EXPLORING 50 YEARS OF RESEARCH ON STUDENTS' IDENTITIES IN HIGHER EDUCATION CONTEXT: A BIBLIOMETRIC ANALYSIS

[1]Nadia Tun & [*2]Zulaikha Khairuddin

[1] Centre of English Language Studies, Faculty of Languages and Communication, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Nerus, Terengganu, Malaysia.

[2] Academy of Language Study, Universiti Teknologi MARA, Shah Alam Campus, 40450 Shah Alam, Selangor, Malaysia.

*Corresponding author: zulaikha5497@uitm.edu.my

ABSTRACT

Background and Purpose: Higher education serves not merely as a means of imparting knowledge, skills, and societal values to students; it also offers individuals the chance to explore the world, discover their identity and position, gain a deeper understanding of others, and develop appropriate responses to diverse perspectives. Hence, it is vital to acknowledge and understand the significance of students' identities as it is not an inherent or fixed concept; instead, it emerges and transforms through diverse academic experiences encountered by students from all around the world. The objective of this study is to examine various methods to identify the global research trends related to students' identities in higher education. This bibliometric study examines various methods to identify the global research trends related to students' identities in higher education, including annual publication output, leading countries, prolific authors, and co-occurring author keywords.

Methodology: A bibliometric study was conducted by utilising VOSviewer software (version 1.6.18) in comprehending the worldwide studies on students' identities in higher education. The study analysed publication trends from as far back as 1974 up to 2022, using data from the Scopus database. Data collection took place between July 1 and July 5, 2023, using the Scopus database. Data collection occurred between July 1 and July 5, 2023. The study examined publication trends, including time, author, and institution mapping, keyword clustering, and identifying emerging terms.

Findings: The findings revealed a notable increase in researchers' attention towards studying students' identities in higher education, reflected in a rise in related articles. Various theoretical frameworks, including Community of Practice, Acculturation, Self-Concept, Habitus, and Situated Learning, have been extensively examined to provide insights into understanding these identities. These frameworks underscore the impact of social, cultural, and psychological factors on how students perceive themselves and engage within academic settings.

Contributions: This bibliometric analysis enhances our understanding of students' identities in higher education, guiding future research and providing useful advice for professionals, researchers, and policymakers in the educational sector. Exploring students' identities offers valuable insights into collaborative networks and potential partners, facilitating further advancements in the field.

Keywords: Students' identities, higher education, bibliometrics analysis.

1.0 INTRODUCTION

Identity encompasses the significance of meaning ascribed to oneself from both personal introspection and external perspectives. This intricate concept allows for the existence of various facets of identity, such as personal, social, and professional identities (Ibarra & Petriglieri, 2010). Nias (1989) claimed that individuals have fixed, static and stable identities which was also mentioned by Berkman et al. (2017) within the field of psychology and education. However, according to Gee (2003), identity is always constructed and negotiated in the different contexts individuals are in, hence, it is dynamic and flexible (Nadia & Zurina, 2023; Park, 2007; Hawkins, 2005; Kasworm, 2010; Lee, 2003; Côté & Levine, 2002; Antaki, 2008). Gee (2001) and Brown et al. (2005) argue that identities could be discursively constructed (Bamberg et al., 2011; Ruelle & Peverelli, 2017; Bös et al., 2018) depending on contexts they are in (Nadia & Zurina, 2023; Virta et al., 2023; Chen et al., 2022; Seyri & Ali Rezaee, 2022; Antaki, 2008; Lee, 2003) or the social interaction among the members of the society as asserted by (Nadia & Zurina, 2023; Ubaidillah & Widiati, 2021; Farland-Smith, 2012; Lee, 2003; Norton, 1997). It is also how individuals view themselves and others from multiple perspectives and environments (Nadia & Zurina, 2023; Virta et al., 2023; Chen et al., 2022; Seyri & Ali Rezaee, 2022; Archer, 2003; Lee, 2003; Benwell & Stokoe, 2006). The

process of identity construction and negotiation refers to the process by which individuals may construct new identities, return to old ones or transform to new ones (Nadia & Zurina, 2023; Ruelle & Peverelli, 2017; Bös et al., 2018; Jenkins, 2014) which also progresses with time (Norton, 1997). Hence, identity can be described as dynamic, continuous and always changing (Nadia & Zurina, 2023; Park, 2007) as it depends on how individuals position themselves in relation to their experiences (Nadia & Zurina, 2023; Wenger, 1998; Gee, 2001; Grad & Martín Rojo, 2008).

The discussion above is in line with the constructivism approach which perceives identity as having individuals to consistently construct and negotiate their identities during social interactions. Rajadurai (2010a, 2010b) views students as individuals who attempt to make sense of meaning when they interact in relation to themselves and others as they construct and negotiate their identities. Students may experience an ongoing process, whether intentionally or not, in constructing and negotiating their identities as they almost always either interact (or not interact) with other individuals they speak to in the classrooms such as lecturers and other students (Giroir, 2014; Winchester, 2013). Ting-Toomey (1999) also suggests that students may attempt to negotiate their identities as they may or may not assert, define, modify, challenge and/or support how they see themselves in position of themselves and others' (Nadia & Zurina, 2023; Ruelle & Peverelli, 2017; Bös et al., 2018; Jenkins, 2014), particularly as they engage when they are the classrooms, which is considered as a Community of Practice (Nadia & Zurina, 2023; Lantolf & Pavlenko, 2001; Norton & Toohey, 2001). This means that as students interact with teachers or other students, they may gain competence and consequently, negotiate their identities.

2.0 LITERATURE REVIEW

The study of students' identities in higher education is a dynamic and diverse area of research that explores different facets of identity, such as race, ethnicity, gender, sexual orientation, socioeconomic status, and academic affiliations. These investigations seek to comprehend how students' identities impact their encounters, academic achievements, and overall well-being within the realm of higher education. These are some significant research domains associated with students' identities in higher education. Some of the studies conducted on students' identities in higher education emphasise on the context of postgraduate students (Mili & Towers, 2022), high school students (Daher & Shahbari, 2020) and student teachers (Winchester, 2013). Within the context of classrooms, studies were conducted either in multicultural classrooms (Mili & Towers, 2022) or English language classrooms (Zacharias,

2010; Winchester, 2013) or virtual classrooms (Daher & Shahbari, 2020) or online forums (Baxter & Haycock, 2014). There are also studies conducted either in the context of classroom presentations (Gray, 2022), active learning environments and social networking sites (Daher, 2012). Some studies focus on the roles of others or how others view students' identities, for instance, Rumenapp (2016) looks at how teachers view students' identities while Winchester (2013) focuses on the role teachers have on students' identities construction.

As students' identities are constructed and negotiated through socialisation, it is important to discuss and acknowledge the process and elements involved in students' identities construction and negotiation higher education. The process of constructing and negotiating students' identities in higher education is influenced by socialisation. Vygotsky's (1978) Sociocultural Theory (SCT) posits that the mind is socially constructed, influenced by social interactions and internalised experiences. Learning, according to SCT, is a social phenomenon shaped by culture, history, and institutions. SCT suggests that individuals' thinking, and abilities develop through social interactions, with these interactions being internalized and assimilated. In the context of the study, SCT is viewed as social discursive practices influencing identity construction and negotiation. Understanding social structure and cultural interpretations in Sociocultural Theory (SCT) involves examining novices' practical engagements with others, particularly in academic settings (Duranti et al., 2012). Academic Discourse Socialisation (Duff, 2007) (ADS) plays a key role, where members blend into academic discourse communities through interactions and cognitive experiences. Successful socialisation leads to increased participation, adaptability, and community membership. Students' experiences in higher education contribute to their socialisation process within academic contexts, involving negotiation of power and identity.

3.0 RESEARCH DESIGN

This research employs a bibliometric methodology to thoroughly examine scholarly papers pertaining to students' identities in higher education. By using bibliometrics, the study quantitatively investigates patterns and trends within the academic literature on students' identities in higher education, providing a detailed understanding of the field's present condition and potential future trajectories (Jiang et al., 2019). Bibliometric analysis has experienced a significant surge in adoption within the realm of academic research in recent years (Donthu et al., 2020a; Donthu et al., 2021; Khan et al., 2021). This growing prominence can be primarily ascribed to several factors, including the continual advancements in bibliometric software tools such as Gephi, Leximancer, and VOSviewer, alongside the

increased availability and accessibility of scientific databases like Scopus and Web of Science. These technological and informational advancements have notably facilitated the undertaking of bibliometric studies, enabling researchers to delve deeper into the complexities of scholarly discourse and extract valuable insights with greater precision and efficiency.

In the present study, a bibliometric study was conducted by utilising VOSviewer software (version 1.6.18) in comprehending the worldwide studies on students' identities in higher education. Although there has been an interest in investigating students' identities in higher education for the last 25 years, very few studies have been published to investigate the trend based on the global scientific publication perspectives. The objective of this study is to examine various methods to identify the global research trends related to students' identities in higher education. Although the Scopus and WoS databases are widely known to share a significant amount of journal indexing, there are specific journals that have distinct indexing in each of these databases. Currently, Scopus is known for its extensive database, surpassing WoS in abstract and citations of peer-reviewed publications coverage across various scientific fields (Chadegani et al., 2013). On the other hand, Google Scholar, which also boasts an extensive document collection, has two notable drawbacks: it can't specify the number of conference articles, and its search results fluctuate based on user-selected search options. Hence, opting for Scopus as the data source for this study ensures a comprehensive coverage of topics that might not be fully represented in WoS alone. The paper aims to:

i. Analyse the temporal distribution of students' identities in higher education journal articles and conference papers.
ii. Present leading authors, countries, and their academic institutions or affiliations in journal articles on students' identities in higher education.
iii. Highlight frequently recurring keywords in journal articles on students' identities in higher education context.
iv. Provide insights for collaborations and future trends in researching on students' identities in higher education.

3.1 Data Source and Search Strategy

The study analysed publication trends from as far back as 1974 up to 2022, using data from the Scopus database. This type of review provides a unique perspective in grasping the significance of a research trend, diverging from the conventional review articles that mainly concentrate on discussing the latest developments and future recommendations in a particular research field.

Data collection took place between July 1 and July 5, 2023, using the Scopus database. The primary focus of gathering articles was centred around the term "students' identities," and "higher education," which had to be present in both the title and abstract. The articles ranged from the oldest one published in 1974 to the most recent one in 2022. To achieve this, a search string (Ng et al., 2022) was constructed as follows: TITLE-ABS ("identit*" AND student* AND higher education) AND (EXCLUDE (PUBYEAR,2023)) AND (LIMIT-TO (PUBSTAGE,"final")) AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (SRCTYPE,"j")). Initially, this search yielded a total of 2192 articles. However, to exclude review articles from the study, an additional phrase was incorporated into the search string. As a result, 320 articles were identified as potentially irrelevant and were excluded. Phrases like "review," "recent," "progress," "critical," "overview," "bibliometric," and "potentiometric" were used to identify such review articles. To ensure that only research articles remained, the Electronic Identifier (EID) of the review article was utilised to create a new search string. This final search string excluded standard review articles, book chapters, and any other potentially irrelevant studies, ultimately reducing the total number of publications to 2065 articles.

The data for the most productive journals was tabulated and organised based on the total publications within a specific timeline. Additionally, to identify single-country publications (SCP), the keyword "AFFILCOUNTRY" was used to narrow down the search results for each country. After gathering and arranging the information according to various criteria, such as year, source, author, affiliation, country, topic area of interest, and document type, the remaining publications totaled 2065 articles. To ensure a more accurate ranking, bibliometric data from the Scopus website, including total publications and citations, CiteScore, and h-index, were incorporated. It's important to note that separate search strings were executed for each material. Figure 1 illustrates the overall procedure of data mining and the elimination of studies.
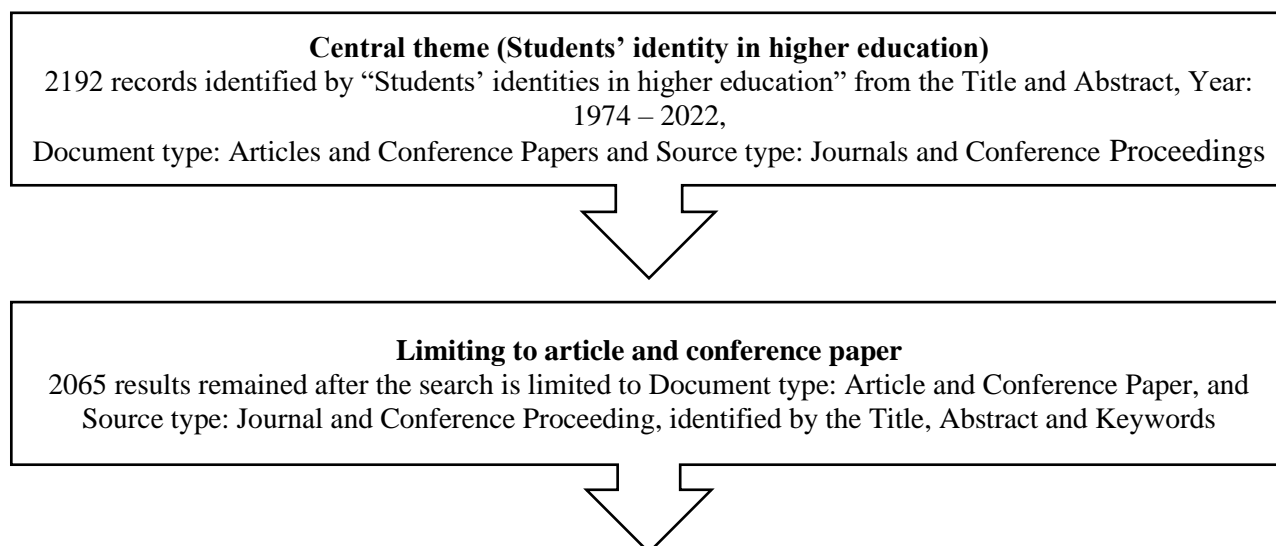
Central theme (Students' identity in higher education)
2192 records identified by "Students' identities in higher education" from the Title and Abstract, Year: 1974 – 2022,
Document type: Articles and Conference Papers and Source type: Journals and Conference Proceedings

Limiting to article and conference paper
2065 results remained after the search is limited to Document type: Article and Conference Paper, and Source type: Journal and Conference Proceeding, identified by the Title, Abstract and Keywords

Figure 1: Research strategy on the data gathering procedure

3.1.1 Bibliometric Maps

The bibliometric mappings were generated based on a dataset comprising 2065 records sourced from the Scopus database and analyzed using the VOSviewer software (version 1.6.18). VOSviewer was chosen as the primary analytical tool for this investigation due to its open-access nature and its compatibility with a diverse array of databases, including Scopus, Web of Science, Dimensions, Lens, and PubMed. These maps were designed to visually represent thematic elements, with countries or author keywords serving as the focal points of interest. The software calculates the strength of relationships between different elements and presents this information as numerical values for straightforward interpretation. Notably, stronger connections between elements are reflected by higher numerical values in the visual representation of links between them.

Using data extracted from 2065 articles, two types of bibliometric maps were created, focusing on co-authorship and co-occurrence. Regarding the assessment of co-authorship, the connection intensity between two elements indicates the extent of shared publications, encompassing journal articles and conference proceedings among the nations. These publications are chosen based on the affiliations of co-authors. Thus, if an article involves authors from diverse nations and affiliations, it is considered a collaborative publication for both countries. Additionally, the cumulative connection strength serves as a gauge for the overall co-authorship influence of a specific country in relation to others. Similarly, the connection intensity connecting two author-provided keywords in co-occurring bibliometric maps signifies the count of instances where the keywords appear jointly in articles. The

development of data representation for both tables and illustrations utilised a comprehensive counting system. A more detailed understanding of the software's capabilities can be acquired from the accompanying instruction manual (Van-Eck & Waltman, 2021).

3.2 Method

For this study, I have adopted Donthu et al.'s (2021) recommended bibliometric analysis methods. This analytical approach helps uncover evolving research trends, highlight emerging topics driving intellectual progress, and identify collaboration patterns among researchers, their nations, and institutions. Recent advancements in bibliometric techniques enable researchers to interpret vast unstructured data, showcasing the growth and impact of relevant publications. Typically, these analyses aim to identify knowledge gaps, update researchers on the latest developments, and inspire new research questions. This study utilised two common approaches: performance analysis (total publications and total citations) and science mapping (citation analysis, co-word analysis and co-authorship analysis) (Donthu et al., 2021). Performance analysis primarily evaluates the contributions of research entities, while science mapping delves into the interconnections among these constituents.

3.2.1 Performance Analysis

Performance analysis involves profiling key components of relevant research, including authors, institutions, journals, and countries (Cobo et al., 2011; Ramos-Rodrígue & Ruíz-Navarro, 2004). This type of analysis, often descriptive in nature, is a fundamental component of bibliometric research (Donthu et al., 2020b), Performance analysis relies on quantitative metrics such as publication counts and citation rates. Various metrics are available for performance analysis. The present study utilised one them; the number of publications and citations per year or per research constituent. Publication count typically signifies productivity, while citation count reflects impact and influence. Additional metrics like citation per publication and the h-index together with citations and publications to gauge the performance of research constituents.

3.2.2 Science Mapping

Science mapping utilises visual aids such as VOSviewer software, helps to elucidate the connections between these research elements. Science mapping delves into collaborative patterns among specified entities, such as authors or institutions, to map out the social dynamics influencing the intellectual landscape of a research topic over a specific period. There are

various techniques for science mapping however, this study only employed 3 of them namely citation analysis, co-word analysis, co-authorship analysis.

Citation analysis, a fundamental technique in science mapping, relies on the premise that citations establish intellectual connections between publications when one cites another (Appio et al., 2014). The impact of a publication is gauged by the number of citations it accrues, facilitating the identification of the most influential works within a research domain. While various methods exist to determine a publication's significance, citations remain the most objective and straightforward measure of its impact (Pieters & Baumgartner, 2002; Stremersch et al., 2007). Through citation analysis, researchers can explore the intellectual dynamics of a field by pinpointing its most influential publications.

Co-word analysis focuses on examining the content of publications by analysing the words used within them. Unlike other techniques such as citation analysis, which focus on cited or citing publications, co-word analysis directly explores the words within the publications themselves. Typically, words for co-word analysis are sourced from author keywords, although they can also be derived from article titles, abstracts, or full texts. This method assumes that words frequently appearing together are thematically related. I investigated emerging themes related to students' identities through co-word analysis. This method assumes that keywords frequently appearing together are thematically linked, forming clusters that represent subtopics with cohesive relationships, akin to "communities of topics," as described by Emich et al. (2020, p. 662). The co-word analysis in this study utilised author keywords from the Scopus database. A total of 4,920 unique keywords were extracted from 2,065 articles during the analysis. To create co-word maps, the author keywords were processed using VOSviewer, ensuring no redundancy. Similar keywords were consolidated into single labels, such as combining "race," "race/ethnicity," into "race." The process resulted in 4,920 refined keywords that were used for constructing the co-word maps. In managing the large number of keywords, a minimum co-word value of 5 was set, which generated 230 keywords. The maps were presented in overlay visualisation mode, where the colour of the items and the strength of the links represented the average number of published articles annually that included the selected keywords.

Co-authorship analysis explores the collaborative interactions among scholars within a research domain. Collaboration through co-authorship represents a formal avenue for intellectual exchange among scholars (Acedo et al., 2006; Cisneros et al., 2018), underscoring the significance of understanding these interactions and associated attributes such as institutional affiliations and countries. Collaboration among scholars has become increasingly

common due to the growing complexity of research methodologies and theories (Acedo et al., 2006), with collaborative efforts often leading to enhanced research outcomes through diverse perspectives (Tahamtan et al., 2016). These collaborative networks, often referred to as "invisible collages," play a crucial role in advancing research endeavours within the field (Crane, 1969). Using VOSviewer software, I performed co-authorship analyses across authors, institutions, and countries. The present study encompasses 73 affiliated countries participating in a co-authorship assessment. The countries were categorised based on the continents of Asia, Europe, America, Africa, and Oceania. The maps were presented using a comprehensive counting system in network visualisation mode. These methods, when integrated with network analysis, provide valuable insights into the bibliometric and intellectual structure of a research field. Baker et al. (2020) and Tunger and Eulerich (2018) offer further insights into these techniques.

4.0 ANALYSIS AND DICUSSION

4.1 Trend of Research Interest and Productivity of Publication in Students' Identities in Higher Education
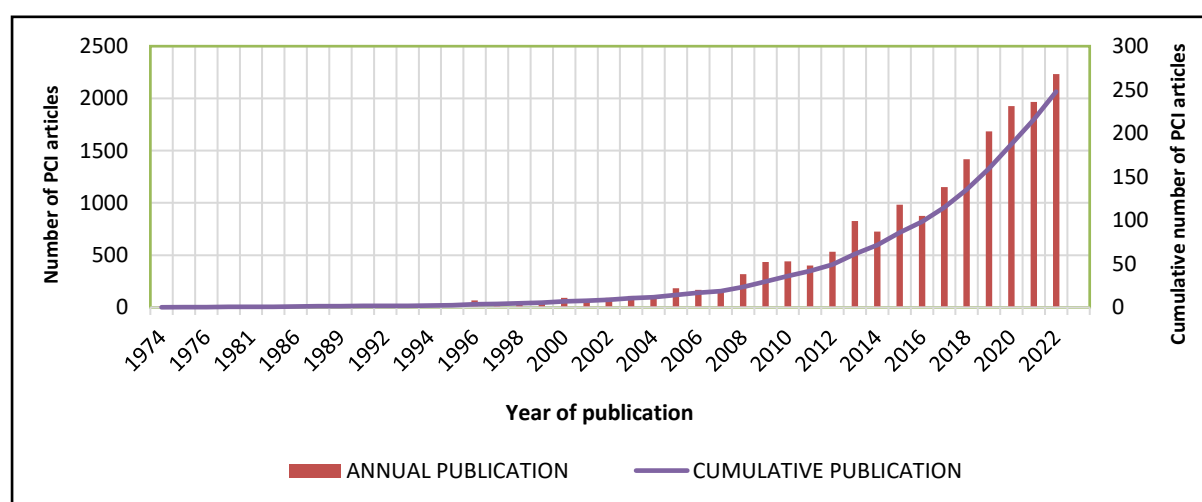
Figure 2: The annual and cumulative numbers of research articles on students' identities in higher education indexed in SCOPUS from 1974 to 2022

The paragraph highlights a 50-year span of research on students' identities in higher education, noting fluctuations in publication output. The figure presented indicates a modest rise in interest in 1995, followed by an increase in 1996 and a subsequent decline from 1997 to 1999. However, a significant upturn in research emerged in 2000, with a notable surge in publications after the low point of 1997-1999 when only 16 articles were published. Post-2000, there was a

consistent and rapid growth in publications, underscoring an escalating recognition of the importance of students' identities in higher education. Notably, from 2002 to the end of 2022, there was an annual increase of almost 10 or more articles. This trend signifies an expanding awareness of the significance of studying students' identities in the higher education context.

Although there was a slight decrease in the annual number of published articles at some point, a subsequent spike in publications was observed. The consistent annual publications contributed to a large cumulative total. It is anticipated that this number will continue to rise in the coming years as awareness among researchers continues to expand. However, one drawback identified by the author is that many of these articles are not published in open access, making it challenging for more researchers to recognize the significance of researching into students' identities in higher education. Based on the 2065 documents analysed, only 799 of them were identified as open access. In highlighting this issue and attempting to increase the number of citations with broader accessibility, it is proposed that more articles should be published in the open access format.

The findings from the Scopus database indicated that several articles included in the study were published in languages other than English. For example, 1912 articles were published in English, followed by Spanish (50), Russian (31), Chinese (20), Portuguese (19), Afrikaans (6), French (6), Lithuanian (6), Korean (4), Croatian (3), Estonian (2), German (2), Malay (2), Swedish (2) and 1 respectively in Arabic, Bosnian, Dutch, Hungarian, Italian, Persian, Polish, Slovenian, Thai and Ukrainian. Despite most of these articles not in English, they were still indexed in Scopus, requiring their titles and abstracts to be in English.

4.2 Preferred Journals

Table 1: The top 10 most productive journals on students' identities in higher education research with their most cited articles from 1974 to 2022

RANK	JOURNAL	TOTAL PUBLICATION (%)	CITESCORE 2022	THE MOST CITED ARTICLE	TIMES CITED	FWCI	PUBLISHER
1	Studies in Higher Education	55 (2.7)	10.8	Student Writing in Higher Education: An academic literacies approach (Lea & Street, 1998)	1062	5.85	Taylor & Francis
2	Journal of College Student Development	37 (1.8)	2.4	Social support, self-esteem, and stress as predictors of adjustment to university among first-year undergraduates (Friedlander et al., 2007)	307	4.2	Johns Hopkins University Press
3	Teaching in Higher Education	36 (1.7)	5.7	Building student belonging and engagement: insights into higher education students' experiences of participating and learning together (Masika & Jones, 2016)	133	6.75	Taylor & Francis
4	Higher Education	33 (1.6)	8.6	Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices (Oleson & Hora, 2014)	169	6.25	Springer Nature
5	Journal of Diversity in Higher Education	30 (1.5)	5.3	Learning race in a U.S. Context: An emergent framework on the perceptions of race among foreign-born students of colour (Fries-Britt et al., 2014)	70	2.49	APA

6	Frontiers in Psychology	26 (1.3)	4.5	Bridging and bonding interactions in higher education: Social capital and students' academic and professional identity formation (Jensen & Jetten, 2015)	40	1.25	Frontiers Media S.A.
7	Journal of Further and Higher Education	26 (1.3)	4.5	Context and identity: Exploring adult learners' experiences of higher education (Askham, 2008)	58	0.79	Taylor & Francis
8	Higher Education Research and Development	25 (1.2)	6	Re-conceptualising graduate employability: the importance of pre-professional identity (Jackson, 2016)	190	11.11	Taylor & Francis
9	Race Ethnicity and Education	18 (0.9)	6.3	Ethnic choosing': Minority ethnic students, social class and higher education choice (Ball et al.,2002)	182	6.28	Taylor & Francis
10	British Journal of Sociology of Education	15 (0.7)	3.2	Emotional journeys: Young people and transitions to university (Christie, 2009)	90	3.74	Taylor & Francis

The data collected from Table 1 reveals that the top 10 most productive journals are owned by five different publishers. Among these publishers, Taylor & Francis has the highest number of journals, with six in total; Studies in Higher Education, Teaching in Higher Education, Journal of Further and Higher Education, Higher Education Research and Development, Race Ethnicity and Education, and British Journal of Sociology of Education. The remaining four journals are published by Johns Hopkins University Press, Springer Nature, APA and Frontiers Media S.A.

The journal with the highest output is 'Studies in Higher Education', which has published 55 articles, accounting for 2.7% of the total publications. It is followed by 'Journal of College Student Development' with 37 articles (1.8%), 'Teaching in Higher Education with 36 articles (1.7%), 'Higher Education' with 33 articles (1.6%), and 'Journal of Diversity in Higher Education' with 30 articles (1.5%). The remaining journals have published fewer than 30 articles respectively. The highest output journal, 'Studies in Higher Education' has the highest total number of citations, with 1062 whereas the Journal of College Student Development' had 307 citations for its most cited article, placing it in second position. On the other hand, 'British Journal of Sociology of Education' has the least cited article, with only 90 citations, placing it in tenth position.

Based on the findings of CiteScore 2022, six journals had an average CiteScore above 5. The Journal of Studies in Higher Education, ranked 1st on the list, achieved the highest CiteScore of 10.8 compared to other journals. On the other hand, 'British Journal of Sociology of Education' had the lowest CiteScore at 3.2. Many researchers consider CiteScore as a measurement to gauge the journal's receiving number of audiences each year. This measurement is significant for authors when deciding where to publish their work. CiteScore, an alternative to Clarivate Analytics' Impact Factor, measures a journal's influence by gathering and comparing citation information from the Scopus database. However, it is essential to recognize that CiteScore should not be the sole basis for analysing a journal's impact as other factors should also be considered. In addition to CiteScore, it is essential to assess whether the journal could effectively reach the appropriate audience and make meaningful contributions to advancements in their field.

4.3 Leading Countries and Top Institutions

RANK	COUNTRY	TPc	SCP (%)	THE MOST PRODUCTIVE INSTITUTIONS	(TPi)
1	United States	653	584 (89%)	University of California	17
2	United Kingdom	434	358 (82%)	King's College London	15
3	Australia	167	126 (75%)	Griffith University	12
4	South Africa	91	82 (90%)	University of Cape Town	13
5	China	63	40 (63%)	Peking University Health Science Center	3
6	Spain	62	42 (68%)	Universitat Rovira i Virgili	4
7	Canada	61	39 (64%)	University of Calgary	3
8	Russian Federation	48	41 (85%)	HSE University	5
9	Israel	30	29 (97%)	University of Haifa	4
10	Sweden	30	21 (70%)	Linköpings Universitet	6
11	Denmark	29	19 (66%)	Københavns Universitet	11
12	Hong Kong	29	24 (83%)	The Education University of Hong Kong	11
13	Malaysia	29	21 (72%)	Universiti Kebangsaan Malaysia	7
14	Brazil	28	21 (75%)	Universidade Federal do Rio de Janeiro	2
15	Ireland	28	21 (75%)	Trinity College Dublin	6

*TPc: Total publications of a single country, TPi: Total publications of a given academic institution, SCP: Single-country publications

Figure 3: The top 15 countries and academic institutions that make the most articles on

students' identities in higher education publications

Figure 3 illustrates the top 15 countries that have made significant contributions to the field of students' identities in higher education research. Among all the worldwide publications, the United States, United Kingdom, Australia, and South Africa accounted for approximately 60% of the total. These countries emerged as major players in this subject. The United States stands

out as the leading country with 653 publications in terms of the total publication output of each country (TPc). The United Kingdom (434) secured the third position and followed by Australia (167) and South Africa (91). The rank order remains for the most productive institutions for the publications where University of California from the United States was the leading institution with 17 in total in regards with the total publications of a given academic institution (TPi). Additionally, Trinity College Dublin is the least productive institution for the publications with only 6 publications in total.

Among the countries listed, China and Canada had the lowest total of single country publications (SCP) with 63% and 64% each. A high SCP value indicates strong collaboration within a country. To reiterate, most publications are authored solely by individuals affiliated with that country, without involvement from authors of other countries. Despite having the lowest SCP rates, both China and Canada demonstrated extensive international collaborations. Having a broad network of international collaborations offers several advantages, such as expanding knowledge exchange, sharing theoretical and practical expertise, and contributing to a constructive university ranking strategy. The United States stood out with the highest SCP rate of 89%, signifying that all more than 2/3 of its publications were from within the country without any involvement from other nations. Although China had the lowest SCP rate at 63%, it still managed to produce 23 articles for its publications through international collaborations involving other countries, securing the 5th position among the most productive countries.
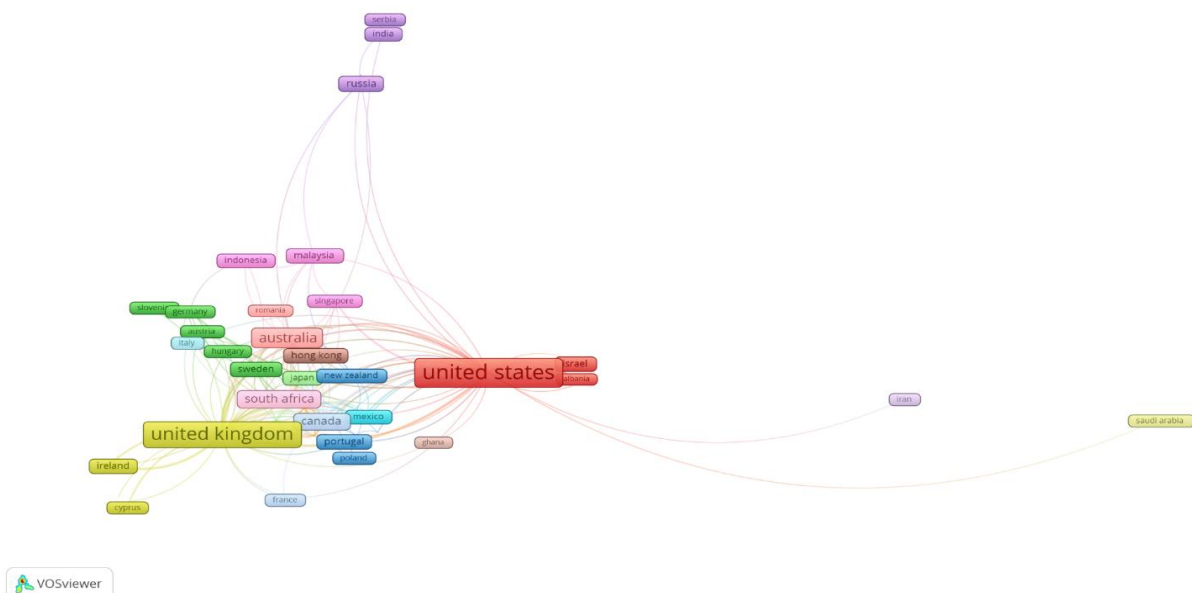

Figure 4: Co-authorship mapping in network visualisation mode. The following URL can be used to open Figure 4 in online VOSviewer: https://tinyurl.com/2ybt7z53

Based on the Figure 4 above, the countries are organised based on their respective regions. In VOSviewer, the connections between countries indicate their relationship, which is determined by the distance and thickness of the links. These factors represent the extent of collaborative work in students' identities in higher education research. When countries are closer together on the map, it indicates a stronger collaboration in terms of joint publications. Similarly, thicker lines signify higher levels of collaboration. Figure 4 shows different regions represented by colours: red for Asian countries, green for Europe, blue for America, yellow for Africa, and purple for Oceania. The number of countries contributing to students' identity in higher education research in each continent is as follows: Asia (24), Europe (25), America (8), Africa (5) and Oceania (3).

It was found that the United States had the highest affiliation with 40 countries and a total of 78 co authorships. Following that, the United Kingdom, with 36 links and 98 co authorships, Australia had 19 links and 48 co authorships, Spain had 18 links and 31 co authorships, and Sweden had 14 links and 20 co authorships. Among the 62 countries in the database, over 80% of them have published fewer than 10 articles in collaboration with partners from other countries. There are various factors that could influence the effectiveness of a country's international collaboration. These factors include the extent of a researcher's connections with colleagues from foreign countries, the presence of visiting researchers, access to international grants, the involvement of foreign postgraduate students, and more. Therefore, enhancing and expanding these contributing elements will play a crucial role in creating a versatile and dynamic research project that can yield outstanding results and future benefits.

4.4 Leading Authors

Table 2 below presents the most productive authors in the field of students' identities in higher education research. Among the 10 authors listed, 4 were affiliated with the United Kingdom, followed by Denmark with 3 authors, and the United States, Australia and South Africa with 1 author respectively. The first publications in this field ranged from 1988 to 2017, with 9 authors serving as the first author and 1 author as co-author. Readers may decide whether the total number of citations, the h-index, or other criteria should be used to evaluate the article's quality. The h-index value generally appraises an author's productivity and the impact of their contributions in a specific research area.

Bennett, Dawn A., an author from Australia, secured the top position on the list with the count of 111 publications since the year 2006. Notably, these works amassed a total of 1567 citations, earning Dawn with an h-index of 22. The second and third spots were claimed by

Busher, Hugh from University of Leicester, United Kingdom and Duran, Antonio A. from Arizona State University, United states with varying years of first publication. Taking the fourth and fifth spots were researchers Madsen, Lene MÃ¸ller and Holmegaard, Henriette Tolstrup, affiliated with Københavns Universitet, Denmark starting in the year of 2001 and 2010 each. Bornman, Elirea and Ulriksen, Lars M. took the sixth and seventh place respectively in the year of 1990 and 2009 onwards. However, Bornman, Elirea took a role as the coauthor in 30 publications in total. The United Kingdom's representation continued with a trio Holton, Mark (ranked 8th), Tobbell, Jane (ranked 9th) and Preece, Siân (ranked 10th), associated with different institutions: University of Plymouth, University of Huddersfield and UCL Institute of Education. Despite being ranked fourth, Madsen, Lene MÃ¸ller has received significantly more citations, with a total of 746, compared to Duran, Antonio A., third in rank, who acquired 391 citations. A similar trend is also observed for Ulriksen, Lars M., ranked seventh, who secured the fourth-highest number of citations at 605.

Table 2: List of the most prolific authors in the field of students' identities in higher education research

NO.	AUTHOR	SCOPUS AUTHOR ID	YEAR OF 1ST PUBLICATION	TOTAL PUBLICATION	H-INDEX	TOTAL CITATION	CURRENT AFFILIATION	COUNTRY
1	Bennett, Dawn A.	55574191781	2006 (a)	111	22	1567	Bond University	Australia
2	Busher, Hugh	14044541100	1988 (a)	68	15	907	University of Leicester	United Kingdom
3	Duran, Antonio A.	57194707974	2017 (a)	58	11	391	Arizona State University	United States
4	Madsen, Lene MÃ¸ller	56268382500	2001 (a)	53	14	746	Københavns Universitet	Denmark
5	Holmegaard, Henriette Tolstrup	39361422800	2010 (a)	33	10	512	Københavns Universitet	Denmark
6	Bornman, Elirea	8687528400	1990 (c)	30	9	308	University of South Africa	South Africa
7	Ulriksen, Lars M.	35325395800	2009 (a)	29	12	605	Københavns Universitet	Denmark
8	Holton, Mark	56032427800	2013 (a)	23	13	459	University of Plymouth	United Kingdom
9	Tobbell, Jane	17347171600	2003 (a)	17	10	441	University of Huddersfield	United Kingdom
10	Preece, Siân	35225268600	2009 (a)	15	9	275	UCL Institute of Education	United Kingdom

*Role in co-authorship, superscripts (a) – First author (c) – C-author

4.5 Author Keywords

The VOSviewer software compiled a dataset containing 4904 author keywords. Following analysis, keywords with similar or related meanings were re-labeled to avoid repetition. A total of 203 keywords qualified for inclusion (met the threshold) in the bibliometric maps, requiring a minimum of 5 occurrences each.

*The colour indicates the average publication year of the documents in which the keywords occur

Figure 5: A bibliometric map generated according to the author keywords co-occurrence in overlay visualisation mode. The following URL can be used to open Figure 4 in VOSviewer: https://tinyurl.com/234lqw6e

4.5.1 Terminology and Concept

As illustrated in Figure 5, it was discovered that the term 'higher education' holds the highest frequency among author-provided keywords, occurring 401 times and interconnected with 175 links to other associated keywords. The rest of the keywords order are as follows; 'identity' (198 occurrences, 144 links), 'student' (57 occurrences, 52 links), 'professional identity' (53 occurrences, 48 links) and 'education' (52 occurrences, 52 links). Furthermore, these terms are also associated with specific countries or regions. For instance, Australia (5 occurrences, 3 links), China (13 occurrences, 17 links), South Africa (11 occurrences, 16 links), Europe (5 occurrences, 7 links) and Mexico (5 occurrences, 6 links). This situation highlights the existence of locations that share a relevant connection with the subject matter investigation.

Researchers can benefit from these location-based associations to outline the limitations of the study as illustrated in Figure 5.

The term 'higher education' is frequently chosen as a keyword by authors, likely because it indicates that the term was widely used and connected to most of the keywords shown in the database. The term also represents the population that is within the setting of higher education in numerous articles, where it is also the only targeted setting for the study. This is followed by the term 'identity' ranked second (198 occurrences and 144 links) and 'student' (57 occurrences, 52 links) which are quite widely used by the researchers. Some of the keywords associated with the term 'identity' are 'professional identity' (53 occurrences), 'gender' (48 occurrences), 'ethnicity' (37 occurrences), 'student experience' (18 occurrences), 'agency' (15 occurrences). 'students' transition' (6 occurrences), etc. The term 'student' is often linked to the keywords like 'gender' (48 occurrences), 'diversity' (46 occurrences), 'social identity' (26 occurrences), 'learning' (19 occurrences), e-learning (14 occurrences), 'academic achievement' and 'communication' (8 occurrences) respectively, etc.

We also identified a few keywords attributed to different kinds of identities that are linked to 'higher education' namely; 'professional identity' (53 occurrences), 'social identity' (26 occurrences), 'academic identity' (18 occurrences) 'learner identity' (17 concurrences) and 'teacher identity' (11 occurrences). These identities are also linked with other keywords that are attributed to some identified theories that are probably utilised in investigating students' identities in higher education. This includes 'community of practice' (23 occurrences) (Wenger, 1998), 'acculturation' (11 occurrences) (Berry, 1980), 'self-concept' (Rogers, 1959) (9 occurrences), 'habitus' (8 occurrences) (Bourdieu, 1990), 'situated learning' (7 occurrences) (Lave & Wenger, 1991), etc. The theories of Community of Practice, Acculturation, Self-Concept, Habitus, and Situated Learning provide frameworks for understanding students' identities in higher education. They highlight the social, cultural, and psychological factors that shape how students perceive themselves and interact within academic communities. By exploring these theories, researchers gain insights into how students navigate cultural shifts, develop self-perceptions, and construct knowledge within authentic learning contexts, ultimately informing efforts to support their academic and personal development.

There are also significant keywords that are directly correlated with the term 'identity' which are 'gender'(48 occurrences), 'diversity' (46 occurrences), 'ethnicity' (37 occurrences), 'transition' (32 occurrences), 'community of practice' (23 occurrences), 'student experience' (18 occurrences),'agency' (15 occurrences), 'engagement' (12 occurrences), 'academic performance' (7 occurrences), etc. This study also discovered that among the keywords

mentioned, some are associated directly with the term 'student' namely; 'gender' (48 occurrences), 'social identity' (26 occurrences), 'e-learning' (14 occurrences), 'academic achievement' (8 occurrences).

In conclusion, the terms and keywords with regards to the students' identities in higher education presented by the researchers in their articles in this study offer a valuable avenue for other scholars or researchers in determining their scope of study and potential theories in exploring students' identity in higher education.

4.6 Limitation of Study

The data output provided by Scopus may not encompass all articles related to students' identities as the search is limited to students' identities in higher education found in titles and abstracts. This limitation arises likely because some researchers may employ different responses or expressions to convey their research focus instead of the specified terms mentioned in this study and the setting chosen for the study is only in higher education. Furthermore, VOSviewer might have excluded certain author keywords in the co-word analysis during the database screening due to incomplete information in the journals. Thus, it is advisable for future research to incorporate databases from alternate sources such as Web of Science and Google Scholar. This approach would enhance the comprehensiveness and interest value of the bibliometric analysis.

5.0 CONCLUSION

This research aimed to offer an overall perspective on the students' identities in higher education research trends spanning from 1974 to 2022, utilising data from the Scopus database, encompassing a total of 2065 articles from reputable journals and conference proceedings. Before gathering the data, various refinement steps were undertaken, including the exclusion of irrelevant and review articles. Visual representations of the textual materials were created using VOSviewer software, presented in the form of figures. Additionally, tables displaying the most prolific journals and authors were organised based on the Scopus website's framework. The initial decade showed a slow growth in publications, with a noticeable acceleration beginning in 2002. Based on this observed pattern, it's predicted that interest in students' identities in higher education research will continue to rise for the foreseeable future. Notably, the United States, the United Kingdom and Australia were the leading contributors in terms of article count, presenting significant opportunities for nations with fewer publications to engage in global collaborations and enhance their knowledge base. The study also identified

prevalent research topics, highlighting the focus on students' identities in higher education settings. Among these, some theories on identities were also identified in the study. The study aims to serve as a valuable resource for researchers exploring the potential spectrum of exploring students' identities in higher education across the globe.

REFERENCES

Acedo, F. J., Barroso, C., Casanueva, C., & Galan, J. L. (2006). Co-authorship in management and organizational studies: An empirical and network analysis. *Journal of Management Studies, 43*(5), 957–983.

Antaki, C. (2008). Identity ascriptions in their time and place: 'Fagin' and 'The Terminally Dim'. In C. Antaki & S. Widdicombe (Eds.), *Identities in talk* (pp. 71-86). SAGE Publications Ltd.

Appio, F. P., Cesaroni, F., & Di Minin, A. (2014). Visualizing the structure and bridges of the intellectual property management and strategy literature: A document co-citation analysis. *Scientometrics, 101*(1), 623–661.

Archer, M. S. (2003). *Structure, agency and the internal conversation.* Cambridge University Press.

Baker, H. K., Kumar, S., & Pandey, N. (2020). A bibliometric analysis of managerial finance: A retrospective. *Managerial Finance, 46*(11), 1495–1517.

Bamberg, M., De Fina, A., & Schiffrin, D. (2011). Discourse and identity construction. In K. Luyckx, S. Schwartz, & V. Vignoles (Eds.) *Handbook of identity theory and research, vol 1* (pp. 177-199). Springer.

Baxter, J. A., & Haycock, J. (2014). Roles and student identities in online large course forums: Implications for practice. *The International Review of Research in Open and Distributed Learning, 15*(1), 21-40.

Benwell, B., & Stokoe, E. (2006). *Discourse and identity*. Edinburgh University Press.

Berkman, E. T., Livingston, J. L., & Kahn, L. E. (2017). The identity-value model of self-regulation: Integration, extension, and open questions. *Psychological Inquiry, 28*(2–3), 157–164.

Berry, J. W. (1980). Acculturation as varieties of adaptation. In A. Padilla (Ed.), *Acculturation: Theory, models and findings* (pp. 9-25). Westview Press.

Bös, B., Kleinke, S., Mollin, S., & Hernández, N. (2018). *The discursive construction of identities on- and offline: Personal – group – collective*. John Benjamins Publishing Company.

Brown, B. A., Reveles, J. M., & Kelly, G. J. (2005). Scientific literacy and discursive identity: A theoretical framework for understanding science learning. *Science Education, 89*(5), 779–802.

Bourdieu, P. (1990). *The logic of practice*. Polity.

Chadegani, A. A., Salehi, H., Yunus, M., Farhadi, H., Fooladi, M., Farhadi, M., & Ale Ebrahim, N. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. *Asian Social Science, 9*(5), 18–26.

Chen, J., Tan, T., & Lei, J. (2022). Exploring learner identity in the blended learning context: A case study of collaborative writing. *System, 108*(August), 102841.

Cisneros, L., Ibanescu, M., Keen, C., Lobato-Calleros, O., & Niebla-Zatarain, J. (2018). Bibliometric study of family business succession between 1939 and 2017: Mapping and analyzing authors' networks. *Scientometrics, 117*(2), 919–951.

Cobo, M. J., L´opez-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics, 5*(1), 146–166.

Côté, J. E., & Levine, C. G. (2002). *Identity formation, agency, and culture: A social psychological synthesis***.** Erlbaum.

Daher, W., & Shahbari, J. A. (2020). Secondary students' identities in the virtual classroom. *Sustainability, 12*(11), 4407.

Daher, W. (2012). Educational and cultural identities in virtual social networks. *International Journal of Cyber Ethics in Education, 2*(1), 57–70.

Donthu, N., Kumar, S., Pandey, N., & Gupta, P. (2021). Forty years of the International Journal of Information Management: A bibliometric analysis. *International Journal of Information Management, 57*(C), 102307.

Donthu, N., Kumar, S., & Pattnaik, D. (2020a). Forty-five years of Journal of Business Research: A bibliometric analysis. *Journal of Business Research, 109*(1), 1–14.

Donthu, N., Reinartz, W., Kumar, S., & Pattnaik, D. (2020b). A retrospective review of the first 35 years of the International Journal of Research in Marketing. *International Journal of Research in Marketing, 38*(1), 232-269.

Duff, P. A. (2007). Second language socialization as sociocultural theory: Insights and issues. *Language Teaching, 40*(4), 309–319.

Emich, K. J., Kumar, S., Lu, L., Norder, K., & Pandey, N. (2020). Mapping 50 years of small group research through small group research. *Small Group Research, 51*(6), 659–699.

Farland-Smith, D. (2012). Personal and social interactions between young girls and scientists: Examining critical aspects for identity construction. *Journal of Science Teacher Education, 23*(1), 1-18.

Gee, J. P. (2001). Identity as an analytic lens for research in education. *Review of Research in Education, 25*(1), 99–125.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy.* Palgrave Macmillan.

Giroir, S. (2014). Narratives of participation, identity, and positionality: Two cases of Saudi learners of English in the United States. *TESOL Quarterly, 48*(1), 34–56.

Grad, H., & Martín-Rojo, L. (2008). Identities in discourse: An integrative view. In R. Dolón, & J. Todolí (Eds.), *Analysing identities in discourse* (pp. 3-28). John Benjamins Publishing Co.

Gray, R. J. (2022). *Genre practices, multimodality and student identities*. Palgrave Macmillan.

Hawkins, M. R. (2005). Becoming a student: Identity work and academic literacies in early schooling. *TESOL Quarterly, 39*(1), 59–82.

Ibarra, H., & Petriglieri, J. (2010). Identity work and play. *Journal of Organizational Change Management, 23*(1), 10–25.

Jiang, Y., Ritchie, B. W., & Benckendorff, P. (2019). Bibliometric visualisation: An application in tourism crisis and disaster management research. *Current Issues in Tourism, 22*(16), 1925-1957.

Jenkins, R. (2014). *Social identity* (2nd ed.). Routledge.

Kasworm, C. E. (2010). Adult learners in a research university: Negotiating undergraduate student identity. *Adult Education Quarterly, 60*(2), 143-160.

Lantolf, J. P., & Pavlenko, A. (2001). (S)econd (L)anguage (A)ctivity Theory: Understanding second language learners as people. In M. Breen (Ed.), *Learner contributions to language learning: New directions in research* (pp. 141-158). Longman.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.

Lee, S. K. (2003). Exploring the relationship between language, culture and identity. *GEMA Online Journal of Language Studies, 3*(2), 1–13.

Mili, & Towers, E. (2022). *Teaching in higher education*. Routledge.

Nadia, T., & Zurina, K. (2023). The construction and negotiation of agentive identity: A case study of a Malaysian student. *Journal of Nusantara Studies, 8*(2), 375-396.

Nias, J. (1989). Teaching and the self. In *Perspective on teacher professional identity* (pp. 151-171). Falmer Press.

Norton, B. (1997). Language, identity and the ownership of English. *TESOL Quarterly, 31*(3), 409–429.

Norton, B., & Toohey, K. (2001). Changing perspectives on good language learners. *TESOL Quarterly, 35*(2), 307–323.

Park, J. E. (2007). Co-construction of nonnative speaker identity in cross-cultural interaction. *Applied Linguistics, 28*(3), 339–360.

Pieters, R., & Baumgartner, H. (2002). Who talks to whom? Intra-and interdisciplinary communication of economics journals. *Journal of Economic Literature, 40*(2), 483–509.

Rajadurai, J. (2010a). "Malays are expected to speak Malay": Community ideologies, language use and the negotiation of identities. *Journal of Language, Identity & Education, 9*(2), 91–106.

Rajadurai, J. (2010b). Speaking English and the Malay community. *Indonesia and the Malay World, 38*(111), 289–301.

Ramos-Rodrígue, A. R., & Ruíz-Navarro, J. (2004). Changes in the intellectual structure of strategic management research: A bibliometric study of the Strategic Management Journal, 1980–2000. *Strategic Management Journal, 25*(10), 981–1004.

Rogers, C. (1959). A theory of therapy, personality relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science Formulations of the person and the social context* (Vol. 3, pp. 184-256). McGraw Hill.

Ruelle, O., & Peverelli, P. (2017). The discursive construction of identity through interaction on social media in a Chinese NGO. *Chinese Journal of Communication, 10*(1), 12-37.

Rumenapp, J. C. (2016). Analyzing discourse analysis: Teachers' views of classroom discourse and student identity. *Linguistics and Education, 35*(1), 26-36

Seyri, H., & Ali Rezaee, A. (2022). PhD students' identity construction in face-to-face and online contexts. *Research in Post-Compulsory Education, 27*(1), 48-65.

Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *Journal of Marketing, 71*(3), 171–193.

Tahamtan, I., Safipour Afshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics, 107*(3), 1195–1225.

Ting-Toomey, S. (1999). *Communicating across cultures*. Guilford Press.

Tunger, D., & Eulerich, M. (2018). Bibliometric analysis of corporate governance research in German-speaking countries: Applying bibliometrics to business research using a custom-made database. *Scientometrics, 117*(3), 2041–2059.

Ubaidillah, M. F., & Widiati, U. (2021). Identity construction and classroom participation of an Indonesian graduate student in an Australian university. *Journal of English Education and Linguistics Studies, 8*(1), 147-174.

Van-Eck, N. J., & Waltman, L. (2021). *Manual for VOSviewer version 1.6.17.* Leiden University.

Virta, J., Hökkä, P., Eteläpelto, A., & Rasku-Puttonen, H. (2023). Informal learning contexts in the construction of physical education student teachers' professional identity. *European Physical Education Review, 29*(1), 22-39.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge University Press.

Winchester, J. (2013). The potential impact of the teacher on student identities in the classroom in an English language teaching context. *TESOL Journal, 4*(4), 697-716.

Zacharias, N. (2010). Acknowledging learning multiple identities in EFL classroom. *K@ta, 12*(1), 26-41.