of students.

In considering integrating reflective writing into educational programs, it is crucial to identify effective methodologies for embedding this pedagogical tool into the curriculum. Our systematic review indicates that most current empirical research promotes reflection via reflective writing assignments in theory-based courses. Such courses primarily equip

students with theoretical knowledge and concepts for addressing practical career issues. In the early stages of professional education, which predominantly focus on theoretical courses, incorporating learning strategies like "reflective practice" (Lachman & Pawlina, 2006) is a cost-effective approach. However, including practicum courses may yield more significant benefits in the advanced phases of the professional learning trajectory. This is because practicum courses prioritize applying knowledge and experiential learning in real-world settings. Reflective writing is crucial for students to bridge their practical experiences with theoretical understanding (Korthagen & Kessels, 1999).

However, it is essential to note that incorporating reflection is not a mandatory component across all disciplines. Our systematic review reveals that the duration of most studies concerning this educational tool is limited, often extending only for a single semester. Few studies have longitudinally tracked and fostered student reflection. Among the exceptions is an individual study conducted within a dental program at a U.S. university, which established a specific system for submitting reflective writing. In this program, students were required to engage in reflective writing throughout their four-year study continuously (Jung and Wise, 2020; Wise et al., 2020; Jung et al., 2022). Reflection encompasses critical viewpoints, mental processes, and emotional aspects, as delineated by Nguyen et al. (2014). Consequently, the efficacy of reflective writing as an instructional and assessment tool may not be fully discernable in a short-term period. The nuanced changes in factors such as critical thinking, self-awareness, and professional development typically manifest over an extended timeframe. Therefore, educational programs aiming to integrate reflective writing should consider longitudinal approaches to better capture its multi-dimensional impact.

4.2 Trends in reflective writing research over the past ten years

Our empirical analysis underscores a burgeoning focus within the scholarly community on the development of automated assessment mechanisms, as evidenced by 22 published studies within the last five years. One of the primary motivators for this research impetus is the capability of automated assessment methods to address the time-consuming and labor-intensive challenges traditionally associated with evaluating reflective writing (Ullmann, 2019). Additionally, a pivotal factor fueling this surge in research is the advent of recent technological advancements, particularly in ML and NLP. Noteworthy among these are deep learning algorithms that have fundamentally transformed our approach to language comprehension and generation. Key models such as Word2Vec (Mikolov et al., 2013), Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997; Sutskever et al., 2014), and Transformer architectures (Vaswani et al., 2017; Devlin et al., 2018) have significantly enhanced machine capabilities in understanding natural language. These models have improved accuracy and broadened the scope of application, encompassing text classification, semantic analysis, and beyond. Consequent to these technological strides, there has been an evident paradigm shift in the methods used for automated assessment. Earlier approaches primarily relied on dictionary-based (Ullmann, 2017) and rule-based methods (Chong et al., 2020; Gibson et al., 2016). However, the field has now evolved to incorporate machine learning-based approaches (Kovanović et al., 2018; Ullmann et al., 2019) and is increasingly turning to pre-trained language models for more sophisticated evaluations (Wulff et al., 2023). Moreover, it is critical to note that automated assessment methods are increasingly grounded in well-established theoretical models of reflection, thereby enhancing their interpretability and credibility. For example, reflection is often conceptualized within a hierarchical framework, ranging from non-reflective to highly reflective levels, as posited by Ip et al. (2012), Kember (1999), and Mezirow (1991). Other frameworks offer a more multidimensional and process-oriented understanding of reflection, such as those proposed by Gibbs (1988), Kolb (1984), and Mansvelder-Longayroux et al. (2007). These two types of theories can provide students with holistic and analytical assessments. Whatsmore, these theoretical underpinnings lend robustness to automated assessment methods, making them not merely algorithmic "black boxes" but interpretable and educationally meaningful tools.

Additionally, during our systematic review, we earmarked "Reflective Writing as an Assessment Tool" as a critical domain warranting further exploration within the realm of higher education. The primary impetus for integrating reflective writing as an assessment mechanism is twofold: first, to gauge students' performance meticulously, and second, to validate the effectiveness of a diverse array of educational interventions. Our systematic review robustly substantiates the utility of reflective writing as a versatile and effective assessment strategy. This methodology has proven instrumental in evaluating various educational outcomes, including but not limited to academic achievement (Tsingos-Lucas and Schneider, 2017), emotional self-regulation (Toivonen et al., 2017), and professional development (Azimi et al., 2019; Wise and Cui, 2019). One of the most compelling attributes of reflective writing is its inherent capability to marry surface-level and in-depth learning experiences seamlessly. This harmonized approach is especially laudable for its inclusivity, as it accommodates and recognizes the idiosyncrasies among individual learners. Complementing its multiple strengths, reflective writing exhibits considerable sustainability and transferability, thus rendering it applicable across diverse educational milieus. Continued engagement with this form of assessment yields manifold benefits by nurturing indispensable skills such as critical thinking, problem-solving aptitude, and self-management. Reflective writing in educational settings is subject to various challenges, including excessive subjectivity, ambiguous evaluation criteria, and the significant time investment required for instruction and assessment. However, our review suggests that advancements are being made to mitigate some of these issues. A noteworthy development pertains to the creation and validation of objective assessment scales explicitly designed to counterbalance the issue of excessive subjectivity. Additionally, the deployment of machine learning algorithms for automatic reflective writing assessment shows promise in reducing subjectivity and administrative burden. Nevertheless, there are inherent limitations that must be acknowledged. For instance, not all students possess the aptitude or willingness to engage in deep reflection, which can negatively impact their learning experience and assessment outcomes. Consequently, the pedagogical design of courses incorporating reflective writing assignments must be thoughtfully conceived. Strategies should focus not only on the development of more reliable assessment tools but also on

fostering a classroom environment that actively encourages student engagement in reflective practices.

4.3 Assessment methods for reflective writing in higher education

While qualitative content analysis has traditionally served as the benchmark for evaluating reflective writing in higher education, burgeoning advancements in ML and NLP present compelling alternatives. Qualitative content analysis has long been lauded for its ability to provide intricate and nuanced insights into reflective writings, a perspective underscored by studies such as Poldner et al. (2014). However, this methodological richness comes with the trade-off of labor-intensiveness and potential subjectivity, often necessitating considerable time and specialized expertise for effective implementation (Ullmann et al., 2019). Conversely, ML and NLP-based technologies offer the advantage of promptness, delivering immediate analyses. Beyond merely dissecting syntactic structures, these computational methodologies extend their analytical reach to multidimensional aspects like semantic interpretation and sentiment analysis, as highlighted in emerging studies like Solopova et al. (2023). Furthermore, our review substantiates that many AI functionalities designed for assessing reflective writing are anchored in solid theoretical frameworks. This adherence to theoretical rigor enhances the methodological robustness of AI-based assessments and lends them the capacity to standardize evaluations. In doing so, they can potentially ameliorate some of the subjectivity often associated with traditional qualitative content analysis.

However, adopting ML and NLP in this realm is challenging. These algorithms are primarily data-driven and depend heavily on large, annotated datasets for training (Nehyba and Štefánik, 2023). Ironically, generating these datasets often requires initial human annotation based on qualitative content analysis, forming a circular dependency. Moreover, current machine learning models may fall short in capturing the intricate cognitive processes involved in reflective thinking—a facet where qualitative approaches excel. Equally concerning is the issue of interpretability. While machine learning algorithms excel at providing quantitative assessments, they often obfuscate the underlying rationale. This lack of transparency can impede instructional decision-making, as educators and policy-makers might find it challenging to understand the "why" and "how" behind algorithmically generated evaluations. At the same time, the opaque nature of AI-based automated assessments can cause anxiety among students and may erode trust in the assessment process (Zhang et al., 2023).

5. Conclusion

This systematic review synthesizes the findings of 88 empirical studies and contributes significantly to ongoing research on reflective writing in higher education with respect to its application, themes, and assessment. The review highlights three main aspects: the integration of reflective writing in higher education, the important domains of current research, and the methodologies used for reflective writing evaluation. Firstly, our selection of eligible studies reveals that the application of reflective writing varies considerably across subject disciplines, applications of reflective writing, educational programs, and study durations. Most studies primarily employed one-semester-long theoretical courses as the instructional setting, with reflective writing assignments as the dominant approach. Most of these studies have focused on health professions education. Moreover, the scope of these studies extended beyond the mere evaluation of reflective writing itself. Researchers also investigated its utility in assessing educational outcomes, and a subset of these studies further examined factors contributing to the quality of such writing. Furthermore, a notable trend in assessment methodologies is the burgeoning application of AI-based approaches, particularly ML and NLP. These emerging technologies are gaining traction for their potential to offer rapid, scalable, and more objective evaluations (e.g., Ullmann et al., 2019; Nehyba and Štefánik, 2023; Wulff et al., 2023).

While this systematic review offers several insights into the application of reflective writing in higher education, it is crucial to acknowledge its limitations. Firstly, one key constraint pertains to the generalizability of the results. The search language for this study was restricted to English, thus primarily yielding research from English-speaking countries, most notably the United States. Furthermore, the broad scope of this review, encompassing methodologies, research themes, and assessment methodologies in higher education, inherently circumscribes the depth of discussion achievable for each aspect. While this expansive approach provides a comprehensive overview, it potentially dilutes the nuanced understanding of each subject area. Lastly, resource constraints resulted in the entire review process being conducted by a single researcher. Even though stringent inclusion-exclusion criteria were followed, and attempts were made to ensure objectivity, the single-researcher format inherently introduces the potential for personal bias. This limitation cannot be entirely mitigated.

The findings of this systematic review bear significant implications for the instructional application of reflective writing within higher education contexts. Firstly, the results reveal that reflective writing is predominantly concentrated within single-semester theory-based courses and mainly focused on health professions education. Such a focus suggests that educators and curriculum designers should contemplate expanding reflective writing into other disciplinary realms and consider its longitudinal integration across multiple semesters. Secondly, the review substantiates the efficacy of reflective writing as a robust mechanism for assessing student performance and the impact of various educational interventions. Hence, our findings posit reflective writing as a potential multidimensional assessment tool that can transcend traditional evaluation forms. This insight encourages educational stakeholders to reconsider the scope and dimensions in which reflective writing can serve as an alternative or complementary assessment method. Lastly, the emergence of AI-based assessment tools in reflective writing offers promising opportunities for personalized instruction and scalable evaluation.

In light of the findings garnered from this systematic review, several avenues for future research emerge as

particularly salient. First of all, while the current inves-tigation substantiates the value and efficacy of reflective writing within short-term course frameworks, there must be a discernible gap in understanding its impact on longitudinal learning outcomes and professional development trajectories. Future studies could extend this line of inquiry by longitudinally assessing students' academic and vocational performance after completing reflective writing tasks. Such endeavors would elucidate the long-term dividends accrued from this instructional strategy. Next, there exists an opportunity for the liter-ature to broaden its scope by examining factors that influence reflective writing at various levels—personal, interpersonal, and social. A nuanced investigation into these aspects will deepen our comprehension of what constitutes high-quality reflective writing. Last but not least, the advent of ML and NLP in the assessment landscape of reflective writing is a promising yet under-investigated domain. Upcoming research could contribute by scrutinizing the effectiveness and lim-itations of these emergent technologies. Developing and validating algorithmic models designed to more accurately discern the key elements and quality indicators intrinsic to reflective writing would be of particular interest. This could mitigate limitations and promote greater transparency and reliability in AI-driven assessment mechanisms.

Data Availability Statement

The dataset generated and analyzed during the current study is available from the corresponding author upon request.

Acknowledgments

Conflicts of interest

The authors declare no conflict of interest.

Use of AI and AI-assisted Technologies

No AI tools were utilized for this paper.

Open Access

References

Adams, J., Ari, M., Cleeves, M., & Gong, J. (2020). Reflective writing as a window on medical students' professional identity development in a longitudinal integrated clerkship. *Teaching and Learning in Medicine*, *32*(2), 117-125. https://doi.org/10.1080/10401334.2019.1687303

Alt, D., Raichel, N., & Naamati-Schneider, L. (2022). Higher education students' reflective journal writing and lifelong learning skills: insights from an exploratory sequential study. *Frontiers in Psychology*, *12*, 707168. https:// doi.org/ 10.3389/fpsyg.2021.707168

Alrashidi, H., Ullman, T. D., & Joy, M. (2020). An empirical evaluation of a Reflective Writing Framework (RWF) for Reflective Writing in Computer Science Education. In *2020 IEEE Frontiers in Education Conference* (FIE) (pp. 1–9). IEEE. https://doi.org/10.1109/FIE44824.2020.9273975

Alrashidi, H., Ullman, T. D., & Joy, M. (2020). An empirical evaluation of a Reflective Writing Framework (RWF) for Reflective Writing in Computer Science Education. In *2020 IEEE Frontiers in Education Conference* (FIE) (pp. 1–9). IEEE. https://doi.org/10.1109/FIE44824.2020.9273975

Alsina, Á., Ayllón, S., & Colomer, J. (2019). Validating the Narrative Reflection Assessment Rubric (NARRA) for re-flective narratives in higher education. *Assessment & Evaluation in Higher Education*, *44*(1), 155–168. https://doi.org/10.1080/02602938.2018.1486391

Alsina, Á., Ayllón, S., Colomer, J., Fernández-Peña, R., Fullana, J., Pallisera, M., Pérez-Burriel, M., & Serra, L. (2017). Improving and evaluating reflective narratives: A rubric for higher education students. *Teaching and Teacher Ed-ucation*, *63*, 148–158. https://doi.org/10.1016/j.tate.2016.12.015

Azimi, E., Kuusisto, E., Tirri, K., & Hatami, J. (2019). How do student teachers reflect on their practice through practicum courses? A case study from Iran. *Journal of Education for Teaching*, *45*(3), 277–289. https://doi.org/10.1080/0958 9236.2019.1599511

Beigman Klebanov, B., Burstein, J., Harackiewicz, J. M., Priniski, S. J., & Mulholland, M. (2017). Reflective Writing About the Utility Value of Science as a Tool for Increasing STEM Motivation and Retention – Can AI Help Scale Up? *International Journal of Artificial Intelligence in Education*, *27*(4), 791–818. https://doi.org/10.1007/s40593-017-0141-4

Bell, A., Kelton, J., McDonagh, N., Mladenovic, R., & Morrison, K. (2011). A critical evaluation of the usefulness of a coding scheme to categorise levels of reflective thinking. *Assessment & Evaluation in Higher Education*, *36*(7), 797–815. https://doi.org/10.1080/02602938.2010.488795

Blatt, B., Plack, M., Maring, J., Mintz, M., & Simmens, S. J. (2007). Acting on reflection: The effect of reflection on students' clinical performance on a standardized patient examination. *Journal of General Internal Medicine*, *22*, 49-54. https://doi.org/10.1007/s11606-007-0110-y

Boud, D., Keogh, R., & Walker, D. (Eds.). (2013). *Reflec-tion: Turning experience into learning*. Routledge.

Bowman, M. (2021). A framework for scaffolding academic reflective writing in dentistry. *European Journal of Dental Education: Official Journal of the Association for Dental Education in Europe*, *25*(1), 35–49. https://doi.org/10.1111/eje.12575

Braun, U. K., Gill, A. C., Teal, C. R., & Morrison, L. J. (2013). The utility of reflective writing after a palliative care experience: Can we assess medical students' professionalism? *Journal of Palliative Medicine*, *16*(11), 1342–1349. https://doi.org/10. 1089/jpm.2012.0462

Brown, A., Jauregui, J., Ilgen, J. S., Riddell, J., Schaad, D., Strote, J., & Shandro, J. (2020). Does the medium matter? Evaluating the depth of reflective writing by medical students on social media compared to the traditional private essay using the REFLECT rubric. *Western Journal of Emergency Medi-cine*, *21*(1), 18. https://doi.org/10.5811/westjem.2019.11.44263

Bruno, A., Galuppo, L., & Gilardi, S. (2011). Evaluating the reflexive practices in a learning experience. *European Journal of Psychology of Education*, *26*, 527-543. https://doi.org/10.1007/s10212-011-0061-x

Burnett, A. J., & Akerson, E. (2019). Preparing future health professionals via reflective pedagogy: a qualitative instrumental case study. *Reflective Practice*, *20*(5), 571-583. https://doi.org/10.1080/14623943.2019.1651711

Cengiz, C., Karataş, F. Ö., & Yadigaroğlu, M. (2014). The Investigation of Pre-service Science Teachers' Reflective Journals. *Procedia - Social and Behavioral Sciences*, *116*, 3297–3302. https://doi.org/10.1016/j.sbspro.2014.01.751

Chen, I., & Forbes, C. (2014). Reflective writing and its impact on empathy in health professions education: systematic review. *Journal of Educational Evaluation for Health Professions*, *11*. https://doi.org/10.3352/jeehp.2014.11.20

Chen, Y., Yu, B., Zhang, X., & Yu, Y. (2016). Topic modeling for evaluating students' reflective writing. In D. Gašević, G. Lynch, S. Dawson, H. Drachsler, & C. Penstein Rosé (Eds.), *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 1–5). ACM Press. https://doi.org/10.1145/2883851.2883951

Chong, C., Sheikh, U. U., Samah, N. A., & Sha'ameri, A. Z. (2020). Analysis on Reflective Writing Using Natural Language Processing and Sentiment Analysis. *IOP Conference Series: Materials Science and Engineering*, *884*(1), 12069. https://doi.org/10.1088/1757-899X/884/1/012069

Christof, E., Köhler, J., Rosenberger, K., & Wyss, C. (2018). *Mündliche, schriftliche und theatrale Wege in der Praxisreflexion: Beiträge zur Professionalisierung pädagogischen Handelns*. hep der Bildungsverlag. https://doi.org/10.5281/zenodo.1451400

Chua, I. S., Bogetz, A. L., Bhansali, P., Long, M., Holbreich, R., Kind, T., Ottolini, M., Park, Y. S., Lineberry, M., & Hirshfield, L. E. (2019). The patient experience debrief interview: how conversations with hospitalized families influence medical student learning and reflection. *Academic Medicine: Journal of the Association of American Medical Colleges*, *94*(11S Association of American Medical Colleges Learn Serve Lead: Proceedings of the 58th Annual Research in Health professions education Sessions), S86-S94. https://doi.org/10.1097/ACM.0000000000002914

Cochran, G. L., Brookes, D. T., & Kramer, L. H. (2013). A framework for assessing learning assistants' reflective writing assignments. In *AIP Conference Proceedings* (pp. 15–18). AIP. https://doi.org/10.1063/1.4789640

Coleman, D., & Willis, D. S. (2015). Reflective writing: The student nurse's perspective on reflective writing and poetry writing. *Nurse Education Today*, *35*(7), 906-911. https://doi.org/10.1016/j.nedt.2015.02.018

Cui, Y., Wise, A. F., & Allen, K. L. (2019). Developing reflection analytics for health professions education: A multidimensional framework to align critical concepts with data features. *Computers in Human Behavior*, *100*, 305–324. https://doi.org/10. 1016/j.chb.2019.02.019

Demmans Epp, C., Akcayir, G., & Phirangee, K. (2019). Think twice: exploring the effect of reflective practices with peer review on reflective writing and writing quality in computer-science education. *Reflective Practice*, *20*(4), 533–547. https://doi.org/10.1080/14623943.2019.1642189

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, 1, 4171-4186. https://doi.org/10.48550/arXiv.1810.04805

Dors, T. M., Van Amstel, F. M., Binder, F., Reinehr, S., & Malucelli, A. (2020). Reflective Practice in Software Development studios: findings from an ethnographic study. In *2020 IEEE 32nd Conference on Software Engineering Education and Trai-ning* (pp. 1-10). IEEE. https://doi.org/10.1109/CSEET49119.2020.9206217

Dressler, J. A., Ryder, B. A., Monteiro, K., Cheschi, E., Connolly, M., Miner, T. J., & Harrington, D. T. (2019). "Tweet"-format reflective writing: a hidden needs assessment? *The American Journal of Surgery*, *217*(2), 314-317. https://doi.org/10.1016/j.amjsurg.2018.09.018

Dunne, J. (2019). Improved levels of critical reflection in Pharmacy Technician student work-placement assessments through emphasising graduate attri-butes. *Journal of Teaching and Learning for Graduate Employability*, *10*(2), 1–14. https://doi.org/10.21153/jtlge2019vol10no2art637

Fan, X., Luo, W., Menekse, M., Litman, D., & Wang, J. (2017). Scaling Reflection Prompts in Large Classrooms via Mobile Interfaces and Natural Language Processing. In G. A. Papadopoulos, T. Kuflik, F. Chen, C. Duarte, & W.-T. Fu (Eds.), *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (pp. 363–374). ACM. https://doi.org/10.1145/3025171.3025204

Findyartini, A., Anggraeni, D., Husin, J. M., & Greviana, N. (2020). Exploring medical students' professional identity formation through written reflections during the COVID-19 pandemic. *Journal of Public Health Research*, *9*, jphr-2020. https://doi.org/10.4081/jphr.2020.1918

Fischer, M. A., Haley, H.-L., Saarinen, C. L., & Chretien, K. C. (2011). Comparison of blogged and written reflections in two medicine clerkships. *Health Professions Education*, *45*(2), 166–175. https://doi.org/10.1111/j.1365-2923.2010.03814.x

Franco, R. S., dos Santos Franco, C. A. G., Severo, M., Ferreira, M. A., & Karnieli-Miller, O. (2022). Reflective writing in the teaching of communication skills for medical students—a systematic review. *Patient Education and Counseling*, *105*(7), 1842-1851. https://doi.org/10.1016/j.pec.2022.01.003

Gibbs, G. 1988. *Learning by doing: A guide to teaching and learning methods*. Oxford University Press.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). Reflective writing analytics for actionable feedback. In A. Wise, P. H. Winne, G. Lynch, X. Ochoa, I. Molenaar, S. Dawson, & M. Hatala (Eds.), *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 153–162). ACM. https://doi.org/10.1145/3027385.3027436

Gibson, A., Kitto, K., & Bruza, P. (2016). Towards the discovery of learner metacognition from reflective writing. *Journal of Learning Analytics*, *3*(2), 22–36. https://doi.org/10.18608/jla.2016.32.3

Gläser-Zikuda, M., Hagenauer, G., & Stephan, M. (2020). The potential of qualitative content analysis for empirical educational research. In *Forum: Qualitative Social Research* (Vol. 21, No. 1, p. 20). DEU. https://doi.org/10.17169/fqs-21.1.3443

Hanlon, C. D., Frosch, E. M., Shochet, R. B., Buckingham Shum, S. J., Gibson, A., & Goldberg, H. R. (2021). Recognizing Reflection: Computer-Assisted Analysis of First Year Medical Students' Reflective Writing. *Medical Science Educator*, *31*(1), 109–116. https://doi.org/10.1007/s40670-020-01132-7

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Houston, C. R. (2016). Do Scaffolding Tools Improve Reflective Writing in Professional Portfolios? A Content Analysis of Reflective Writing in an Advanced Preparation Program. *Action in Teacher Education*, *38*(4), 399–409. https://doi.org/10.1080/01626620.2016.1226201

Ip, W. Y., Lui, M. H., Chien, W. T., Lee, I. F., Lam, L. W., & Lee, D. (2012). Promoting self-reflection in clinical practice among Chinese nursing underg-raduates in Hong Kong. *Contemporary nurse*, *41*(2), 253-262. https://doi.org/10.5172/conu. 2012.41.2.253

Isaac, C., Behar-Horenstein, L., Lee, B., & Catalanotto, F. (2015). Impact of Reflective Writing Assignments on Dental Students' Views of Cultural Competence and Diversity. *Journal of Dental Education*, *79*(3), 312–321. https://doi.org/10.1002/j.0022-0337.2015.79.3.tb05886.x

Jung, Y., & Wise, A. F. (2020). How and how well do students reflect? Multi-dimensional automated reflection assessment in health professions education. In *LAK 2020 Conference Proceedings - Celebrating 10 years of LAK: Shaping the Future of the Field - 10th International Conference on Learning Analytics and Knowledge* (pp. 595-604). (ACM International Conference Proceeding Series). Association for Computing Machinery. https://doi.org/10.1145/3375462.3375528

Jung, Y., Wise, A. F., & Allen, K. L. (2022). Using theory-informed data science methods to trace the quality of dental student reflections over time. *Advances in Health Sciences Education: Theory and Practice*, *27*(1), 23–48. https://doi.org/10.1007/s10459-021-10067-6

Kember, D. (1999). Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *International Journal of Lifelong Edu-cation*, *18*(1), 18-30. https://doi.org/10.1080/026013799293928

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall

Korthagen, F. A. J., & Kessels, J. P. A. M. (1999). Linking Theory and Practice: Changing the Pedagogy of Teacher Education. *Educational Researcher, 28*(4), 4–17. https://doi.org/10.3102/0013189x028004004

Korthagen, F. A., Kessels, J., Koster, B., Lagerwerf, B., & Wubbels, T. (2001). *Linking practice and theory: The pedagogy of realistic teacher education*. Rout-ledge.

Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018). Understand students' self-reflections through learning analytics. In A. Pardo, K. Bartimote-Aufflick, G. Lynch, S. B. Shum, R. Ferguson, A. Merceron, & X. Ochoa (Eds.), *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 389–398). ACM. https://doi.org/10.1145/3170358.3170374

Lachman, N., & Pawlina, W. (2006). Integrating professionalism in early health professions education: the theory and application of reflective practice in the anatomy curriculum. *Clinical Anatomy*, *19*(5), 456-460. https://doi.org/10.1

002/ca.20344

Lin, C.-W., Lin, M.-J., Wen, C.-C., & Chu, S.-Y. (2016). A word-count approach to analyze linguistic patterns in the reflective writings of medical students. *Health Professions Education Online*, *21*, 29522. https://doi.org/10.3402/meo.v21.29522

Liu, M., Kitto, K., & Buckingham Shum, S. (2021). Combining factor analysis with writing analytics for the formative assessment of written reflection. *Computers in Human Behavior*, *120*, 106733. https://doi.org/10.1016/j.chb.2021.106733

Liu, M., Shum, S. B., Mantzourani, E., & Lucas, C. (2019). Evaluating Machine Learning Approaches to Classify Pharmacy Students' Reflective Statements. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Lecture Notes in Computer Science. Artificial Intelligence in Education* (Vol. 11625, pp. 220–230). Springer International Publishing. https://doi.org/10.1007/978-3-030-23204-7_19

Lucas, C., Smith, L., Lonie, J. M., Hough, M., Rogers, K., & Mantzourani, E. (2019). Can a reflective rubric be applied consistently with raters globally? A study across three countries. *Currents in Pharmacy Teaching & Learning*, *11*(10), 987–994. https://doi.org/10.1016/j.cptl.2019.06.004

Mamede, S., Schmidt, H. G., & Penaforte, J. C. (2008). Effects of reflective practice on the accuracy of medical diagnoses. *Health Professions Education*, *42*(5), 468-475. https://doi.org/10.1111/j.1365-2923.2008.03030.x

Mansvelder-Longayroux, D. D., Beijaard, D., & Verloop, N. (2007). The portfolio as a tool for stimulating reflection by student teachers. *Teaching and Teacher education*, *23*(1), 47-62. https://doi.org/10.1016/j.tate.2006.04.033

Mayne, L. (2012). Reflective writing as a tool for assessing teamwork in bioscience: Insights into student performance and understanding of teamwork. *Biochemistry and Molecular Biology Education*, *40*(4), 234-240. https://doi.org/10.1002/bmb.20621

Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in Health Sciences education*, *14*, 595-621. https://doi.org/10.1007/s10459-007-9090-2

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space.* https://doi.org/10.48550/arXiv.1301.37 81

Miller-Kuhlmann, R., O'Sullivan, P. S., & Aronson, L. (2016). Essential steps in developing best practices to assess reflective skill: A comparison of two rubrics. *Medical Teacher*, *38*(1), 75–81. https://doi.org/10.3109/0142159X.2015.1034662

Mitchell, K. M., Roberts, T., & Blanchard, L. (2021). Reflective writing pedagogies in action: a qualitative systematic review. *International journal of nursing education scholarship*, *18*(1). https://doi.org/10.1515/ijnes-2021-0057

Moniz, T., Arntfield, S., Miller, K., Lingard, L., Watling, C., & Regehr, G. (2015). Considerations in the use of reflective writing for student assessment: Issues of reliability and validity. *Health Professions Education*, *49*(9), 901–908. https://doi.org/10.1111/medu.12771

Moon, J. A. (2013). *Reflection in learning and professional development: Theory and practice*. Rout-ledge.

Naeger, D. M., Hua, E. W., Ahearn, B., & Webb, E. M. (2015). Reflective writing: a potential tool to improve interprofessional teamwork with radiologists. *Academic Radiology*, *22*(10), 1221-1225. https://doi.org/10.1016/j.acra.2015.07.006

Nehyba, J., & Štefánik, M. (2023). Applications of deep language models for reflective writings. *Education and Information Technologies*, *28*(3), 2961-2999. https://doi.org/10.1007/s10639-022-11254-7

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176-1189. https://doi.org/10.1111/medu.12583

Núñez-Canal, M., de Obesso, M. D. L. M., & Pérez-Rivero, C. A. (2022). New challenges in higher education: A study of the digital competence of educators in Covid times. *Technological Forecasting and Social Change*, *174*, 121270. https://doi. org/10.1016/j.techfore.2021.121270

Ottenberg, A. L., Pasalic, D., Bui, G. T., & Pawlina, W. (2016). An analysis of reflective writing early in the medical curriculum: The relationship between reflective capacity and academic achievement. *Medical Teacher*, *38*(7), 724–729. https://doi.org/10.3109/0142159X.2015.1112890

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 89. https://doi.org/10.1186/s13643-021-01626-4

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*. http://hdl.handle.net/2152/31333

Poldner, E., van der Schaaf, M., Simons, P. R.-J., van Tartwijk, J., & Wijngaards, G. (2014). Assessing student teachers' reflective writing through quantitative content analysis. *European Journal of Teacher Education*, *37*(3), 348–373. https://doi.org/10.1080/02619768.2014.892479

Poon, L. K. M., Li, Z., & Cheng, G. (2017). Topic Classification on Short Reflective Writings for Monitoring Students' Progress. In S. K. Cheung, L. Kwok, W. W. Ma, L.-K. Lee, & H. Yang (Eds.), *Lecture Notes in Computer Science. Blended Learning. New Challenges and Innovative Practices* (Vol. 10309, pp. 236–246). Springer International Publishing. https://doi.org/10.1007/978-3-319-59360-9_21

Pryjmachuk, S., McWilliams, C., Hannity, B., Ellis, J., & Griffiths, J. (2019). Transitioning to university as a nursing student: Thematic analysis of written reflections. *Nurse Education Today*, *74*, 54–60. https://doi.org/10.1016/j.nedt.2018.12.003

Rogers, R.R. (2001). Reflection in Higher Education: A Concept Analysis. *Innovative Higher Education, 26*, 37-57. https://doi.org/10.1023/A:1010986404527

Rojí, R., Noguera-Tejedor, A., Pikabea-Díaz, F., Carrasco, J. M., & Centeno, C. (2017). Palliative care bedside teaching: A qualitative analysis of medical students' reflective writings after clinical practices. *Journal of Palliative Medicine*, *20*(2), 147-154. https://doi.org/10.1089/jpm.2016.0192

Ryan, M., & Ryan, M. (2013). Theorizing a model for teaching and assessing reflective learning in higher education. *Higher Education Research & Development, 32,* 244-257. https://doi.org/10.1080/07294360.2012.661704

Savicki, V., & Price, M. V. (2015). Student Reflective Writing: Cognition and Affect Before, During, and After Study Abroad. *Journal of College Student Development*, *56*(6), 587–601. https://doi.org/10.1353/csd.2015.0063

Savicki, V., & Price, M. V. (2021). Reflection in Transformative Learning: The Challenge of Measurement. *Journal of Transformative Education*, *19*(4), 366–382. https://doi.org/10.1177/15413446211045161

Schön, D. A. (1983). *The reflective practitioner*. Jossey-Bass.

Shum, S. B., Sándor, Á., Goldsmith, R., Wang, X., Bass, R., & McWilliams, M. (2016). Reflecting on reflective writing analytics. In D. Gašević, G. Lynch, S. Dawson, H. Drachsler, & C. Penstein Rosé (Eds.), *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 213–222). ACM Press. https://doi.org/10.1145/2883851.2883955

Solopova, V., Gruszczynski, A., Rostom, E., Cremer, F., Witte, S., Zhang, C., ... & Landgraf, T. (2023). *PapagAI: Automated Feedback for Reflective Essays.* https://doi.org/10.48550/arXiv.2307.0752

Springer, D. G., & Yinger, O. S. (2019). Linguistic Indicators of Reflective Practice Among Music Education Majors. *Journal of Music Teacher Education*, *28*(2), 56–69. https://doi.org/10.1177/1057083718786739

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, *27*. https://doi.org/10.48550/arXiv.1409.3215

Talwalkar, J. S., Moriarty, J. P., & Ellman, M. S. (2019). Students' experiences with death and dying prior to medical school: A content analysis of students' written reflections. *The American Journal of Hospice & Palliative Care*, *36*(11), 999–1007. https://doi.org/10.1177/1049909119847965

Taylor-Haslip, V. (2010). Guided reflective journals depict a correlation to the academic success of nursing students. *Teaching and Learning in Nursing*, *5*(2), 68–72. https://doi.org/10.1016/j.teln.2010.01.002

Toivonen, A. K., Lindblom-Ylänne, S., Louhiala, P., & Pyörälä, E. (2017). Medical students' reflections on emotions concerning breaking bad news. *Patient Education and Counseling*, *100*(10), 1903-1909. https://doi.org/10.1016/j.pec.2017.05.036

Tsingos-Lucas, C., Bosnic-Anticevich, S., Schneider, C. R., & Smith, L. (2017). Using Reflective Writing as a Predictor of Academic Success in Different Assessment Formats. *American Journal of Pharmaceutical Education*, *81*(1), 8. https://doi.org/10.5688/ajpe8118

Ullmann, T. D. (2017). Reflective writing analytics. In A. Wise, P. H. Winne, G. Lynch, X. Ochoa, I. Molenaar, S. Dawson, & M. Hatala (Eds.), *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 163–167). ACM. https://doi.org/10.1145/3027385.3027394

Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, *29*(2), 217-257. https://doi.org/10.1007/s40593-019-00174-2

van Beveren, L., Roets, G., Buysse, A., & Rutten, K. (2018). We all reflect, but why? A systematic review of the purposes of reflection in higher education in social and behavioral sciences. *Educational Research Review*, *24*, 1–9. https://doi.org/10.1016/j.edurev.2018.01.002

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 5998-6008. https://doi.org/10.48550/arXiv.1706.03762

Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P. (2012). Fostering and evaluating reflective capacity in health professions education: Developing the REFLECT rubric for assessing reflective writing. *Academic Medicine: Journal of the Association of American Medical Colleges*, *87*(1), 41–50. https://doi.org/10.1097/ACM.0b013e31823b55fa

Wise, A. F., & Cui, Y. (2019). Top Concept Networks of Professional Education Reflections. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 260–264). ACM. https://doi.org/10.1145/3303772.3303840

Wise, A.F., Reza, S. & Han, R. J. (2020). Becoming a dentist: Tracing professional identity development through mixed-methods data mining of student reflections. In *Proceedings of the 13th international conference of the learning sciences.* Nashville, TN: International Society of the Learning Sciences. https://doi.dx.org/10.22318/icls2020.294

Woldt, J. L., & Nenad, M. W. (2021). Reflective writing in dental education to improve critical thinking and learning: A systematic review. *Journal of Dental Education*, *85*(6), 778-785. https://doi.org/10.1002/jdd.12561

Wright, L., & Lundy, M. (2012). Blogging as a tool to promote reflection among dietetic and physical therapy students during a multidisciplinary international service-learning experience. *Journal of Allied Health*, *41*(3), 73E-78E. https://pubmed.ncbi.nlm.nih.gov/22968780/

Wulff, P., Buschhüter, D., Westphal, A., Nowak, A., Becker, L., Robalino, H., ... & Borowski, A. (2021). Computer-based classification of preservice physics teachers' written reflections. *Journal of Science Education and Technology*, *30*, 1-15. https://doi.org/10.1007/s10956-020-09865-1

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2023). Utilizing a pretrained language model (BERT) to classify preservice physics teachers' written reflections. *International Journal of Artificial Intelligence in Education*, *33*(3), 439-466. https://doi.org/10.1007/s40593-022-00290-6

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education*, *45*, 477-501. https://doi.org/10.1023/ A:1023967026413

Young, K., James, K., & Noy, S. (2016). Exploration of a reflective practice rubric. *Asia-Pacific Journal of Cooperative Education*, *17*(2), 135-147. https://www.ijwil.org/files/APJCE_17_2_135_147.pdf

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: a multigroup analysis. *International Journal of Educational Technology in Higher Education*, *20*(1), 49. https:// doi.org/10.1186/s41239-023-00420-7