ISSN 2411-7390

jle

JOURNAL OF LANGUAGE & EDUCATION

Volume 6 Issue 4, 2020

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Contents

Tikhonova, E., & Raitskaya, L. (2020). Mapping the Current Research Agenda on Scholarly Publishing: Scopus-Indexed Reviews. *Journal of Language and Education, 6*(4), 4-10. https://doi.org/10.17323/jle.2020.11743

Mapping the Current Research Agenda on Scholarly Publishing: Scopus-Indexed Reviews

Elena Tikhonova[1,2] , Lilia Raitskaya[3]

[1] National Research University Higher School of Economics

[2] Peoples' Friendship University of Russia (RUDN University)

[3] Moscow State Institute of International Relations (MGIMO University)

Correspondence concerning this article should be addressed to Elena Tikhonova, National Research University Higher School of Economics, 11 Pokrovsky Bulvar, Pokrovka Complex, room S313, Moscow, 109028, Russian Federation. E-mail: etihonova@hse.ru

Nearly ten years ago, scholarly publishing came to the fore in research on scientific communication spurred by the evolving Open Science system, the reinvention of peer reviews, and new attitudes to scholarly publications in the ranking-based academic environment. Here, the JLE editors revisit the field of scholarly publishing and identify the most popular areas where potential JLE authors might have difficulty. In this editorial, Scopus-indexed reviews are analysed to map the prevailing trends. The editorial review shows that the trends include open access, peer review transparency, the changing role of libraries in scholarly publishing, CrossRef's initiatives, outsourcing and skills lacking in publishing, open-access monographs, and the role of commercial publishers.

Keywords: scholarly publishing, scholarly writing, open science, open access, DOAJ, language choice, peer review, academic journals

Introduction

Scholarly publishing has been researched more intensely over the last ten years as a separate field of study. The reality of "publish or perish" has motivated researchers to seek more publications in high-ranking journals as these activities determine a researcher's status in the quantitative systems of university rankings. Research on scholarly publishing is highly influenced by the emerging field of Open Science and other trends igniting heated debate. In fact, some scholars feel that 'evidence-based rational debate' is often suppressed by misinformed rhetoric (Tennant et al., 2019).

Some of the topics are highly contested. Others are universal and have been in existence for years. The striking feature of the major research themes is brand-new aspects of the old phenomena.

The Open Science movement has been changing the landscape of science at large in both academic and scholarly publishing. Open Science as a new philosophy gravely transforms attitudes and processes, and attaches new understanding to old phenomena. Some heated topics have emerged, with misinformed rhetoric preventing scholarly publishing from fast-track transformations. Whereas open access is the most visible part of Open Science, "efforts have extended to the availability of open data and software" (Wolfram, Wang, Hembree, & Park, 2020, p.1034).

Open-access publishing has become one of the centrepieces in research on scholarly publishing. Open access (OA) dates back the beginning of the 21st century. It came into being as a set of practices and principles of research output distribution in electronic journals that was widely spawned in the early 2000s. With all the colour-naming OA systems in the world, the gist of it is linked to readers' and society's access to science and new knowledge dissemination.

Being part and parcel of scholarly publishing, peer review remains one of the delicate and vital spheres of scholarly publishing. Although it has been a professional and social filter and barrier to low-quality research for 300-plus years, peer review occasionally involves human failings and does not always prove the relevance and accuracy of the publication (Tennant et al., 2019). The research community put forward a lot of criticism and drawbacks of the peer review process, including bias, peer review fraud, and unfairness (Wolfram, Wang, Hembree, & Park, 2020); unnecessary delays (Benos et al., 2007); as well as a preponderance of incompetent reviewers, a lack of constructive criticism, and editorial passivity (Sciullo & Duncan, 2019).

Peer review is under scrutiny from all of the parties concerned. New progressive forms are being tested. There are many recommendations on how to improve the process and make peer review more effective and bias-free. It is suggested that reviewers and editors should be properly trained and open review forms be introduced (Sciullo & Duncan, 2019). Payment and praise for reviewers are being discussed by the wider academic community.

Peer review aims to improve research and its reporting quality, reducing "the risk of publishing research that is sloppy, erroneous or, at worst, fabricated" (Wolfram, Wang, Hembree, & Park, 2020, p.1034).

Open Science has led to changes in peer review, offering open peer review where the previously hidden process has become public. Peer reviewers are being held more responsible for their assessment and recommendations, as with this form of peer review their names are known to the public. The traditional blind peer review effect of the "black box", thus, has worn off.

In this short editorial review, we map the most essential trends in research on scholarly communication by answering the following research question: What are the recent trendsetting themes for research on scholarly publishing?

Methodology

Scholarly publishing research is represented in journals ranging by their scope from communication and linguistics to education and medicine. Although there are a few basic problems at their core, various aspects of scholarly publishing occasionally come out linked to a specific field or subject area. As major peer-reviewed journals on scholarly publishing and aligned topics are indexed with Scopus, we decided to rely on the database, studying the research agenda. To be more exact, we would like to illustrate our mapping of the mainstream directions of study by reviews published in Scopus-indexed journals.

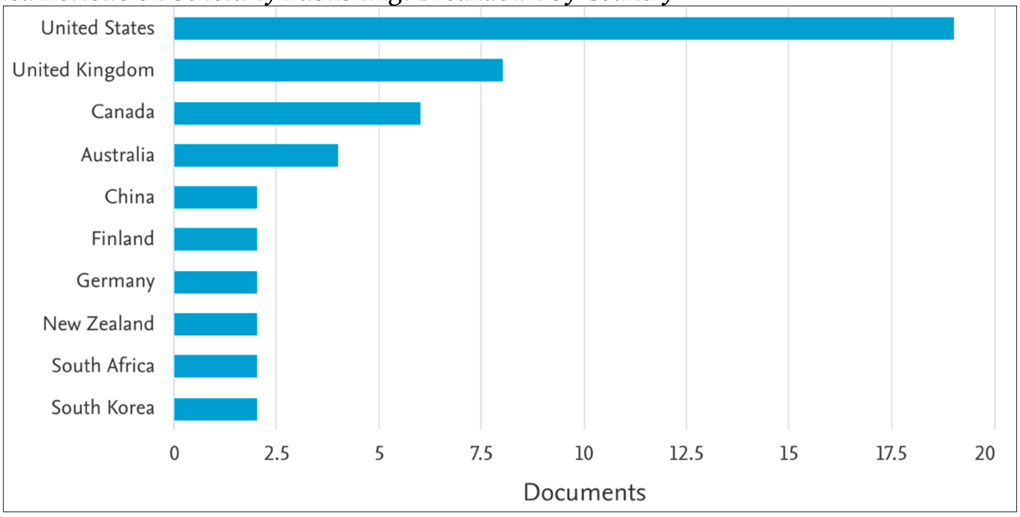
As the recent noticeable upturn in the number of studies related to scholarly publishing started in 2012, we limited ourselves to the period between 2012 and 2020 (although number of papers for 2020 is not complete yet, it covers most of the publications). We started our search for documents (article title, abstracts, and keywords), with "scholarly publishing" as a key term. It brought 1,363 documents spreading between 1934 and 2020. Starting from 2012, more than 50 new publications on scholarly publishing were indexed in Scopus yearly, with a total of 826 documents (2012-2020).

As reviews tend to give a deep analysis of the field, we studied the reviews published in the *Journal of Scholarly Publishing* (Q2, n=21), *Science Editing* (Q2, n=5), *Insights: the UKSG Journal* (Q2, n=4), *Information Services and Use* (Q2, n=3), *Learned Publishing* (Q1, n=2), *Libri* (Q2, n=2), and another 14 journals (n=1).

The affiliations of the reviews are rather numerous, with three of them belonging to Princeton University and the University of Sydney each. Fordham University and University of Technology Sydney had two affiliations each. The other universities (n=80) affiliate one review each.

The reviews were unevenly spread geographically, with 19 reviews coming from the USA, eight from the UK, eight from Canada, and four from Australia. Thus, 37 (72.5 %) out of 47 reviews were Anglophone (See *Figure* 1).

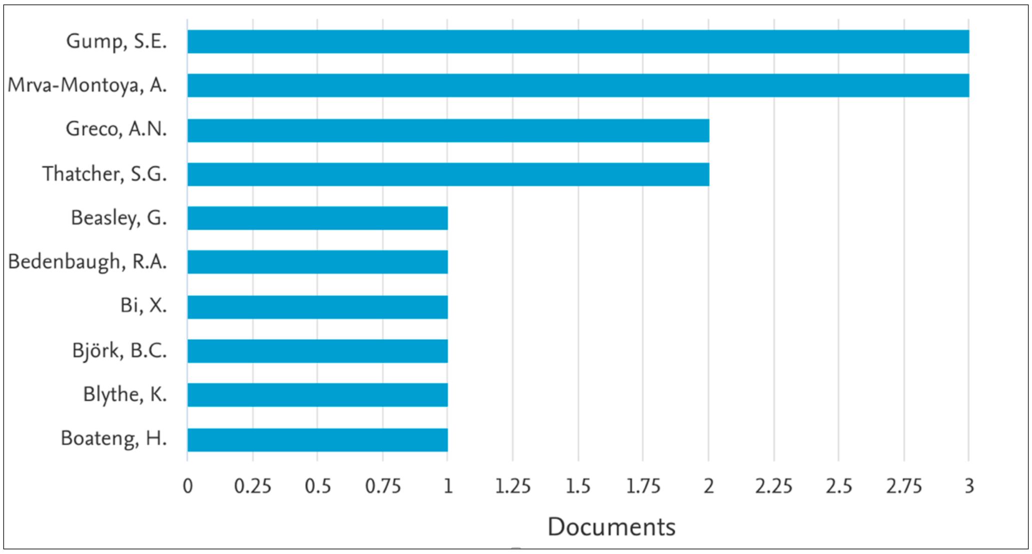
Figure 1
Scopus-Indexed Reviews on Scholarly Publishing: Breakdown by Country



Note. Adapted from Scopus Database, 2020. Copyright 2020 by Scopus.

S.E. Gump and A. Mrva-Montoya authored three of the 51 reviews each, and A.N. Greco and S.G. Thatcher had two publications each (See *Figure 2*). None of the other authors had more than one review each.

Figure 2
Scopus-Indexed Reviews on Scholarly Publishing: Documents Per Author



Note. Adapted from Scopus Database, 2020. Copyright 2020 by Scopus.

While filtering the selected reviews, we found that some of them were rather irrelevant to the field of scholarly publishing. Thus, four publications were omitted from our analysis.

Some publications aim to give an overview of the trends in scholarly publishing or research carried out in one subject area. Tennant et al. (2019) put forward ten hot topics around scholarly publishing in an attempt to sum up the most sensitive issues. Another review focused on updating journal editors on publishing trends and providing readers with the latest research (Smart, 2014).

Results

Open Science and Open Access in Scholarly Publishing

Fifteen of the 47 reviews directly or partially focused on open access publishing and Open Science challenges for scholarly publishing. The OA reviews ranged from political frameworks for open access publishing in Canada (Papillon et al., 2019; Price & Puddephatt, 2017) to open access and publishing research funding in Germany (Mayer, 2013), OA monograph publishing (Thatcher, 2015), mega-journals and the open-access business model (Spezi et al., 2017), OA-related hot topics and key trends in scholarly publishing (Smart, 2014; Tennant et al., 2019), subscription-based journals in the OA environment (Laakso, Solomon, & Björk, 2016), predatory publishing and open-access journals (Smith, 2017), the economic perspectives of open access (Frankland & Ray, 2017), OA hybrid model (Mrva-Montoya, 2017), criteria for joining the Directory of Open-Access Journals (Bi, 2017), a national OA journal system in South Korea (Park & Seo, 2016), challenges for the labour in OA publishing (Eve, 2017), article processing charges (Beasley, 2016), etc.

Mega-journals as a new phenomenon in the open-access business model landscape were reviewed using four factors: their scale, disciplinary scope, peer review policy, and economic model (Spezi et al., 2017).

A comprehensive review dealing with hot topics arising from Open Science helps us identify the most heated topics addressed at four levels: policy, science, education, and academic communication. It covers such top priority issues as preprints, copyright transfers, new approaches to peer review, predatory publishing, and the global databases legitimacy (Tennant et al., 2019).

Some of the publications we selected are limited to local or national experiences or environments. They are rather valuable as we learned a lot about evidence in various states and on diverse aspects. A review limited to Spain aimed to consider effective ways for national scholarly sources towards open access, open-peer review, and altmetrics (Segado-Boj, Martin-Quevedo, & Prieto-Gutierrez, 2018). South Korea has a national OA publication platform for distributing research outcomes in Science, Technology, Engineering and Mathematics (STEM). The system encompassing a model and a platform that increases the open-access awareness of researchers as they witness improved knowledge dissemination (Park & Seo, 2016).

Although open-access scholarly publishing in Canada is a growing trend in scientific communication, subscription-based journals prevail and remain trendsetters in social sciences and humanities. Price and Puddephatt (2017)analysed the competing resistance on the part of OA journals in the new challenge-based setting.

The cost of scholarly content for the underfinanced science and research in Sub-Saharan Africa remains a great obstacle to African scholars' bigger contribution to world science (Malapela, 2017).

Seeking sometimes opposite aims (contributions to science versus profit), open-access journals follow various business models and fall into diverse groups (top-ranking journals, predatory journals, new-born journals, established journals, etc.). As predatory publishing, an enormous scholarly publishing obstacle, is widely associated with open-access journals, a new challenge has emerged; journals seek to be whitelisted be researchers. As Open Science aims to deliver scholarly contributions to the wider global readership, most open-access journals think of high-quality publishing as their top priority. All practices are scrutinized through (non)-predatory criteria. The next review calls our attention to comparing OA business models given their potential for abuse and predatory publishing (Smith, 2017).

The Directory of Open Access Journals serves to support the best practices in open-access publishing, improving its quality via ever-rising criteria (Bi, 2017).

One of the reviews dwells upon the economics of e-journals versus traditional (print) sources, outlining that OA publishing in e-journals brings about added costs and increases the potential of misinformation (Frankland & Ray, 2017).

Mrva-Montoya's (2017) review considers an OA hybrid model for a new university press. It combines diversified funding and sustainability with rigorous peer review, and high-standard editing and publishing.

Libraries: A New Strategy for Survival

A few reviews (n=4) considered cloud computing in academic library practices (Mavodza, 2013), library coalitions (Lippincott, 2016), the impact of the serials crisis on library budget shortfalls (Greco, 2015), and the role of librarians in scholarly publishing reforms (Bedenbaugh, 2014). The review by S.K. Lippincott presented the Library Publishing Coalition (LPC), a professional association. With the purpose of knowledge dissemination, it supports library publishers who strive to fill in the gaps in the scholarly publishing system.

Peer Review: Promoting Transparency and Fairness

With all the criticism about peer review, it is still considered valuable if it is well managed. New forms of peer review are emerging. Open peer review, with the identities of the parties being made public, raises the effectiveness of the reviewing process. It improves the quality of submissions.

Among the selected publications, there were six relating to peer review as an important process within the scholarly publishing domain. One of the reviews (Sciullo & Duncan, 2019) raised ethical issues within the context of professionalizing peer review. Another review offered a model based on a review board for scientific publishing to balance the interests of the stakeholders (Hagen, 2018).

Monographs: New Practices

We came across several reviews summarizing new approaches to publishing and distributing monographs. Academic books are mainly brought out by university presses or by commercial publishers. The reviews analyse monographic purchasing trends in academic libraries (Jones & Courant, 2014), the economic aspect of monographic price setting (Greco & Spendley, 2016), the introduction of online content platforms along with traditional print books that essentially enhances research dissemination and raises feasibility of producing books (Mrva-Montoya, 2015), and moves towards more inclusive practices in the university press community (Coggins, Fosado, Henry, & Manaktala, 2020).

As university presses are going through tough times due to decreasing demand for their print books, they tend to turn to open-access or hybrid strategies, ensuring rigorous peer review, high-quality editing and production, and effective marketing (Mrva-Montoya, 2017). More arguments on the background of the crisis of scholarly monograph publishing explain the failures of some moves and projects before the OA business model was introduced (Thatcher, 2015).

A Mixed Bag of Studies

The area of research on scholarly publishing encompasses other issues.

CrossRef was established in 2000 to provide reference linking services. At present, it actively works out and introduces additional services in scholarly publishing, including Cited-By linking, CrossCheck plagiarism screening, and CrossMark update identification (Lammey, 2014).

Some researchers focused on the linguistic aspects of scholarly publishing. In China, researchers' language choices in humanities and social sciences are shifting to the English language in international journals, although it was shown that researchers still have very limited foreign language aptitude (Zheng, & Gao, 2016). Another review (Uysal, 2014) scrutinized the macro-level state policy in Turkey towards scholarly publishing. It revealed that there is a growing dominance of English in academic and scholarly publishing, with some measures to support the national language in science.

One of the gaps in scholarly publishing is found with regard to skilled professionals. Open access and the growing market for publications demand skilled labour. But the human resources needed to meet the high professional standards are insufficient (Eve, 2017). Outsourcing partly provides for some skills in scholarly publishing and communication (Matthews, 2017). The editing aspect of scholarly publishing is rarely discussed, but it plays a very special and vital role in producing high-profile journals (Fretz, 2017).

Conclusion

In their analysis, the JLE editors discovered that the most discussed issues in the research domain of scholarly publishing covered open access initiatives and consequences for academic journals, university presses, commercial publishers, and all other stockholders; their attitudes to open access; a new form of improved peer review corresponding to the needs of the scholarly publishing community and Open Science; new practices and services applied by libraries to remain a part of scholarly publishing; predatory publishing in the context of predatory open access; moves by DOAJ and CrossRef to enhance open access; national/local practices in scholarly publishing. JLE expects that some of our potential writers would like to pinpoint their interests and conduct research on some of the hot topics enumerated above.

References

Beasley, G. (2016). Article processing charges: A new route to open access? *Information Services and Use, 36*(3-4), 163-170. https://doi.org/10.3233/ISU-160815

Bedenbaugh, R. A. (2014). Librarian, interrupted: A plea for scholarly publishing reform. *Contemporary Sociology, 43*(5), 615-618. https://doi.org/10.1177/0094306114546118

Benos, D.J., Bashari, E., Chaves, J.M., Gaggar, A., Kapoor N., LaFrance, M., Mans, R., Mayhew, D., McGowan, Polter, S.A, Qadri, Y., Sarfare, S., Schultz, K., Splittgerber, R., Stephenson, J., Tower, C., Walton, R.G., & Zotov, A. (2007). The ups and downs of peer review. *Advances in Physiology Education*, *31*, 145-152. https://doi.org/10.1152/advan.00104

Bi, X. (2017). Quality open access publishing and registration to directory of open access journals. *Science Editing, 4*(1), 3-11. https://doi.org/10.6087/kcse.82

Coggins, N. L., Fosado, G. C., Henry, C., & Manaktala, G. (2020). Towards inclusive scholarly publishing: Developments in the university press community. *Insights: The UKSG Journal, 33.* https://doi.org/10.1629/UKSG.506

Eve, M. P. (2017). Open publication, digital abundance, and scarce labour. *Journal of Scholarly Publishing, 49*(1), 26-40. https://doi.org/10.3138/jsp.49.1.26

Frankland, J., & Ray, M. A. (2017). Traditional versus open access scholarly journal publishing: An economic perspective. *Journal of Scholarly Publishing, 49*(1), 5-25. https://doi.org/10.3138/jsp.49.1.5

Fretz, M. J. H. (2017). Speaking of editing: The nomenclature of copy-editing. *Journal of Scholarly Publishing, 48*(4), 243-267. https://doi.org/10.3138/jsp.48.4.243

Greco, A. N. (2015). Academic libraries and the economics of scholarly publishing in the twenty-first century: Portfolio theory, product differentiation, economic rent, perfect price discrimination, and the cost of prestige. *Journal of Scholarly Publishing, 47*(1), 1-43. https://doi.org/10.3138/jsp.47.1.01

Gump, S. E. (2018). Special section on the value of scholarly book reviews. *Journal of Scholarly Publishing, 50*(1), 1-7. https://doi.org/10.3138/jsp.50.1.01

Hagen, N. (2018). Review boards for scientific publishing. *Journal of Scholarly Publishing, 49*(4), 419-434. https://doi.org/10.3138/jsp.49.4.03

Jones, E. A., & Courant, P. N. (2014). Monographic purchasing trends in academic libraries. *Journal of Scholarly Publishing, 46*(1), 43-70. https://doi.org/10.3138/jsp.46.1.003

Laakso, M., Solomon, D., & Björk, B. (2016). How subscription-based scholarly journals can convert to open access: A review of approaches. *Learned Publishing, 29*(4), 259-269. https://doi.org/10.1002/leap.1056

Lammey, R. (2014). CrossRef developments and initiatives: An update on services for the scholarly publishing community from CrossRef. *Science Editing, 1*(1), 13-18. https://doi.org/10.6087/kcse.2014.1.13

Lippincott, S. K. (2016). The library publishing coalition: Organizing libraries to enhance scholarly publishing. *Insights: The UKSG Journal, 29*(2), 186-191. https://doi.org/10.1629/uksg.296

Malapela, T. (2017). Access to scholarly research information in sub-Saharan Africa: A review. *Libri, 67*(1), 1-13. https://doi.org/10.1515/libri-2016-0060

Matthews, L. E. (2017). Outsourcing - measuring the status quo. *Insights: The UKSG Journal, 30*(2), 92-100. http://doi.org/10.1629/uksg.368
Mavodza, J. (2013). The impact of cloud computing on the future of academic library practices and services. *New Library World, 114*(3), 132-141. https://doi.org/10.1108/03074801311304041
Mayer, K. U. (2013). Open access improves returns to public research funding: A perspective from Germany. *Information Services and Use, 33*(1), 3-10. https://doi.org/10.3233/ISU-130690
Mrva-Montoya, A. (2015). Beyond the monograph: Publishing research for multimedia and multiplatform delivery. *Journal of Scholarly Publishing, 46*(4), 321-342. https://doi.org/10.3138/jsp.46.4.02
Mrva-Montoya, A. (2017). Open access strategy for a 'new' university press: A view through the stakeholder lens. *Journal of Scholarly Publishing, 48*(4), 221-242. https://doi.org/10.3138/jsp.48.4.221
Papillon, M., O'Neill, B., Bourque, M., Marland, A., & White, G. (2019). Open access and academic journals in Canada: A political science perspective. *Canadian Journal of Political Science, 52*(4), 903-922. https://doi.org/10.1017/S0008423919000799
Park, M., & Seo, T. (2016). Creating a national open access journal system: The Korean journal publishing service. *Journal of Scholarly Publishing, 48*(1), 53-67. https://doi.org/10.3138/jsp.48.1.53
Price, T., & Puddephatt, A. (2017). *Power, emergence, and the meanings of resistance: Open access scholarly publishing in Canada, 48,* 95-115. https://doi.org/10.1108/S0163-239620180000048008
Sciullo, N. J., & Duncan, M. (2019). Professionalizing peer review: Suggestions for a more ethical and pedagogical review process. *Journal of Scholarly Publishing, 50*(4), 248-264. https://doi.org/10.3138/jsp.50.4.02
Segado-Boj, F., Martín-Quevedo, J., & Prieto-Gutiérrez, J. J. (2018). Attitudes toward open access, open peer review, and altmetrics among contributors to Spanish scholarly journals. *Journal of Scholarly Publishing, 50*(1), 48-70. https://doi.org/10.3138/jsp.50.1.08
Smart, P. (2014). The big picture: Scholarly publishing trends 2014. *Science Editing, 1*(2), 52-57. https://doi.org/10.6087/kcse.2014.1.52
Smith, K. L. (2017). Examining publishing practices: Moving beyond the idea of predatory open access. *Insights: The UKSG Journal, 30*(3), 4-10. https://doi.org/10.1629/uksg.388
Spezi, V., Wakeling, S., Pinfield, S., Creaser, C., Fry, J., & Willett, P. (2017). Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. *Journal of Documentation, 73*(2), 263-283. https://doi.org/10.1108/JD-06-2016-0082
Tennant, J. P., Crane, H., Crick, T., Davila, J., Enkhbayar, A., Havemann, J., Thacker, P. D., Vanholsbeeck, M. (2019). Ten hot topics around scholarly publishing. *Publications, 7*(2), 34. https://doi.org/10.3390/publications7020034
Thatcher, S. G. (2015). Open-access monograph publishing and the origins of the office of digital scholarly publishing at Penn State University. *Journal of Scholarly Publishing, 46*(3), 203-223. https://doi.org/10.3138/jsp.46.3.01
Uysal, H. H. (2014). English language spread in academia: Macro-level state policies and micro-level practices of scholarly publishing in turkey. *Language Problems and Language Planning, 38*(3), 265-291. https://doi.org/10.1075/lplp.38.3.03uys
Wolfram, D., Wang, P., Hembree, A., & Park, H. (2020). Open peer review: Promoting transparency in open science. *Scientometrics, 125*(2), 1033-1051. https://doi.org/10.1007/s11192-020-03488-4
Zheng, Y., & Gao, A. X. (2016). Chinese humanities and social sciences scholars' language choices in international scholarly publishing: A ten-year survey. *Journal of Scholarly Publishing, 48*(1), 1-16. https://doi.org/10.3138/jsp.48.1.1

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Alimorad, Z., & Yazdani, S. (2020). Self-Image Improvement and Iranian EFL Learners' Oral Performance: Effects on Complexity, Accuracy, and Fluency. *Journal of Language and Education, 6*(4), 11-29. https://doi.org/10.17323/jle.2020.9772

Self-Image Improvement and Iranian EFL Learners' Oral Performance: Effects on Complexity, Accuracy, and Fluency

Zahra Alimorad, Shiva Yazdani
Shiraz University

Correspondence concerning this article should be addressed to Dr. Zahra Alimorad, Shiraz University, Eram Square, Shiraz, Iran, E-mail: zahra.alimorad@shirazu.ac.ir

The present study investigated the effects of self-image improvement on the quality of Iranian EFL learners' oral production. To this aim, 30 lower-intermediate university students were randomly divided into an experimental (EG) and control (CG) group. The main focus for both groups was enhancing the students' oral performance, however only the EG received the treatment via self-image improving techniques. Their self-image was measured with the aid of the Offer Self-Image Questionnaire (OSIQ) and their oral performance was examined using two parallel IELTS speaking tests before and after the treatment. It was found that by receiving self-image training, the participants' self-image improved, which led to a parallel improvement in their speaking skill. More specifically, the EG improved significantly in oral complexity and two components of oral fluency, mean length of run (MLR) and speech rate A (SRA). However, regarding oral accuracy and speech rate B (SRB), no significant difference was observed between the two groups.

Keywords: accuracy, complexity, fluency, Iranian EFL learners, oral performance, self-image, quality

Introduction

The remarkable role of English is ever growing; never has there been a language so widely spoken as English (Crystal, 2003). Globalization, the development of communicative language teaching, and national and international academic demands are among the many reasons that have led English to become an international language with a prominent role in human lives (Knapp, 2015).

The rapid growth of this global trend coupled with the need for achieving proficiency in oral English ability has led to some scholars in the field focus on discovering the best ways to help improve students' ability to communicate in English (Richard, 2008). To date, psychological factors have received a lot of attention by researchers (Juhana, 2012) due to their seemingly indisputable role in determining improvements in second/foreign language (S/FL) learners' oral proficiency (Haidara, 2014). Researchers have pointed to the relationship between and the effects of a number of these psychological constructs and the level of S/FL oral proficiency. To put it more clearly, studies have been done on the effects of or the relationship between English oral proficiency and self-confidence (Gurler, 2015; Juhana, 2012), learners' attitudes toward English (Norton & Toohey, 2001; Zeinivand, Azizifar, & Gowhary, 2015), self-esteem (Ahour & Hassanzadeh, 2015; Devi & Qiqiehi, 2016; Koosha, Ketabi, & Kassaian, 2011), anxiety (Mai, 2011; Woodrow, 2006), motivation (Ihsan, 2016; Juhana, 2012), shyness (Al Nakhalah, 2016; Ostovar, Safaee, & Sobhanifar, 2015), fear of making mistakes (Sato, 2003), personality types, introversion /extroversion (Mai, 2011; Meihua, 2012), and self-efficacy (Khatib & Maarof, 2015; Liu, 2013).

To communicate successfully in English is to possess a fair level of oral proficiency which, as Omaggio (1986) asserts, equals accurate and functional verbal communication in the target language. Successful communication in English requires learners to produce fluent and accurate language of an appropriate level of structural complexity. Many L2 practitioners and SLA (Second Language Acquisition) researchers hold that S/FL oral proficiency consists of three principal components: complexity, accuracy, and fluency, or CAF (Ellis, 2008; Norris & Ortega, 2009).

Generally, fluency is defined as the ability to produce speech with ease, at a normal rate, without hesitations and unnecessary pausing, manifested in continuous, smooth, and native-like speech (De Jong, 2012; Ellis & Barkhuizen, 2005). As another component of the CAF triad, accuracy (or correctness) is regarded as the most internally consistent one and it refers to the correct and error-free use of language as well as the extent to which S/FL performance deviates from target language norms or sticks to target-like language use (Housen & Kuiken, 2009). Complexity refers to both linguistic and cognitive complexity in SLA literature. Cognitive complexity is the relative difficulty with which language elements are processed during S/FL performance, while linguistic complexity refers to formal and linguistic properties of S/FL elements (Housen & Kuiken, 2009) interpreted as varied, complex, rich, elaborated, and sophisticated target language produced by the learner (Ellis & Barkhuizen, 2005).

Researchers recommend different methods for measuring each of these components. Accuracy can be measured in either a specific or a general manner (Ellis & Barkhuizen, 2005). The former deals with the accurate production of a specific form (suitable for focused tasks) while the latter is inclusive of the learner's overall accuracy (Iwashita, 2010), which is sufficiently informative (Ahmadian & Tavakoli, 2011) and recommended over the specific one (Ellis & Barkhuizen, 2005). According to Delaney (2012), general accuracy can be measured as the ratio of error-free clauses to the total number of clauses.

SLA researchers measure complexity according to syntactic and lexical complexity (Norris & Ortega, 2009). Lexical complexity is measured by D, a sophisticated TTR (the type-token ratio) measure that adjusts for the length of text (McKee, Malvern, & Richards, 2000) and can be measured by CLAN software (Li, Chen, & Sun, 2015). Syntactic complexity is also measured as the ratio of total clauses to AS-units (Analysis of Speech Unit: a main clause and any attached subordinate clauses or sub-clausal units) (Delaney, 2012).

Fluency is measured as the mean length of run (MLR) and speech rate, which is the most common and trusted way according to previous studies (e.g., Li et al., 2015, Skehan & Foster, 2005). According to Li et al. (2015), MLR is measured by the mean number of syllables produced in utterances between pauses with the standard of 0.28 seconds being considered the cutoff point of a pause. Speech rate includes speech rate A (SRA), which is the ratio of syllables to the complete time it takes to produce them, and speech rate B (SRB), which is similar to SRA; however, syllables, words, phrases, or clauses that are repeated, reformulated, or replaced are not considered.

Literature Review

The available literature is replete with various psychological constructs beginning with self, such as self-confidence, self-concept, self-efficacy, and self-esteem from which self-image can conspicuously be distinguished. Self-image is a mental blueprint and a cognitive representation of self, personality, mental functioning, social attitudes and a base upon which one's personality, behaviors, and life circumstances are built (Di Blasi et al., 2015; Maltz, 1960; Proctor, 2002) which is created from an individual's beliefs about oneself (Maltz, 1960; Proctor, 2002); beliefs that have unconsciously been formed from past experiences of success, failure, happiness, and humiliation, especially in childhood (Maltz, 1960; McMullen, 2013; Proctor, 2002).

Self-image is thus a multidimensional construct manifested through performing in various social settings and psychological functioning. Therefore, this construct should be measured as a multifaceted concept as well (Offer, Ostrov, Howard, & Atkinson, 1988). Offer et al.'s (1988) conceptualization of self-image encompasses five main selves: 1) Psychological self which comprises concerns, feelings, wishes, and fantasies 2) Social Self which revolves around the individual's social life 3) Sexual self and an individual's sexual drives 4) Familial self and the feelings and attitudes individuals have toward their families, and 5) Coping self and how individuals cope with their world. Although each self shows a different aspect of self-image, together, they provide a thorough depiction of one's self-image (Offer et al., 1988).

Once formed, self-image acts as an inner thermostat directly controlling one's behaviors, actions, feelings, abilities, life circumstances, goals, end results, and even others' reactions to oneself. Ideas that are inconsistent with self-image would be rejected, not believed and not acted on. While human beings proceed to act upon their self-image without questioning its validity, no one would be able to outperform or act despite it. Even if one manages to briefly escape, they will be snapped back quickly (Maltz, 1960; Proctor, 2002) just as the person with a failed self-image will find a way to fail in spite of all their conscious efforts, willpower, and good intentions (Hirsch, Meynen & Clark, 2004; Maltz, 1960; Proctor, 2002). Nonetheless, self-image is a dynamic construct (Offer et al., 1988) that, if understood and managed, can be changed and improved to suit life purposes, and become the key to living without limit (Maltz, 1960; Proctor, 2002) since "it controls mind, just as surely as mind controls heartbeats" (Proctor, 2002, p. 7).

As an example of the early studies on the relationship between self-image and educational performance, Sopis (1965) investigated the self-image of students with reading problems. She hypothesized that students' poor performance was the result of their inaccurate definition of themselves. To examine this, she placed 240 students from grades two, three, four, and five in some reading sessions. Participants defined themselves using the Colvin Silhouette test as having high, average, or low self-images in reading. The results indicated a direct correlation between pupils' performance and their level of self-image.

In another study, Glazier (1998) witnessed that students in remedial English writing courses suffered from poor self-image. They had been labeled failures before and those remedial courses were only second reminders. The researcher proposed six ways to help those disheartened students improve their self-image and English writing skills. Alleviating failure, freeing students from inhibitions, and giving students a sense of confidence and success were among the self-image improvement principles proposed and used by the researcher. As a result of implementing the proposed principles in writing courses at Western Illinois University, many students in those courses showed significant improvement. Later, their other writing instructors also reported that the students performed even better than many regular students. They were said to have become confident, handed in their papers without failure, prepared well-organized papers, come to classes motivated, become ready to learn, and taken criticism well because they believed they could improve; and the biggest change was said to be that of their attitudes about themselves.

Using a different approach, Lee and Oxford (2007) narrowed down the perspective and focused only on English learning self-image and its effects on students' English performance. More than 500 Korean students from both traditional and vocational schools participated in the study, and their self-image was assessed using the Background Information Questionnaire. The results showed that the vocational students' poor performance was the result of their impaired English-learning self-image.

Sandi-Urena and Gatlin (2013) also investigated the effects of self-image on academic achievement. They examined the association between the self-image of 24 laboratory graduate teacher assistants (GTAs) and their performance in the labs as laboratory instructors. Information on self-image and teaching styles were collected using interviews, memos, training artifacts, laboratory manuals, and syllabi. The findings showed that GTAs' performance was in accordance with their type of self-image.

Reviewing the related literature, the present researchers did not find any studies on self-image and academic performance for EFL learners and, in particular, their speaking performance. Therefore, to fill this gap, the present study can serve as a stepping stone for further studies on this under-researched area of enquiry.

Building upon the notion that the prime cause of success or failure is one's own self-image (Maltz, 1960; Proctor, 2002), and that many underachieving learners are simply the victims of their own negative self-image, this study aimed at helping English learners improve their oral proficiency through, first, improving their self-image using psychological techniques. To this end, attempts were made to transform students' self-image into more realistic, positive, and success-oriented ones by employing scientific, and psychological techniques put forth by Maltz (1960). Considering the objectives, the present study was motivated by two research questions:

RQ1: Does Iranian EFL learners' self-image improve as a result of receiving *self-image psychological training techniques?*

RQ2: Is there any significant difference between the performance of students who received the self-image improvement techniques with those who did not in terms of (a) the accuracy of their oral production, (b) the complexity of their oral production, (c) the fluency of their oral production?

Methodology

Participants

Thirty undergraduate university students agreed to participate in this study. To determine their proficiency level, Oxford's QPT (Quick Placement Test) was given to the students. All of the participants' (female = 20, male = 10) native language was Persian and their age ranged from 20 to 25. To determine the effects of the treatment, they were randomly assigned into a control group (CG, N=15) and an experimental one (EG, N=15). All of the participants in both groups were assured of the confidentiality of their identities throughout the study as well as in the report of its results.

Instruments

Two main instruments were used to collect the needed data. However, as mentioned above, since it was necessary to determine the proficiency level of the students in order to ensure that the two groups were homogeneous prior to the treatment, the QPT was used as a prerequisite for the experiment. This test consisted of two parts. Part One (questions 1-40) was taken by all candidates and was aimed at participants who were at or below an intermediate level. Part Two was geared towards more proficient participants and those who scored more than 35 on the first part of the test (Geranpayeh, 2003).

As one of the main instruments, two parallel IELTS speaking interviews were used as pre- and post-speaking tests. While ideally it would have been better to use the same measurement utilized in the pre-test (i.e., IELTS speaking Test 1) to examine students' progress in their speaking skills as the result of the experiment, the practice and memory effect bias could have not been avoided (Ary, Jacobs, Sorensen Irvin, & Walker, 2019; Cliffordson, 2004; Geving, Webb, & Davis, 2005; Hausknecht, Halpert, Di Paolo, & Gerrard, 2007; Lievens, Buyse, & Sackett, 2005). That is why two different, but parallel, IELTS speaking tests were utilized in this study. IELTS tests in general, and IELTS speaking test in particular, have a very structured nature and are of the same level of difficulty (Brown & Hill, 2007). The tests were drawn from the book IELTS 11 (Appendix A). In this study, all of the participants in both the EG and CG participated in the oral interviews conducted at the beginning and end of the experiment by the second researcher under the direct supervision of the first researcher who was also the main supervisor of this study. The second researcher was a trained interviewer and conducted all the oral speaking tests herself.

The oral interviews were all recorded and transcribed in order to assess different components of CAF. In order to determine the inter-coder reliability of the oral speaking tests, the second researcher was trained by her supervisor to code the data. Then, a randomly selected sample of the data was given to another M.A. student who had already done the same coding for his M.A. thesis and was familiar with the procedures. The correlation between the two sets of coding was calculated and turned out to be 98%, which was satisfactory for the purposes of the current study.

Moreover, the participants' level of self-image was assessed using an adapted version of the Offer Self-Image Questionnaire (OSIQ) originally developed by Daniel Offer and his colleagues in 1961. It was a 99-item inventory. Six items were omitted due to their being culturally inappropriate for the context of Iran. The questionnaire was then translated by the second researcher into Persian and after a back-translation by a graduate English student from the Department of Foreign Languages and Linguistics of Shiraz University, some items were modified. The final draft of the translated questionnaire was approved by an assistant professor of TEFL, who was an expert in the field, from the same department. Responses were elicited on a 6-point Likert–type scale form *describes me very well* to *does not describe me at all*. Based on the data from the 30 participants of the study, its Cronbach alpha reliability index was estimated using SPSS software and turned out to be 0.788, which seemed acceptable for the purposes of the current study.

Materials

Maltz's (1960) worksheet was used as a self-improvement aid in this study. To help people improve their self-image, in his book *Zero Resistance Living*, Maltz (1999) presents some psychological techniques in the form of a 12-session worksheet. Each lesson in the worksheet is divided into two main parts. The first part contains an introduction to the lesson plus some inspirational anecdotes and scientific facts about human self-image and psychological behavior while the second part is devoted to exercises for each lesson (See a sample of Lesson 1 of this worksheet as an example in Appendix B).

Treatment

Given that the students were undergraduate students of diverse majors, it was envisaged that they were busy taking different classes. Therefore, to avoid interfering with their university schedule, the experiment was conducted in the evenings. During the experimental treatment, the second researcher orally presented the first part of each self-image improvement lesson to the participants in the EG in their native language. This part usually lasted between ten to fifteen minutes. The rationale behind using their native language during self-image improvement training periods was to avoid providing EG students with extra exposure to English as compared with the CG participants. PowerPoint presentations were used when needed. The next part of the lesson, which went on for about twenty minutes (the recommended time span proposed by Maltz, 1960), and which consisted of self-image improvement exercises, was translated into Persian and handed to the participants each session. Once the self-image improvement part of the EG's class was over (about the first thirty minutes of each session), the class proceeded with English speaking activities that were identical in both the EG and CG. The entire experiment lasted for 12 sessions (as recommended by Maltz, 1960), three days a week for two weeks. This study schedule was adopted in order to avoid interrupting the participants' university schedule and their impending final exams. In order to determine the integrity of the treatment, a professor of psychology at Shiraz University agreed to help the researchers in this regard. He was consulted with prior to the experiment; and, during repeated and continuous sessions, he passed on the required therapeutic skills to the second researcher who was responsible for carrying out the self-image treatments. The self-image treatment sessions were regularly reported to that professor so that adherence to the manual would be checked as well. Therefore, benefiting from the direct supervision of the psychology professor and by closely following his advice and recommendations, the two researchers of this study were able to maintain the adherence/competence protocol (Mcleod et al., 2017; Waltz, Addis, Koerner, & Jacobson, 1993) regarding the psychological part of their experiment (i.e., the self-image improving treatment).

As mentioned above, after presenting the self-image improving techniques in the EG, their class proceeded with English speaking activities that were identical to those of the CG. That is, to improve their English-speaking skill, the same procedure was followed in both the EG and CG. A point that must be borne in mind is that since the participants' command of English was at a low-intermediate level, holding the class as a mere chat class was not an option. For this reason, some parts of the two books *Top Notch Fundamental B* and *New Interchange 2* were selected to be taught to the learners so that the participants could learn to give speeches on various topics. Speaking topics were elicited from the *New Interchange* book. Selected topics for speaking from the *New Interchange* book were introduced, proper and related structures were explained, and then related vocabulary for each topic was taught.

The participants were encouraged to hold discussions with their peers rather than simply offer answers to the instructor. Moreover, since learners had little to no exposure to authentic spoken English outside the classroom, they were asked to watch two seasons of the four-season Disney original teen series *Wizards of Waverly Place* each week. They were then asked to give oral and written summaries and to write down all the new words and expressions they heard and understood. Occasionally, some parts of the series were played in the class as a way to teach the participants new vocabulary and everyday expressions. Furthermore, to meet the purpose of the experiment, the EG was asked to watch the aforementioned series and then, as their self-image homework practice, imagine themselves being able to speak just like those native speakers they had watched and heard.

Data Collection and Analysis Procedures

The selected sample was randomly assigned to an EG and a CG. The OSIQ was given to the participants in both groups at both the outset and the end of the study to assess any improvements in their self-image. It is worth mentioning that no time limit was defined for the completion of the questionnaires. Likewise, two IELTS speaking tests with similar topics were selected as parallel tests for the pre- and post-tests in the study. Participants' responses to the oral IELTS interviews were recorded and transcribed and the quality of their productions was measured by tapping its components; namely, accuracy, fluency, and complexity. Accuracy was measured as the ratio of error-free clauses to the total number of clauses (Delaney, 2012). Following Foster and Skehan's (1996) definition, an error-free clause is "a clause in which there is no error in syntax, morphology, or word order. Errors in lexis were counted when the word used was incontrovertibly wrong. In cases of fine decisions of appropriacy, no error was recorded" (p. 310).

Two of the most commonly used fluency measures were used in this study: MLR and speech rate (Li et al., 2015). MLR was measured by counting the number of syllables produced between pauses lasting more than 0.28 seconds. Speech rate was also divided into SRA and SRB. The former was measured as the ratio of syllables to the time it took the participant to do the task, and for measuring the latter, the same procedure was followed except that syllables, words, phrases, or clauses that were repeated, reformulated, or replaced were not considered (Li et al., 2015).

Complexity was measured in terms of the two categories of syntactic complexity and lexical diversity. Syntactic complexity was measured as the ratio of clauses per AS-unit. Lexical complexity was also measured by calculating the D value, which is considered a more reliable measure than type-token ratio (Malvern & Richards, 2002) using CLAN software (Delaney, 2012; Li et al., 2015). To analyze the data, the experimental treatment was considered as the independent variable and values were assigned to the participants according to their presence in either the EG or CG. More specifically, to uncover any differences between these two groups in terms of their improvements regarding oral performance accuracy and given that oral accuracy had only one aspect (i.e., the ratio of error-free clauses to the total number of clauses), which was considered as the dependent variable, an independent samples t-test was run[1]. Moreover, since oral complexity was measured by analyzing the two components of lexical diversity and syntactic complexity, that is, there were two dependent variables, a MANOVA test was run[2]. Additionally, as oral fluency was assessed by calculating its three components (i.e., MLR, SRA, and SRB), another MANOVA test was conducted[3]. Before embarking on MANOVA, its assumptions needed to be checked[4]. To check multivariate normality, Mahalanobis Distances was also calculated. Finally, the homogeneity of variance-covariance matrices was examined by considering Box's M Test of Equality of Covariance Matrices and Levene's Test of Equality of Error Variances was used to make sure that the assumption of equality of variance for each variable was not violated[5]. All of the above-mentioned calculations were done using SPSS software.

Results

To ensure the homogeneity of the two groups prior to the experiment, two independent samples t-tests and two MANOVAs were run on the scores attained from the first administration of the OSIQ as well as the participants' pre-test oral scores. The results indicated that the two groups were homogenous in terms of accuracy, complexity, and fluency as well as their self-image at the outset of the study.

Effect of the Psychological Training Techniques on Learners' Self-Image

To examine the effectiveness of self-image psychological training, independent samples t-tests were run on data attained from pre- and post-test administration of the OSIQ in both the EG and CG. These results are displayed in Table 1.

Table 1
Independent Samples t-test for Self-image Pre- and Post-test Results

	Pre-test						Post-test							
	N	**M**	**SD**	**f**	**t**	**Sig.**	**N**	**M**	**SD**	**f**	**t**	**Sig.**	**Effect size**	**Cohen's d**
EG	15	311.80	18.27	0.02	1.21	.23	15	316.33	12.75	.08	1.26	.000	.58	1.5
CG	15	302.38	22.58				15	292.13	15.54					

[1] Pallant, J. (2011). *SPSS Survival manual: A step by step guide to data analysis using SPSS for windows.* McGrew Hill.
[2] Ibid.
[3] Ibid.
[4] Ibid.
[5] Ibid.

As depicted in Table 1, the results of the independent samples t-tests revealed that although the two groups had been homogeneous in terms of their self-image when the study commenced (p=.23>0.05), a significant difference in the self-image level of the EG (M = 316.3, SD = 12.7) and CG (M = 292.13, SD = 15.5; t (28) = 1.26, p=0.00<0.05, two-tailed) was observed after the treatment. Moreover, by measuring Cohen's d value (d = 1.5), it was further confirmed that the changes between the two groups regarding improvements in terms of the quality of their self-image was rather meaningful (Plonsky & Oswald, 2014).

Comparing EG and CG oral English accuracy

In order to determine whether the two groups' level of oral accuracy underwent any changes during the course of the experiment, two paired samples t-tests were run and the results are provided in Tables 2 and 3.

Table 2
Paired Samples Statistics

		Mean	N	SD	Std. Error Mean
Pair 1 (EG)	Pretest	.2993	15	.10437	.02695
	Posttest	.3960	15	.11096	.02865
Pair 2 (CG)	Pretest	.3020	15	.06570	.01696
	Posttest	.3613	15	.07337	.01895

Table 3
Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper			
Pair 1	Pretest – Posttest (EG)	-.09667	.08723	.02252	-.14497	-.04836	-4.292	14	.001
Pair 2	Pretest-Posttest (CG)	-.05933	.02764	.00714	-.07464	-.04403	-8.315	14	.000

As revealed in these two tables, there was a statistically significant increase in their scores from pre-test (M=.29, SD=.10) to post-test (M=.39, SD=.11), t (14) = -4.29, p <. 0005 (two-tailed) for the EG and a parallel increase in the scores from pre-test (M=.30, SD=.06) to post-test (M=.36, SD=.07), t (14) = -8.31, p <. 0005 (two-tailed) for the CG. The mean increase in accuracy scores of the EG and CG was -.09 and -.05 with a 95% confidence interval ranging from -.14 to -.04 and -.07 to -.04, respectively. Ethically speaking, we did expect that learning would happen in both groups after a period of instruction but what we were more concerned about was the effect of the treatment on the accuracy scores of the EG compared to the CG, which was investigated using an independent samples t-test.

Table 4
Independent Samples t-test for Oral English Accuracy Pre- and Post-test Results

	Pre-test					
	N	M	SD	f	t	Sig.
EG	15	0.29	0.10	1.98	-0.08	.93
CG	15	0.30	0.06			

	Pre-test						
	N	M	SD	f	t	Sig.	Cohen's d
EG	15	0.29	0.10	1.98	-0.08	.93	0.27
CG	15	0.30	0.06				

Table 4 indicates that in terms of the accuracy of their oral production, no statistically significant difference was seen after the implementation of the treatment in the EG (M= 0.39, SD= 0.11) and CG (M= 0.36, SD= 0.07; t (28) = 1.009, p=.32>0.05, two-tailed). Cohen's d value (d = 0.27) further indicated that the difference between the two groups in terms of their oral accuracy was not meaningful (Plonsky & Oswald, 2014). To answer the first part of the second research question, it must be pointed out that although oral accuracy in the EG improved (M pre-test=0.29, M post-test=0.39), this improvement was not significantly different from that of the CG (M pre-test=0.30, M post-test=0.36).

Comparing EG and CG Oral English Complexity

Prior to running MANOVA, its assumptions had to be checked[6] using SPSS. It was found that all assumptions were met (Mahalanobis distance value=4.38<13.82; Box's test=0.916>0.001; Levene's test=0.899, 0.647>0.05), which let the researchers proceed to run the MANOVA test itself. Table 5 shows the output of this test.

Table 5
Multivariate Tests for Oral English Complexity

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilks' Lambda (Pre-test)	.830	2.769[a]	2.000	27.000	.081	.170
Wilks' Lambda (Post-test)	.567	10.316[b]	2.000	27.000	.000	.433

As Table 5 depicts, despite being homogeneous in terms of their oral English complexity at the outset of the study, F (2, 27)=2.769, p=.081>0.05, Wilks' Lambda=.830, eta squared=.170, a statistically significant difference was found between the EG and CG with respect to the effect of the treatment on both dependent variables of oral complexity, F (2, 27) = 10.31, p=.000< 0.05, Wilks' Lambda= 0.56, eta squared= 0.43. To examine the main effect of the treatment on the degree of change for each dependent variable, independent samples t-tests were run on data from the pre- and post-interviews for both the EG and CG. These results are displayed in Table 6.

Table 6
Independent Samples t-test for Oral Complexity Pre- and Post-test Data

	Pre-test			Post-test				Cohen's d
	Groups	Mean	SD	Mean	SD	Sig.	t	
LC	EG	56.57	4.05	72.00	7.80	0.001	3.89	1.46
	CG	51.18	7.74	61.10	7.52			
SC	EG	1.12	0.09	1.22	0.07	0.000	3.98	1.42
	CG	1.11	0.7	1.12	0.07			

The results in Table 6 show that participants' oral complexity in terms of lexical complexity in the EG (M= 72.00, SD= 7.80) and CG (M= 61.10, SD= 7.52; t (28) = 3.89, p= 0.001<0.05) was significantly different. Moreover, Cohen's d value (d = 1.46) revealed that the difference between the two groups in terms of their lexical complexity was meaningful (Plonsky & Oswald, 2014). In terms of syntactic complexity, a statistically significant difference was also seen between the EG (M= 1.22, SD= 0.07) and CG (M= 1.12, SD= 0.07; t (28) = 3.98, p=.000<0.05). Measuring Cohen's d value (d = 1.42) further indicated that the changes between the EG and CG regarding their level of syntactic complexity was rather meaningful (Plonsky & Oswald, 2014).

Comparing the EG's and CG's oral English Fluency

In this section, to overcome the problem of correlated observations, p-values less than .01 were considered statistically significant, as recommended by Stevens (2009). Furthermore, before conducting MANOVA, its assumptions were checked and all were met (Mahalanobis distance value=10.94<16.27; Box's test=0.418>0.001; Levene's test=0.32, 0.38, 0.39>0.05). Table 7 shows the output for these results.

Table 7
Multivariate Tests for Oral English Fluency

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Wilks' Lambda (Pre-test)	.972	.248[a]	3.000	26.000	.862	.028
Wilks' Lambda (Post-test)	.256	25.150[b]	3.000	26.000	.000	.744

As Table 7 shows, the two groups were homogeneous in terms of their oral English fluency before the treatment, F(3, 26)=.248, p=.862>0.01, Wilks' Lambda=.972, eta squared=.028; however, a statistically significant difference was detected between the two groups' oral fluency, F (3, 26) = 25.150, p=.000<0.01, Wilks' Lambda= 0.256, eta squared= 0.74. To examine the effect of the treatment on each of the dependent variables of fluency separately,

[6] Pallant, J. (2011). *SPSS Survival manual: A step by step guide to data analysis using SPSS for windows.* McGrew Hill.

as follow-up analyses, three independent samples t-tests were run on MLR, SRA, and SRB. These results are displayed in Table 8.

Table 8
Independent Samples t-test on Post-test Results of Oral English Fluency

		Mean	SD	Sig.	t	Cohen's d
MLR	EG	3.53	0.74	0.000	4.30	1.81
	CG	2.48	0.58			
SRA	EG	68.02	13.22	0.000	5.77	0.27
	CG	65.67	8.57			
SRB	EG	74.70	11.10	0.569	5.77	2.62
	CG	54.56	7.66			

As shown in Table 8, a statistically significant difference was found in terms of MLR and SRA scores ($p=.000<0.01$). On the contrary, no statistically significant changes were seen in the scores of the EG ($M= 74.70$, $SD= 11.10$) and CG ($M= 54.56$, $SD= 7.66$; $t\ (28) = 5.77$, $p=0.569> 0.1$, two-tailed) regarding SRB. Moreover, and drawing on Plonsky and Oswald (2014), to supplement information from statistical significance tests, the "practical significance" (p. 879) of oral fluency measurement was addressed as well. As a result, Cohen's d value was measured for each oral fluency variable (i.e., MLR, SRA, and SRB), and interpreted based on Plonsky and Oswald's (2014) L2 specific benchmark proposed for interpreting practical significance in L2 studies; where a d value in the neighborhood of 0.40 is generally considered small, 0.70 medium, and 1.00 large. Therefore, and as shown in Table 8, practical significance was found in terms of MLR ($d= 1.81$) and SRB ($d= 2.62$). However, no practical significance was seen in the EG and CG regarding their SRA ($d=0.27$). In other words, and statistically speaking, the EG improved in terms of MLR and SRA. But regarding practical changes, the EG significantly progressed in terms of MLR and SRB but no practically significant improvement was observed in terms of SRA.

Discussion

As a result of the self-image treatments, and according to the findings, it can convincingly be argued that self-image training techniques were conducive to developing stronger and better-formed self-images for the EG participants who received the treatment. Comparing the CG's mean score gained on the first administration of the OSIQ ($M = 302.38$, $SD = 22.58$) and the second one ($M = 292.13$, $SD = 15.54$), it would be safe to say that not only did the level of self-image in the CG not remain constant, but it diminished slightly, which is in line with Maltz's (1960) remark that overall, the unconscious mind (where self-image is located (Maltz, 1960)) tends to draw on negativities when overlooked rather than on positive thoughts and feelings. Therefore, a mindful approach towards self-image is highly recommended (Maltz, 1960, Proctor, 2002). Improvement in self-image as a result of receiving an intervention is in line with previous studies conducted by Hintikka et al. (2006) and Ryum, Vogel, Walderhaug, and Stiles (2015) in which participants'/patients' levels of self-image improved significantly, almost to that of a normal person, after undergoing the treatments.

As for changes in terms of learners' levels of oral accuracy, the findings enjoy the support of previous research that reported the rather slow development of accuracy over other oral production components like fluency due to the existence of the trade-off effect (Skehan, 2009). However, the same finding is not in line with those of Li et al. (2015) and Vercellotti (2015) who reported a rather concurrent change across all CAF components. Nevertheless, this finding may be justified on the grounds that the participants might have focused more on meaning and not form, and as Skehan (2009) suggested, a trade-off exists between fluency and accuracy. Thus, due to the shortage of memory to pay attention to both the accuracy and fluency of their oral production (Skehan, 2009), they were not able to achieve significant improvements in terms of oral accuracy within the time span of the experiment. The same finding, however, contradicts the results of Meisel, Clahsen, and Pienemann (1981) who reported the positive effects of learners' motivation, as another psychological construct, on reaching a high level of oral accuracy.

To answer the second part of research question two, it can be argued that the findings showed a statistically significant difference between the performance of the EG and CG in terms of both lexical and syntactic complexities. Improvement in terms of oral complexity is in line with a number of previously conducted research projects on other psychological constructs. Li et al. (2015) reported improvement in complexity as a result of receiving speaking treatment. The results are also in line with Park and Lee's findings (2005) which showed that participants with high levels of self-confidence and low anxiety had stronger English oral proficiency in terms of ease of speech, interaction, delivery, and complex use of language. Nevertheless, to the best of the present researchers' knowledge, no study has examined the effects of self-image improvement on S/FL learners' oral complexity. Therefore, this area of inquiry warrants further investigation to confirm or refute these findings.

With regard to the learners' oral fluency, the results indicated that the EG statistically improved in terms of MLR and SRA, but practically (Plonsky & Oswald, 2014) only progressed in terms of MLR and SRB. One possible justification for the small d value for SRA ($d = 0.27$) is that this research was conducted in a classroom context in contrast to the controlled environment of a laboratory where greater rigor and experimental control is feasible; hence higher d values (Cohan, 1988). Also, since the CG in this study was not a true control group but a comparison group (they received the traditional treatment), the effect size for SRA was not that large (Norris, 2012; Plonsky & Oswald, 2014). The nature of the research must also be put in perspective when interpreting effect size values (Lipsey, 2009; Vacha-Haase & Thompson, 2004). Therefore, the lower d value of oral accuracy ($d = 0.27$) compared to the high d value of oral fluency (in general), can be due to the very existence of the trade-off between oral fluency and oral accuracy as first reported by Skehan (2009).

All that being said, it must be borne in mind that, at times, the minimal time, costs, and resources needed to implement a certain technique, despite its low effect size values, and considering the potential beneficial results of implementation, justify the interventions (Plonsky & Oswald, 2014); as in the case of Lee and Huang (2008) and Alsadhan's (2011) study, which showed that implementing their input enhancement and noticing techniques proved to be useful in practice despite the initial low d values of .22 and .30. In this study, too, implementing self-image improvement techniques as a way to enhance the quality of English learners' oral fluency was cost effective and feasible. And considering the possible positive effects, it sounds fine to apply these techniques despite the low numerical d values in the case of SRA ($d= 0.27$), and /or oral accuracy ($d= 0.27$).

Overall, regarding fluency improvements, the results were in line with some other studies in the field. For example, Skehan (2009) found that as complexity improves, so does fluency. In this study, too, a trade-off was detected between form (complexity, accuracy) and fluency, and between complexity and accuracy. However, the same findings were not congruent with Vercellotti's (2015) research. In her research, after receiving the treatment, all CAF components improved in a linear fashion and no trade-off was observed among them. Nevertheless, she also reported a linear relationship between complexity and fluency. Our findings were also in line with studies by Thai and Boers (2015) and De Jong (2012), who showed the positive effects of 4/3/2 activities on fluency. In terms of improvement regarding SRA and SRB, our findings were in line with that of Li et al.'s (2015). They concluded that in the case of pre-task planning time –another speaking technique used in the treatment- MLR and SRB improved while SRA did not. The same findings were in accordance with Yurong and Yin (2008) who found a correlational relationship between self-esteem, motivation, anxiety, personality type, and fluency. Nevertheless, it seems necessary to conduct more studies to support these findings, which points to the need for further research in this area.

Conclusion

The results of this study showed that students' self-image could determine their success, or lack thereof, when it comes to their English-speaking skills in that the EG students managed to incorporate within their realistic and healthy self-images the image of a successful language learner as well. Like most other studies, however, the present study is subject to several limitations that could restrict the generalizability of its findings. Firstly, the study was limited to the low-intermediate English proficiency level and factors such as the age and gender of the participants were not controlled. Secondly, the sample size was rather small and the sample was not randomly selected as the participants had signed up for the classes voluntarily. Despite these limitations, the findings can have valuable implications for the theory and practice of language teaching. From a theoretical point of view, the results of the present study supported the idea that self-image is dynamic and flexible and can be managed and improved by receiving proper psychological training and implementing appropriate techniques. From a practical point of view, on the other hand, the present findings provide evidence on the effects of self-image as a psychological factor affecting English learners' oral performance. These findings could draw the practitioners' attention to redesigning their teaching practices by taking into account the need for improving their learners' self-image. In this sense, this research is applicable especially to English speaking courses for a range of proficiency levels where the focus is on one or all CAF components. A second important practical implication of this study could be geared toward pre- and in-service teacher training courses where they can learn about the important role of self-image so that they could converge self-image improving techniques into their teaching approaches. As the third practical implication, the findings could also be beneficial to curriculum and syllabus designers, making them aware of the potential for incorporating these techniques into the design of curricula and syllabi in an attempt to achieve more satisfactory outcomes.

As the results of this study demonstrated the role self-image could play in the quality of low-intermediate Iranian learners' L2 oral performance, future studies could be designed to cover other oral proficiency levels, i.e. advanced, upper-intermediate, and beginner levels. Additional work is also required to determine whether this treatment could be as effective as it was here in the case of other language skills (reading, writing, and listening) or any of their subskills. Finally, future studies can examine the long-term effects of implementing self-image training techniques on the students' oral proficiency by administering delayed post-tests in addition to the immediate ones.

Conflict of Interests

The authors declare that they have no conflicts of interest.

References

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35-59. https://doi.org/10.1177/1362168811425433

Ahour, T., & Hassanzadeh, Z. (2015). An investigation of the relation between self-esteem, indirect strategy use and Iranian intermediate EFL learners' oral language proficiency. *Theory and Practice in Language Studies, 5*(2), 442-451. https://doi.org/10.17507/TPLS.0502.28

Al Nakhalah, A. (2016). Problems and difficulties of speaking that encounter English language students at Al Quds Open University. *International Journal of Humanities and Social Science Invention, 5*(12), 96-101.

Alsadhan, R. O. (2011). *Effect of textual enhancement and explicit rule presentation on the noticing and acquisition of L2 grammatical structures: A meta-analysis* [Unpublished doctoral dissertation]. Colorado State University.

Ary, D., Jacobs, L. C., Sorensen Irvin, C. K., & Walker, D. A. (2019). *Introduction to research in education* (10th ed.). Wadsworth, Cengage Learning.

Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing, 20*(1), 1-25. https://doi.org/10.1191/0265532203lt242oa

Brown, A., & Hill, K. (2007). Interviewer style and candidate performance in the IELTS oral interview. In L. Taylor & P. Falvey (Eds.), *IELTS collected papers: Research in speaking and writing assessment* (pp. 37-61). Cambridge University Press.

Cliffordson, C. (2004). Effects of practice and intellectual growth on performance on the Swedish Scholastic Aptitude Test (SweSAT). *European Journal of Psychological Assessment, 20,* 192–204. https://doi.org/10.1027/1015-5759.20.3.192

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.

De Jong, N. H. (2012). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics, 36*(2), 223–243. https://doi.

org/10.1017/S0142716413000210
Delany, T. (2012). Quality and quantity of oral participation and English proficiency gains. *Language Teaching Research, 16*(4), 467 – 482. https://doi.org/10.1177/1362168812455586
Devi, S., & Qiqiehi, S. (2016). The relationship between English language proficiency, academic achievement and self-esteem of non-native-English-speaking students. *International Education Studies, 9*(5), 147-155. https://doi.org/10.5539/IES.V9N5P147
Di Blasi, M., Cavani, P., Pavia, L., Baido, R., Grutta, S., & Schimmenti, A. (2015). The relationship between self-image and social anxiety in adolescence. *Child and Adolescent Mental Health, 20*(2), 74-80. https://doi.org/10.1111/camh.12071
Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.
Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford.
Foster, F., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323. https://doi.org/10.1075/tblt.1.16the
Geranpayeh, A. (2003). A quick review of the English quick placement test. *Research Notes, 12*(3), 8–10.
Geving, A. M., Webb, S., & Davis, B. (2005). Opportunities for repeat testing: Practice doesn't always make perfect. *Applied H.R.M. Research, 10,* 47–56. https://doi.org/10.1037/e518612013-432
Glazier, T. F. (1998). *Improving the poor self-image of the remedial student*. Paper presented at the Conference on College Composition and Communication, Denver, the United States.
Gurler, I. (2015). Correlation between self-confidence and speaking skill of English language teaching and English language and literature preparatory students. *Current Research Journal of Social Sciences, 1*(2), 14-19.
Haidara, Y. (2014). Psychological factors affecting English speaking performance for the English learners in Indonesia. *Universal Journal of Educational Research, 4*(7), 1501-1505. https://doi.org /10.13189/ujer.2016.040701
Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Gerrard, M. O. (2007). Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology, 92,* 373–385. https://doi.org /10.1037/0021-9010.92.2.373
Hintikka, U., Marttunen, M., Pelkonen, M., Laukkanen, L., Viinamaki, H., & Lehtonen. J. (2006). Improvement in cognitive and psychosocial functioning and self-image among adolescent patients with suicide attempts. *BMC Psychiatry, 30*(1), 1-10. https://doi.org /10.1186/1471-244X-6-58
Hirsch, C. R., Meynen, T., & Clark, D. M. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory, 14*(2), 496-506. https://doi.org /10.1080/09658210444000106
Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461-473. https://doi.org/10.1093/applin/amp048
Ihsan, M. (2016). Students' motivation in speaking English. *Journal of English Educators Society, 1*(1), 31-48. https://doi.org/10.11648/j.ijla.20190706.18
Iwashita, N. (2010). Features of oral proficiency in task performance by EFL and JFL learners. In Prior, T., Watanabe, Y. & Lee, S. (Eds.), *Selected proceedings of the 2008 second language research forum: Exploring SLA perspectives, positions, and practices* (pp. 32-47). Cascadilla Proceedings Project.
Juhana, J. (2012). Psychological factors that hinder students from speaking in English class (a case study in a senior high school in south Tangerang, Banten, Indonesia). *Journal of Education and Practice, 3*(12), 100-110.
Khatib, F., & Maarof, N. (2015). Self-efficacy perception of oral communication ability among English as a second language (ESL) technical students. *Procedia - Social and Behavioral Sciences, 204*, 98-104. https://doi.org/ 10.1016/j.sbspro.2015.08.121
Knapp, K. (2015). English as an international lingua franca and the teaching of intercultural communication. *Journal of English as a Lingua Franca, 4*(1), 173-189. https://doi.org/10.1515/jelf-2015-0003
Koosha, B., Ketabi, S., & Kassaian, Z. (2011). The effects of self-esteem, age and gender on the speaking skills of intermediate university EFL learners. *Theory and Practice in Language Studies, 1*(10), 1328-1337. http://dx.doi.org/10.17507/tpls.0601.09
Lee, K. R., & Oxford, R. (2007). Differences in strategy use, language attitudes, and self-image between Korean vocational students and traditional students. *English Language Teaching, 19*(1), 1-28. https://doi.org/10.17936/pkelt.2007.19.1.001
Lee, S., & Huang, H. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*, 307–331. https://doi.org/10.1017/S0272263108080479
Li, L., Chen, J., & Sun, L. (2015). The effects of different lengths of pre-task planning time on L2 learners' oral test performance. *TESOL Quarterly, 49*(1), 38-66. https://doi.org/10.1002/tesq.159
Lievens, F., Buyse, T., & Sackett, P. R. (2005). Retest effects in operational selection settings: Development and

test of a framework. *Personnel Psychology, 58*, 981–1007. https://doi.org/10.1111/j.1744-6570.2005.00713.x
Lipsey, M. W. (2009). Identifying interesting variables and analysis opportunities. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (2nd ed., pp. 147–158). Russell Sage Foundation.
Liu, M. (2013). English Bar as a venue to boost students' speaking self-efficacy at the tertiary level. *English Language Teaching, 6*(12), 27-37. https://doi.org/10.5539/ELT.V6N12P27
Mai, L. T. (2011). *An investigation into factors that hinder the participation of university students in English speaking lessons* [Unpublished master's thesis]. University of Languages and International Studies. repositories.vnu.edu.vn/jspui/bitstream/123456789/42495/1/04051000606.pdf
Maltz, M. (1960). *Psycho-Cybernetics: A new way to get more living out of life*. Prentice Hall Press.
Maltz, M. (1999). *Zero resistance living: The Pscycho-Cybernetics mastery series.* Prentice Hall Press.
Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing, 19*(1), 85–104. https://doi.org/10.1191/0265532202lt221oa
McKee, G., Malvern, D., & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing, 15*(3), 323-337. https://doi.org/10.1093/llc/15.3.323
Mcleod, B., Southam-Gerow, M., Jenson-Doss, A., Hogue, A., Kendall, P., & Weisz, J. (2017). Benchmarking treatment adherence and therapeutic competence in individual cognitive-behavioral treatment for youth anxiety disorders. *Journal of Clinical Child and Adolescence Psychology, 48*(1), 1-13. https://doi.org/10.1080/15374416.2017.1381914
McMullen, T. (2013). *Self-image*. In Encyclopedia of behavioral medicine (vol. 2, pp. 1744-1745). Springer.
Meihua, L. (2012). Predicting effects of personality traits, self-esteem, language class risk-taking and sociability on Chinese university EFL learners' performance in English. *Journal of Second Language Teaching and Research, 1*(1), 30-57. https://doi.org/10.5420/JSLTR.01.01.3318
Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*(2), 109-135. https://doi.org/10.1017/S0272263100004137
Nakatsuhara, F. (2008). Inter-interviewer variation in oral interview tests. *ELT Journal, 62*(3), 266-275. https://doi.org/10.1093/elt/ccm044
Norris, J. M. (2012). Meta-analysis. In C. Chapelle (Ed.), *Encyclopedia of Applied Linguistics* (pp. 3653–3662). Malden, MA: Wiley.
Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555-578. https://doi.org/10.1093/applin/amp044
Norton, B., & Toohey, K. (2001). Changing perspectives on good language learners. *TESOL Quarterly, 35*(2), 307–322. https://doi.org/10.2307/3587650
Offer, D., Ostrov, E., Howard, K., & Atkinson, R. (1988). *The teenage world*. Springer Science.
Omaggio, A. (1986). *Teaching language in context: Proficiency-oriented instruction*. Heinle & Heinle.
Ostovar, A., Safaee, E., & Sobhanifar, A. (2015). The effect of shyness on English speaking scores of Iranian EFL learners. *Journal of Literature, Languages and Linguistics, 12*, 22-28.
Park, H., & Lee, A. R. (2005). *L2 learners' anxiety, self-confidence and oral performance*. Paper presented at the Pan-Pacific Association of Applied Linguistics (PAAL), Japan.
Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. https://doi.org/10.1111/lang.12079
Proctor, B. (2002). *The winner's image*. Arizona, AZ: Life Success Productions.
Richard, J. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.
Ryum, T., Vogel, p., Walderhaug, E., & Stiles, T. (2015). The role of self-image as a predictor of psychotherapy outcome. *Scandinavian Journal of Psychology, 56*(1), 62-8. https://doi.org/10.1111/sjop.12167
Sandi-Urena, S., & Gatlin, T. (2013). Factors contributing to the development of graduate teaching assistant self-image. *Journal of Chemical Education, 90*(10), 1303-1309. https://doi.org/10.1021/ed200859e
Sato, K., (2003). *Improving our students' speaking skills: Using selective error correction and group work to reduce anxiety and encourage real communication*. Akita Prefectural.
Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/APPLIN/AMP047
Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 195–216). John Benjamins.
Sopis, J. (1965). The relationship of self-image of a reader to reading achievement. *Intervention in School and Clinic, 1*(2), 94-113. https://doi.org/10.1177/105345126500100209

Stevens, J. (2009). *Applied multivariate statistics for the social sciences*. Erlbaum.

Thai, C., & Boers, F. (2015). Repeating monologue under increasing time pressure: Effects on fluency, complexity and accuracy. *TESOL Quarterly, 50*(2), 1-25. https://doi.org/10.1002/tesq.232

Yurong, Z., & Yin, N. (2008). The effects of affective factors on oral English fluency of college English students. *Christian English Language Educators Association, 31*(2), 107-116. http://dx.doi.org/10.7575/aiac.ijalel.v.5n.3p.133

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology, 51*, 473–481. https://doi.org/10.1037/0022-0167.51.4.473

Vercellotti, M. (2015). The development of complexity, accuracy and fluency in second language performance: A longitudinal study. *Applied linguistics, 38*(1), 1-23. https://doi.org/10.1558/cj.29513

Waltz, J., Addis, M., Koerner, K., & Jacobson, N. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Counselling and Clinical Psychology, 61*(4), 620-630. https://doi.org/10.1037//0022-006x.61.4.620.

Woodrow, L. (2006). Anxiety and speaking English as a second language. *Regional Language Centre, 37*(3), 308-328. https://doi.org/10.1177/0033688206071315

Zeinivand, T., Azizifar, A., & Gowhary, H. (2015). The relationship between attitude and speaking proficiency of Iranian EFL learners: The case of Darrehshehr city. *Procedia Social and Behavioral Sciences, 199*, 240-247. https://doi.org/10.1016/j.sbspro.2015.07.512

APPENDIX A

IELTS Speaking Test 1 (Pre-test)

PART 1
The examiner asks the candidate about him/herself, his/her home, work or studies and other familiar topics.

EXAMPLE

Friends
- How often do you go out with friends? [Why, Why not?]
- Tell me about your best friend at school.
- How friendly are you with your neighbors? [Why, Why not?]
- Which is more important to you, friends or family? [Why?]

PART 2
(You will have to talk about the topic for one to two minutes. You have one minute to think about what you are going to say. You can make some notes to help you if you wish.)

Describe a writer you would like to meet.
- You should say:
- Who the writer is
- What you know about this writer already
- What you would like to find out about him/her
- And explain why you would like to meet this writer

PART 3
Discussion topics:

Reading and children
Example questions:
What kinds of book are most popular with children in your country? Why do you think that is?
Why do you think some children do not read books very often?
How do you think children can be encouraged to read more?

Reading for different purposes:
Example questions:
Are there any occasions when reading at speed is a useful skill to have? What are they?
Are there any jobs where people need to read a lot? What are they?
Do you think that reading novels is more interesting than reading factual books? Why is that?

APPENDIX A

IELTS Speaking Test 2 (Post-test)

PART 1
The examiner asks the candidate about him/herself, his/her home, work or studies and other familiar topics.

EXAMPLE

Photographs

- What type of photos do you like taking? [Why, Why not?]
- What do you do with the photos you take? [Why, Why not?]
- When you visit other places, do you take photos or buy postcards? [Why, Why not?]
- Do you like people taking photos of you? [Why, Why not?]

PART 2
(You will have to talk about the topic for one to two minutes. You have one minute to think about what you are going to say. You can make some notes to help you if you wish.)

Describe a day when you thought the weather was perfect.

You should say:

- Where you were on this day
- What the weather was like on this day
- What you did during the day
- And explain why you thought the weather was perfect during this day.

PART 3
Discussion topics:

Types of weather
Example questions:
What type of weather do people in your country dislike most? Why is that?
What jobs can be affected by different weather conditions? Why?
Are there any important festivals in your country that celebrate a season or a type of weather?

Weather forecasts
Example questions:
How important do you think it is for everyone to check what the next day's weather will be? Why?
What is the best way to get accurate information about the weather?
How easy or difficult it is to predict the weather in your country? Why is that?

APPENDIX B

Self-image Improvement Worksheet

A sample of Lesson 1

LESSON ONE

The power of imagination
Napoleon Bonaparte once wrote, "the human race is governed by its imagination." Unfortunately for many people, imagination is not good government. It is a dictatorship, an instrument of failure. By not using their imaginations positively, creatively, they are enslaved by allowing their imagination to run wild, without control, to conjure up images more or less at random. These people inevitably suffer from a self-image of failure, and are incapable of reaching their goals.

We are all well aware of the fact that behavior is controlled by the mind, the brain. The brain tells us what to do in the future based on its knowledge of the past. The brain has no future reading or fortune-telling capabilities. It is a marvelous computer, but it is limited to the past and present. The only way it can see into the future is through the imagination.

You have total control over your imagination. You can imagine any way you want to in the future, success or failure. It's all up to you. But all too often, people are not willing to exercise this control. They allow their imagination to destroy their potential by picturing situations from the past in which they have failed. We all carry mental scrapbooks, but instead of preserving the joyous occasions of life, the moments of accomplishments, some people save only their times of failure and frustration.

But it is so easy and simple to take control over your imagination. All you have to remember is that the brain has a hard time distinguishing between real and imagined experiences, if they are portrayed vividly enough.

If we picture ourselves functioning in specific situations, it is nearly the same as actual performance. Mental practice can help you perform better in real life.

OVERVIEW

In this lesson, you will begin to develop three of the major tools of Psycho-Cybernetics that you will use to change your life. They are:
1. Using your imagination creatively
2. Setting clear, specific, and realistic goals
3. Learning to relax your body and mind

Research has demonstrated that the exercises and activities in this program are most effective when done in a relaxed playful state of mind- a childlike sense of openness and curiosity in the mental attitude most conducive to learning.

Do the exercises in the order in which they appear. Try to allow yourself 30 to 45 minutes uninterrupted quiet time at the end of each day to practice. Do only as much each day as feels comfortable to you. Remember: don't strain. Your Psycho-Cybernetics program is meant to be enjoyed as well as studied.

*(For further information read **Chapter Three, Imagination- The First Key TO Your Success Mechanism** in the Psycho-Cybernetics book).*

Principles of Spiritual and Psychological Relaxation
Here are four principles of spiritual and psychological relaxation. Being able to live these principles means being able to live a creative and rewarding life, free of negative influences that hold us back. Keep these principles in mind as you progress through this course. You will notice how much easier they become as you acquire more skills in the practice of Psycho-Cybernetics.

1	**Forgive others.**	No strings attached. No sense of condemnation, no forgiveness on the installment plan. A clean slate.
2	**Forgive yourself.**	You are only human. Accept yourself with kind eyes.
3	**See yourself at your best.**	A person of confidence, not at your worst, a person of frustration.
4	**Keep up with yourself always,**	never with someone else. Stop criticizing others.

Self-inventory worksheet
An important tool in changing your self-image to one that expresses your productive and success-oriented self is to have an accurate, up-to-date list of your good qualities and negative qualities. In the special list below write the good things about yourself as well as the things that you think could be improved.

For example, under assets, you might list qualities such as kind, hardworking, punctual, etc. Write them down no matter how small you think they are.

POSETIVE QUALITIES
1 -
2 -
3 -
4 -
5 -
6 -
7 -
8 -

NEGATIVE QUALITIES
1 -
2 -
3 -
4 -
5 -
6 -
7 -
8 -

ONCE IMPOSSIBLE-NOW EASY

Think of three things you were sure at one time you wouldn't be able to learn that now are easy for you. (For example, tying your shoes, riding a bike, driving a stick-shift car, reading)
1 -
2 -
3 -
(Remember how difficult they once seemed and how easy they are now.)

Now think of at least three things that you would like to be able to do in the future that seem difficult or impossible to you now. (Be sure they are realistic. For example, learning the piano is realistic, but playing a concert at Carnegie Hall next year is not.)
1 -
2 -
3 -

This exercise is to have you realize that you have already accomplished goals in your life you once saw as difficult if not impossible. You can use the confidence of your past successes to help you reach your goals in the present and the future.

An identical world

Imagine the world where everybody is exactly the same. The same looks, age, skills, goals, everything. What would such a world look like, sound like, feel like? Write down your observations briefly. Remember there is no «right» answer.

……
……
……
………………………………………………………………………………………

What does this vision make you think about your unique individuality?

……
……
……
………………………………………………………………………………………

Once impossible-now easy

Memorize this simple phrase: ONCE IMPOSSIBLE-NOW EASY, and use it as your personal slogan, your description of your entire life experience.

Whenever you are confronted with a particularly hard or challenging situation, a new skill you need to learn that seems intimidating or difficult:

1. STOP- do NOT let anxiety and stress build up!
2. Go to the Quiet Room in your mind and…
3. Quickly remember the list of ALL the things that in your entire life were "ONCE IMPOSSIBLE" but are "NOW EASY".
4. SEE that you have lived this transition time and time and time and time again
5. Tell your Servo-Mechanism "here is another one".
6. … and what seems "impossible" would soon be «easy».

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Arzhadeeva, D., & Kudinova, N. (2020). Using Debate to develop Writing Skills for IELTS Writing Task 2 among STEM Students. *Journal of Language and Education, 6*(4), 30-43. https://doi.org/10.17323/jle.2020.10424

Using Debate to Develop Writing Skills for IELTS Writing Task 2 among STEM Students

Daria Arzhadeeva, Natalia Kudinova
National Research University Higher School of Economics

Correspondence concerning this article should be addressed to Daria Arzhadeeva, National Research University Higher School of Economics, 20 Myasnitskaya Ulitsa, 101000, Moscow, Russia.
E-mail: darzhadeeva@hse.ru

The paper focuses on the issue of developing essay writing skills in the context of IELTS preparation and explores the issue of whether academic debate can enhance STEM students' ability to structure their essays, develop a smooth progression of ideas, and provide supported and extended arguments, which, in turn, may result in higher scores for the IELTS Task Response and Coherence and Cohesion categories. To answer this, a study was undertaken in the academic years 2016/2017 and 2017/2018 among STEM undergraduate students in the National Research University Higher School of Economics, Moscow, Russia. The study involved two groups of students (36 students in each): the group that attended regular IELTS preparation classes and the other that, in addition to regular classes, attended debate classes where among other things Toulmin's argument structure was taught. At the beginning and end of the experiment both groups submitted essays that were analysed according to IELTS rubrics for Task Response and Coherence and Cohesion, and the presence or absence of the elements of Toulmin's argument structure. In addition, the essays were assessed by an independent IELTS teacher. An independent-samples t-test and Levene's test were utilised to determine the significance of the collected data. The findings revealed that, on average, the students of the experimental group scored well in Task Response and Coherence and Cohesion, yet some results were inconsistent, which requires further research.

***Keywords*:** debate, academic debate, STEM students, IELTS, Toulmin's structure, argumentation theory

Introduction

The International English Language Testing System (IELTS) is an English proficiency test that is taken by those who plan to work or study in an English-speaking environment. Comprising IELTS Academic and IELTS General Training, the test is accepted by more than 10,000 educational institutions worldwide, as the IELTS official website states.[1] Importantly, many universities in non-English speaking countries require an IELTS certificate as proof of one's command of the language (Lewthwaite, 2007; Sanonguthai, 2011).

IELTS Academic, which is the primary focus of this article, consists of four parts: Listening, Reading, Writing, and Speaking. Out of these four, a special challenge is posed by Writing. According to Test Taker Performance,[2] the mean band score of Academic Writing was the lowest compared to the other three parts across all 40 countries listed: 5.54 for males and 5.64 for females. The results in the Russia were not an exception, with the mean score being 5.97.

To meet international educational standards, the National Research University Higher School of Economics (HSE) in the Russian Federation designed the syllabus of its English course to accommodate the full format of IELTS Academic. As stated in *The Framework for Developing Foreign Languages Communicative Competence among Students of Bachelor's Programs and Specialist's Program* at the National Research University Higher

[1] IELTS. (2018a). IELTS for study. https://www.ielts.org/what-is-ielts/ielts-for-study
[2] IELTS. (2018b). Test taker performance 2018. https://www.ielts.org/teaching-and-research/test taker-performance

School of Economics,[3] by the end of the second year, students are to demonstrate English proficiency at B2-C1 CEFR level, that is, students have to gain IELTS Band 7 to get an excellent mark.

By year 2016, we had been preparing STEM (Science, Technology, Engineering, Mathematics) students for IELTS for two years and had noticed that our students tended to struggle with the Writing part of the exam, particularly with writing essays. We thought that such a situation could be explained by the fact that the language learning strategies employed by STEM students tend to differ from non-STEM students (Han, 2015). For instance, Cheng, Xu, and Ma (2007) found that STEM students use metacognitive strategies less frequently. The study by Han (2015) revealed that for writing skills STEM students "employ pre-writing strategies and attempt to find meanings of new words rather than construct messages in sentences before writing" (p. 96). On the other hand, according to González-Becerra (2019), the literature on this group is rather scarce, which means that more research into the peculiarities of language learning of STEM students is needed to tailor the courses for their needs.

The collected information motivated us to search for new ways to teach writing. Among them was debate structure (Sanongunthai, 2011). Therefore, we made the decision to carry out research to test the assumption that debate enhances the writing skills of STEM students. This paper presents the results of our study.

Theoretical Background

The Writing part of IELTS Academic comprises Task 1 and Task 2, with the Task 2 score being twice as important as that of Task 1.[4] For our research, we focused primarily on Writing Task 2, which requires examinees to write an essay of around 250 words.[5] As of August 18, 2018, on the IELTS official website it is stated that "[i]n Writing Task 2, test takers are given a topic to write about in an academic or semi-formal/neutral style. Answers should be a discursive consideration of the relevant issues".[6] What is of particular pertinence to our study is that "[t]his task assesses the ability to present a clear, relevant, well-organised argument, giving evidence or examples to support ideas and use language accurately".[7]

Four criteria are used to assess essays: Task Response, Coherence and Cohesion, Lexical Resource, and Grammatical Range and Accuracy (see IELTS Scores Guide (2018) for the complete description of the rubrics). It is important to highlight that these four criteria contribute equally to the overall score of the writing part. In this study, we look in detail into Task Response (TR) and Coherence and Cohesion (CC) as it is these criteria, we hypothesize, that can be developed by means of debate.

TR refers to formulating an opinion about a given statement and supporting it with evidence and examples taken from one's experience. To score well in TR, one should cover all of the points raised in the task question and provide well-developed and fully-supported arguments. The CC criterion involves the logical structuring of an essay and use of cohesive devices.[8] To get the highest score in CC, "there needs to be a smooth progression of sentences and ideas" (Soodmand Afshar et al., 2017, p. 7). Interestingly, Coffin (2004) pointed out that "although argument structure is not explicitly part of the IELTS assessment criteria nor a focus for rater training, the overall band descriptors used in candidates' reports make general reference to their ability to argue" (p. 233). This idea is later reiterated by Soodmand Afshar et al. (2017). Coffin (2004) goes on to recommend that English language students preparing for IELTS should be made aware of different argumentative structures.

Argumentative structures are typically studied within argumentation theory. In the framework of this theory there is still no consensus on the definition of the term 'argumentation' (Van Eemeren, 2015). For our purpose, we have chosen the definition suggested by Dupin de Saint-Cyr, Bisquert, Cayrol, and Lagasquie-Schiex:

> Argumentation can be viewed as a process done in order to exchange information together with some justification with the aim to obtain well-justified knowledge, or with the aim to increase

[3] The framework for developing foreign languages communicative competence among students of bachelor's programs and specialist's program in National Research University Higher School of Economics. (2017). National Research University Higher School of Economics.
[4] British Council. (2018). IELTS Scores Guide.
[5] Ibid.
[6] IIbid.
[7] Ibid.
[8] Ibid.

> or decrease the approval of a point of view, or to confront and combine different views. (Dupin de Saint-Cyr et al. 2016, p. 57)

In our view, this definition conveys the core of Task 2, i.e., to provide a point/points of view and support it/them with evidence. Furthermore, such an understanding of argumentation is in accord with the structure and the objectives of debates.

Out of many possible argumentation structures (Coffin, 2004; Stab & Gurevych, 2014), the most practically useful and effective for both written argumentation (Connor & Mbaye, 2002) and debates (Arzhadeeva & Kudinova, 2017) is Toulmin's model of argument structure. According to Toulmin, Rieke, & Janik (1984), the structure of any argument consists of four main and two optional elements: claim, data/grounds, warrants, backing, rebuttal, and quantifier. Claim is understood as an assertion or "a well-defined position … to consider and discuss" (Toulmin, Rieke, & Janik, 1984, p. 29). Data, or grounds, is construed as an underlying foundation that is required for a claim to be accepted as reliable. Normally, data includes statistical data, facts, the results of an experiment, etc. By warrant, Toulmin et al. (1984) mean "statements indicating how the facts on which we agree are connected to the claim or conclusion now being offered" (p. 45). However, warrants have to be validated and to that end backing is introduced. Backing is "the general body of information" (p. 26) utilized to make sure that warrants are reliable. The first optional element of the argument structure is rebuttal, which is viewed by Toulmin et al. (1984) as "the extraordinary or exceptional circumstances that might undermine the force of the supporting arguments" (p. 95). The other optional element is qualifier, which is a phrase or a word that shows the degree of certainty for the arguments used in a debate.

To illustrate, in an opinion essay the aforementioned elements of the argument structure work in the following way. Claim is the position of the writer usually expressed in an introduction. Data is a set of supporting arguments presented in the main body of an essay that include facts, examples, statistics, or expert opinions that students might know about. Warrant is normally a sentence that shows a logical connection between the main idea and each supporting argument. Backing is a fact that is used to confirm the validity of logical connections expressed in warrant. Rebuttal is a counterargument to the position of the writer frequently presented in a separate paragraph. Qualifier is a modal verb, adverb of certainty, or conditional sentence showing how certain the writer is about their position.

In their recent research, Ananda, Arsyad, and Dharmayana (2018) analysed sample essay answers with band score 8-9 from the publicly available IELTS materials from 2017[9] and proved that all these elements are used to write an argumentative essay in IELTS Writing. They identified the most commonly utilized elements of argument in Task 2: claim, data, and warrant; and went on to point out the existence of two structures in the IELTS essay: the simple structure involving claim and data and the strong structure including "claim, data as obligatory elements, warrant as conventional element, backing, qualifier and rebuttal as the optional element" (Ananda et al., 2018, p. 11). Thus, to successfully write an IELTS essay in Task 2, one seems to need to use the strong argument structure.

We looked into a number of existing teaching methods that are believed to improve students' writing performance in order to find ones that help them develop the argument structure. Ostovar-Namaghi and Safaee (2017) conducted a data-driven study, providing a list of methods that, in teachers' opinions, are of particular help to students, among which are exposing candidates to sample answers, teaching grammar and vocabulary as a prerequisite to writing, teaching prefabricated patterns, raising candidates' awareness of scoring criteria, teaching discourse markers, and encouraging learners to develop their content knowledge. It is important to note that this list is based solely on the teachers' experiences. However, there is a significant body of studies conducted mostly by Iranian researchers who "have tested the effectiveness of theory-driven interventions under controlled experimental conditions to come up with universally applicable generalizations" (Ostovar-Namaghi & Safaee, 2017, p. 74). Among those studies are ones devoted to students' exposure to model essays (Abe, 2008; Bahgeri & Zare, 2009; Dickinson, 2013); teaching meta-discourse markers (Allami & Serajfard, 2012); collaborative writing (Shehadeh, 2011); computer-assisted writing (Algraini, 2014; Bani Abdelrahman, 2013); using games (Ocriciano, 2016); corrective feedback (Ganji, 2009; Ketabi & Torabi, 2012; Sanavi & Nemati, 2014); using corpora or corpus-based materials (Smirnova, 2017); and debates (Sanongunthai, 2011).

[9] About IELTS .(2017). Academic IELTS essays sample. https://ieltsmaterial.com/writing/

From the intervention techniques listed above, we decided to focus only on debates, as during debates learners are supposed to utilize Toulmin's argument structure, which is also essential to writing a successful IELTS Writing Task 2.

In our study, we construe academic debate as

> a teaching method which is aimed at developing and enhancing a learner's aspects of personality and performance such as communicative skills, problem solving skills, active listening, critical thinking, creativity, motivation, and adaptability and helps to gain knowledge and overcome stereotypes. (Kudinova & Arzhadeeva, 2020, pp. 4–5)

The prominent feature of academic debate is a structured procedure that is agreed upon beforehand: two opposing teams discuss a controversial topic by presenting arguments in favour and against it; and their performance is assessed by a group of judges.

There are a number of studies showing that debates have a positive impact on the following aspects: students' motivation (Schroeder & Ebert, 1983); communication skills (Alasmari & Ahmed, 2013; Temjakova & Ustinova, 2012); active listening, critical thinking, and creativity (Frankovskaya, 2002; Temjakova & Ustinova, 2012); adaptability (Gaumer Erickson & Noonan, 2016; Kudinova & Arzhadeeva, 2020). In addition, debates enable students to acquire knowledge (Khopiyajnen, 2015; Vo & Morris, 2006) and challenge stereotypes (Schroeder & Ebert, 1983). However, we have not found enough research on the influence of debate on students' performance during the IELTS exam. The only scholar who directly looked into the matter is Sanonguthai (2011), who explored the effectiveness of debate in English as a teaching method used to prepare high school students for the IELTS writing module, specifically Task 2. To that end, Sanonguthai chose 20 senior students who attended an academic writing course that was taught by a specially invited debate coach. The coach tailored the syllabus to the needs of the students. She combined debate sessions with activities aimed at extending their vocabulary and improving their grammar. Sanonguthai observed one class on week 8 and conducted follow-up interviews with both the students and the coach. One month later, the students took a mock IELTS test run by an experienced IELTS teacher. Three students' essays were randomly selected and studied, and the writers of those essays were interviewed. At the end of the study, Sanonguthai drew the following conclusions:

- debate is one of the most effective methods for teaching students to write and think more critically;
- debate hones students' skills to organize their written work and to structure an argument;
- debate extends students' background knowledge of the most common issues discussed in IELTS

However, after reading the research, certain points remained unclear. Sanonguthai's analysis does not provide information on the participants' backgrounds (initial level of English, gender, previous English learning experience, the length of IELTS preparation, previous debate experience), which could be significant as they might have affected the results. The description of the procedure also leaves room for speculation. Sanonguthai does not specify how long the debate course was and how many debate classes per week the students had. The degree of the teacher's and the students' involvement in class debate activities was unequal, as the coach appears to have been more actively engaged than her students. She provided them with content, rephrased their sentences, and restructured their arguments, whereas the students seem to have assumed a passive role. Only six out of 20 gave speeches and it is not entirely clear what the rest did during the debate itself. The conclusion drawn by the scholar from the fact that a few students failed to submit one essay and that this affected their final scores in writing appears to be far-fetched. In light of all this, a new more thorough study is needed to prove or disprove the conclusions Sanonguthai made about the effect of debates on students' performance in IELTS Writing Task 2. This is what the current study intends to do.

Research Hypotheses

In our study we tested the following two hypotheses:

1. Debate improves STEM students' ability to provide well-developed and fully-supported arguments, which can help them to score well in Task Response.
2. Debate enables STEM students to structure their essays well and develop a smooth progression of ideas, which can help them to score well in Coherence and Cohesion.

Methodology

Research Design

To test our hypotheses, in 2016 we organized a debate club called the 'Debating Society' in HSE University in Moscow, Russia. The club functioned from September to March because the IELTS -type exam was scheduled at the end of March. In the 2016/2017 academic year the participants attended weekly club meetings. In the 2017/2018 academic year, we decided to continue the work of the club, but with new participants. That was done to increase the sample size.

Settings

HSE University is the leading university in Russia.[10] It offers IELTS preparation that starts in the first year, but it is quite sporadic and mostly limited to the IELTS Listening and Reading parts. Writing Task 2 comes into focus only in the second-year syllabus, which means that all students have one year of General English and start to intensively prepare for IELTS only in their second year.

Second-year students have a 90-minute English class twice a week; all in all, they have 52 classes from September to March. The classes are held in English and organised around *Complete IELTS Bands 6.5–7.5 Student's Book* by Brook-Hart and Jakeman (2013).[11] In class, students are familiarised with the IELTS exam format, requirements, and assessment criteria. Apart from reading, listening, and speaking, the students are trained to do IELTS Writing Tasks 1 and 2. The classes are taught by CELTA-trained Russian speaking teachers with 9-12 years of teaching experience.

Preparation for IELTS Writing Task 2 involves discussing different essay structures, examining sample answers, and learning cohesive devices, subject-specific vocabulary, and grammar. Every second or third week, students write an essay following the required IELTS format.

Participants

At the beginning of the 2016/2017 academic year, 138 second-year HSE University undergraduates majoring in STEM were invited to take part in the debate club. The number of students who volunteered to come to the first club meeting was 19. After having the nature and purpose of the study explained to them, only 11 students agreed to participate and regularly attended the debate meetings (at least 90% of the classes) throughout the academic year: five male and six female students. In 2017/2018, out of 152 second-year students, 42 students agreed to attend debate classes. Yet, only 25 (15 males and 10 females) of them decided to take part in the study, whereas the rest attended debate classes but did not participate in the study. We decided to view the debate participants from both academic years (36) as an experimental group due to their similarity: they were all Russian-speaking second-year STEM students who followed the same curriculum, needed to meet the same English course requirements, and planned to take the IELTS exam in March. In the experimental group, the English level of the students, which is traditionally checked at the end of the first year when all first-year students sit for the HSE independent English exam,[12] was Intermediate - Upper-Intermediate, corresponding to B1–B2 CEFR (The Common European Framework of Reference for Languages). However, there were a few students whose level of English was Pre-Intermediate (A2). Importantly, all of the participants had learnt English since they were in grade 2 of primary school. Upon completion of the school English program, only one student took a standardised test (the Russian Unified State Exam; level B2 according to the Specification of Test Measures for the Unified State Exam in Foreign Languages[13]) and scored 96 points out of 100. None of the students took any of the internationally recognized tests (B2 First, C1 Advanced, C2 Proficiency, the TOEFL® test, IELTS, etc.). As for the length of IELTS preparation in HSE, it starts in the first year, but is quite sporadic and mostly limited to the Listening and Reading parts of the IELTS exam. Writing Task 2 comes into focus only

---

[10] HSE Among Top Three Russian Universities. (2020, July 6). https://www.hse.ru/en/news/edu/376039157.html

[11] Brook-Hart, G., & Jakeman, V. (2013). Complete IELTS Bands 6.5–7.5: Student's Book. Cambridge: Cambridge University Press.

[12] *The framework for developing foreign languages communicative competence among students of bachelor's programs and specialist's program in National Research University Higher School of Economics*. (2017). National Research University Higher School of Economics.

[13] *Specification of test measures for the Unified State Exam in foreign languages.* (2018). Federal Institute for Educational Measurement (FIPI).

in the second-year syllabus, which means that all of the participants of the study had one year of General English and started to intensively prepare for IELTS only in their second year.

The participants of the control group were also the second-year Russian-speaking STEM students who attended regular English classes, but did not take part in debates. After having learnt about the purpose of the study, they volunteered to participate in it. Thus, the control group consisted of 36 students, 11 students (7 males and 4 females) from the 2016/2017 academic year and 25 students (12 males and 13 females) from the next year. The English level of the students in the control group was similar to that of the experimental group in addition to their language and IELTS preparation background.

Furthermore, we would like to underscore that out of the 36 students in the experimental group only one had a previous experience with Parliamentary debate. Those debates were held in the Russian language.

Procedure

In addition to the regular classes described in the Settings section, the students in the experimental group attended debate classes held once a week from September to March. Importantly, the teachers of the regular classes did not teach the debate classes. The latter were taught by two teachers with three years of debate teaching experience and ten years of English teaching experience.

A typical debate class entailed:

- theoretical input: the students were introduced to the rules of debates, Toulmin's structure of an argument, strong and weak arguments, propositions, rebuttal arguments, points of interest, typical fallacies, hedging, etc.;
- exercises to put theory into practice;
- mini-debates in pairs or groups of three/four (15-20 minutes long).

Every third class was a full-class debate. The topics of the debates covered issues that were common for the IELTS exam. They were normally announced one week before the debate.

Parliamentary debate was chosen as the debate format. In brief, this format involves two teams with opposite opinions debating a topic and each speaker delivering their speech within a certain time period. However, a few alterations were made to tailor the format to the needs and English level of the participants. In every debate, all of the participants were randomly assigned to three teams: Proposition, Opposition, and a team of judges. The Proposition team introduced the motion and spoke for it, providing relevant arguments, statistics, and examples. The Opposition team spoke against the motion countering the arguments of the Proposition and providing their own. The team of judges assessed the performance of the Proposition and Opposition teams and after some discussion choose the winner. Before debating, the participants had 20 minutes to develop a strategy and choose suitable arguments. During the debate, each participant in each group was to speak publicly for 3 or 4 minutes presenting their argument and/or counterargument. While someone was speaking, a member of the opposing team was allowed to ask questions (points of information). In addition, at the end of the debate each team had the chance to ask two or three questions to weaken their opponent's position. It is also important to underscore that the teacher's role during the debate class was limited to organizing, monitoring, and feedback without active involvement in the preparation, performance, or assessment.

At the beginning and the end of the course (in September and March), during regular lessons the students from both the experimental and control groups wrote an IELTS Task 2 essay as a pre- and post-test. The essay topic for both first and final classes was as follows:

> Some people claim that not enough of the waste from homes is recycled. They say that the only way to increase recycling is for governments to make it a legal requirement. To what extent do you think laws are needed to make people recycle more of their waste?[14]

For the purposes of the experiment, the topic was repeated in both academic years: 2016/2017 and 2017/2018.

The essays were collected and analysed according to the criteria, which will be discussed further. In addition, the final essays were graded by one independent teacher with 13 years of IELTS teaching experience. She used the IELTS Task 2 band descriptors for Task Response and Coherence and Cohesion. The Intraclass Correlation Coefficient for Task Response was .947 and for Coherence and Cohesion was .938. The decision to invite an independent expert was made due to the fact that IELTS examinees are officially given the overall score for both Writing Task 1 and Task 2 without specifying the particular score for an essay or scores for individual criteria, which makes it impossible to understand students' performance in Task Response and Coherence and Cohesion in Writing Task 2.

In the analysis of the essays, we relied on the fact that according to Transfer of International English Language Test Results to HSE[15] an 'excellent' mark (8 out of 10) equals IELTS band 7, so in our study we utilized the writing assessment criteria for IELTS band 7 (only Task Response and Coherence and Cohesion). The additional criterion applied to assess the essays was the use of Toulmin's model (Table 1). To simplify the assessment process, we resorted to the binary system of grading where zero (0) stood for the absence of an element and/or its improper use and one (1) stood for the proper use of an element.

Table 1
Assessment Criteria

Task Response (TR) IELTS (Band 7)		Coherence and cohesion (CC) IELTS (Band 7)		Additional Criteria: Toulmin's Model of Argument Structure (TA)	
TR1:	all parts of the task are addressed;	CC1:	information and ideas are organised logically;	TA1:	claim;
TR2:	a clear position is formulated;	CC2:	clear progression throughout;	TA2:	data;
TR3:	main ideas are presented, extended, and supported;	CC3:	a clear central topic is present within each paragraph;	TA3:	warrant;
TR4:	ideas are supported by relevant evidence;	CC4:	a range of cohesive devices is used appropriately.	TA4:	backing;
TR5:	ideas are supported by relevant examples.			TA5:	qualifier;
				TA6:	rebuttal.

To evaluate the results, we acquired the gain score for the experimental group and the gain score for the control group and did an independent-samples t-test ($p < .05$) between the two sets of gain scores. We also used Levene's test to assess the homogeneity of variances drawn from different samples (Tables 3, 5, and 7). All calculations were made using IBM® SPSS® Statistics Subscription Build 1.0.0.1347 64-bit edition.

Results

The results of the experiment are presented in Table 2 through Table 7 and grouped according to the hypotheses. Interestingly, as can be seen from the tables showing descriptive statistics (Tables 2, 4, and 6), in many cases the standard deviation is larger than the mean, which can be attributed to the fact that we used the binary system of grading with the two elements - zero and one.

Hypothesis 1. Debate Improves STEM Students' Ability to Provide Well-Developed and Fully-Supported Arguments, Which Can Help Them to Score Well in Task Response

The results of the students' performance in Task Response are presented in Tables 2 and 3. Table 2 shows the mean values for pre- and post-tests of both experimental and control groups, the mean of the gain score for each criterion in Task Response and the mean for the expert's opinion. In Table 3, the results of Levene's test and an independent-samples t-test for each TR criterion separately are given. Importantly, in all but one case the results of Levene's test exceeded 0.05. This is why, to evaluate the results of TR1, TR2, TR3, and TR5, we used the results of the t-test with equal variances assumed; however, for TR4 we utilised the one with equal variances not assumed.

[15] Transfer of international English language test results to HSE. (2018, September 7). https://www.hse.ru/en/studyspravka/perezach/

Table 2

Task Response: Mean Values of Pre- and Post-tests for Experimental and Control Groups, Mean Gain Score between Two Sets, and Mean Gain Score for Expert's Opinion

Variable	Groupsa	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
TR1	experimental group	0.417 (0.5)	0.917 (0.28)	0.5 (0.609)
	control group	0.583 (0.5)	0.833 (0.378)	0.25 (0.604)
TR2	experimental group	0.75 (0.439)	1 (0.000)	0.25 (0.439)
	control group	0.833 (0.378)	0.917 (0.28)	0.08 (0.439)
TR3	experimental group	0.306 (0.467)	0.611 (0.494)	0.31 (0.525)
	control group	0.361(0.487)	0.222 (0.422)	-0.14 (0.539)
TR4	experimental group	0.028 (0.167)	0.694 (0.467)	0.67 (0.478)
	control group	0.000 (0.000)	0.111 (0.319)	0.11 (0.319)
TR5	experimental group	0.167 (0.378)	0.778 (0.422)	0.61 (0.549)
	control group	0.194 (0.401)	0.278 (0.454)	0.08 (0.604)
TR Expert's Opinion	experimental group			6.61 (0.903)
	control group			5.31 (1.064)

Note. [a] n = 36

Table 3

Task Response: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
TR1	Equal variances assumed	0.69	0.409	1.749	70	0.085	[-0.035, 0.535]	0.412
TR2	Equal variances assumed	2.991	0.088	1.61	70	0.112	[-0.04, 0.373]	0.387
TR3	Equal variances assumed	0.175	0.677	3.368	70	0.001	[0.181, 0.708]	0.804
TR4	Equal variances assumed	25.271	0.000	5.801	70	0.000	[0.365, 0.747]	
	Equal variances not assumed			5.801	60.979	0.000	[0.364, 0.747]	1.378
TR5	Equal variances assumed	1.196	0.278	3.881	70	0.000	[0.257, 0.799]	0.918
TR Expert's Opinion	Equal variances assumed	1.341	0.251	5.612	70	0.000	[0.842, 1.77]	1.317

Note. [a] $p < .05$

A closer look at Table 3 reveals that the results for TR1 and TR2 do not demonstrate a statistically significant increase, while the results of the experimental group for TR3, TR4, and TR5 show a statistical difference when compared to the control group. Such results are supported by a 95% confidence interval (CI), which is greater than zero for all three variables, and the effect size (Cohen's *d*), which is large.

Table 3 also provides information about the results of the t-test for expert's opinion on students' performance in TR. The independent expert assessed the experimental group statistically higher than the control group, which is confirmed by the values of CI and Cohen's *d*.

Hypothesis 2. Debate Enables STEM Students to Structure Their Essays Well and Develop a Smooth Progression of Ideas, Which Can Help Them to Score Well in Coherence and Cohesion.

The results of the students' performance in Coherence and Cohesion are presented in Tables 4 and 5. Table 4 shows the mean values for the pre- and post-tests in both the experimental and control groups, the mean of the gain score for each criterion in CC, and the mean for the expert's opinion. Table 5 outlines the results of Levene's test and an independent-samples t-test for each CC criterion separately, including a 95% CI and the effect size. Similar to the results of Levene's test for TR criteria, there is one case where the results of the test did not exceed 0.05 - CC4. Consequently, for this criterion we used the results of the t-test with equal variances not assumed.

Table 4

Coherence and Cohesion: Mean Values of Pre- and Post-tests for Experimental and Control Groups, Mean Gain Score between Two Sets, and Mean Gain Score for Expert's Opinion

Variable	Groups[a]	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
CC1	experimental group	0.25 (0.439)	0.639 (0.487)	0.39 (0.645)
	control group	0.278 (0.454)	0.361 (0.487)	0.08 (0.649)
CC2	experimental group	0.167 (0.378)	0.389 (0.494)	0.22 (0.591)
	control group	0.111 (0.319)	0.278 (0.454)	0.17 (0.561)
CC3	experimental group	0.222 (0.422)	0.722 (0.454)	0.5 (0.561)
	control group	0.472 (0.506)	0.611 (0.494)	0.14 (0.639)
CC4	experimental group	0.194 (0.401)	0.806 (0.401)	0.61 (0.494)
	control group	0.25 (0.439)	0.667 (0.478)	0.42 (0.692)
CC Expert's Opinion	experimental group			6.11 (1.141)
	control group			5.17 (1)

Note. [a] n = 36

Table 5

Coherence and Cohesion: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
CC1	Equal variances assumed	1.825	0.181	2.004	70	0.049	[0.001, 0.61]	0.479
CC2	Equal variances assumed	0.49	0.486	0.409	70	0.684	[-0.215, 0.326]	0.087
CC3	Equal variances assumed	0.437	0.511	2.548	70	0.013	[0.078, 0.644]	0.599
CC4	Equal variances assumed	7.052	0.01	1.372	70	0.174	[-0.088, 0.477]	
	Equal variances not assumed			1.372	63.356	0.175	[-0.089, 0.478]	0.316
CC Expert's Opinion	Equal variances assumed	0.251	0.618	3.735	70	0.000	[0.44, 1.449]	0.876

Note. [a] p < .05

As can be seen from Table 5, the difference between the groups was significant in two (CC1 and CC3) out of the four criteria. The 95% confidence intervals of [0.001, 0.61] for CC1 and [0.078, 0.644] for CC3 confirm that. Regarding the effect size, Cohen's *d* for CC1 was small, below 0.50, and Cohen's *d* for CC3 was medium, greater than 0.50 but lower than 0.80.

Looking at the independent expert's score, we can see that the mean value of the control group is statistically lower than that of the experimental group, a 95% CI is greater than zero, and the effect size is large.

Additional Criteria: Toulmin's Model of Argument Structure (TA)

Tables 6 and 7 provide information on the students' performance in each structural element of Toulmin's model. Table 6 shows the mean values for the pre- and post-tests in the experimental and control groups and the gain score for each element of Toulmin's structure. Table 7 presents the results of the Levene's test and the t-test for each element. For TA4, the results of the Levene's test did not exceed 0.05, so we interpreted the data based on the results of the t-test with equal variances not assumed.

Table 6

Toulmin's Model of Argument: Mean Values of Pre- and Post-tests for Experimental and Control Groups, and Mean Gain Score between Two Sets

Variable	Groups[a]	M (SD) Pre-Test	M (SD) Post-Test	M (SD) Gain Score
TA1	experimental group	0.722 (0.454)	0.861 (0.351)	0.14 (0.543)
	control group	0.833 (0.378)	0.75 (0.439)	-0.08 (0.649)
TA2	experimental group	0.444 (0.504)	0.778 (0.422)	0.33 (0.632)
	control group	0.694 (0.467)	0.278 (0.454)	-0.42 (0.649)
TA3	experimental group	0.083 (0.28)	0.667 (0.478)	0.58 (0.5)
	control group	0.167 (0.378)	0.389 (0.494)	0.22 (0.637)
TA4	experimental group	0.000 (0.000)	0.556 (0.504)	0.56 (0.504)
	control group	0.000 (0.000)	0.167 (0.378)	0.17 (0.378)
TA5	experimental group	0.222 (0.422)	0.75 (0.439)	0.53 (0.736)
	control group	0.222 (0.422)	0.528 (0.506)	0.31 (0.668)
TA6	experimental group	0.222 (0.422)	0.056 (0.232)	-0.17 (0.507)
	control group	0.222 (0.422)	0.167 (0.378)	-0.06 (0.583)

Note. [a] n = 36

TA1: claim; TA2: data; TA3: warrant; TA4: backing; TA5: qualifier; TA6: rebuttal.

Table 7

Toulmin's Model of Argument: Results of Levene's Test and Independent-Samples T-Test

Variable		Levene's Test for Equality of Variances		T-Test for Equality of Means			95% CI	Cohen's d
		F	Sig.	t	df	p		
TA1	Equal variances assumed	0.59	0.445	1.576	70	0.12	[-0.059, 0.503]	0.368
TA2	Equal variances assumed	0.18	0.672	4.965	70	0.000	[0.449, 1.051]	1.171
TA3	Equal variances assumed	0.276	0.601	2.674	70	0.009	[0.092, 0.63]	0.629
TA4	Equal variances assumed	25.221	0.000	3.704	70	0.000	[0.179, 0.598]	
	Equal variances not assumed			3.704	64.911	0.000	[0.179, 0.599]	0.876
TA5	Equal variances assumed	0.395	0.532	1.341	70	0.184	[-0.108, 0.553]	0.313
TA6	Equal variances assumed	0.001	0.974	-0.863	70	0.391	[-0.368, 0.146]	0.201

Note. [a] $p < .05$

The significant difference between the experimental and control groups can be found in TA2, TA3, and TA4. The 95% CI for all three variables are greater than zero; the effect size for TA3 is medium while that for TA2 and TA4 is large.

Discussion

Hypothesis 1. Debate Improves STEM Students' Ability to Provide Well-Developed and Fully-Supported Arguments, Which Can Help Them to Score Well in Task Response

The findings show that among the students who attended the debate classes there was a far larger proportion of those who presented and extended main ideas in their essays and supported them with relevant evidence and examples (TR3, TR4, and TR5), and the difference between the groups was consistent and substantial (Table 3). Such an outcome may be explained by the fact that during debates the participants were taught to clearly formulate the core ideas and develop them by giving relevant statistics, facts, examples, or experts' opinions. These skills correspond to the elements of data (TA2) and warrant (TA3) within the framework of Toulmin's model of argument structure, which was taught during the debate classes. In the essays of the students who attended debates, data and warrant were used more frequently than in the essays of those who attended only the regular classes (Table 7).

However, participation in debates, seemingly, did not bring any changes to the students' ability to fully address all parts of the task (TR1) and to formulate a clear position in their essays (TR2) (Table 3), which is consistent with the results obtained for the elements of Toulmin's structure. There was no significant difference in the use of claim (TA1) between the two groups (Table 7). The conducted tests did not allow us to say whether all participants of the study managed to perform well or poorly in TR1, TR2, and TA1. However, we might suggest that if they had all performed well, that could be attributed to the fact that these aspects are traditionally focused on in regular IELTS preparation classes.

Although debate had a considerable impact on only three out of five criteria in Task Response, we can still see a positive trend in the way debates enabled students to perform better on their essays. Such results are also confirmed by the independent expert's grading, which showed that the students who attended the debate classes performed better in Task Response than those who only attended regular IELTS preparation classes. This allows us to say that the results of our study support the hypothesis that participation in debate considerably enhances STEM students' ability to give well-developed and fully-supported arguments, helping them to score well in Task Response.

Hypothesis 2. Debate Enables STEM Students to Structure Their Essays Well and Develop a Smooth Progression of Ideas, Which Can Help Them to Score Well in Coherence and Cohesion

Although the results of the t-test demonstrated that the students from the experimental group were better able to organise their thoughts logically (CC1), that difference was trivial (Cohen's $d = 0.479$). Such an outcome does not prove that debates helped students hone this skill. The other aspect that showed a significant improvement was CC3 (a clear central topic being within each paragraph) and, based on the medium effect size, the change was likely to be visible. Regarding CC2 (having a clear progression throughout the essay) and CC4 (using a range of cohesive devices appropriately), the debate classes did not give the students who attended them a competitive advantage over those who did not. Yet according to the independent expert, the students who attended the debate classes wrote more coherent essays and the difference between the groups was substantial.

Interpreting the results for the ability to use a range of cohesive devices appropriately, the lack of a significant difference between the control and the experimental groups can be attributed to two facts: first, cohesive devices are normally studied as a compulsory part of the IELTS preparation course at HSE University; second, the cohesive devices used in debates are more typical of spoken, not written, discourse. As to ensuring a clear progression throughout the essay, it is possible that the students of the experimental group did not differ from those of the control group due to the aspects of debate organisation. In debates, each student puts forward one logically clear and concise argument, not a string of arguments typical of essays. Constructing a clear progression within a range of arguments is a team task, as during the debates members of the team contribute equally to the defense of their main position. That aspect must have resulted in the students of the experimental group being better able to select one central topic within each paragraph. Regarding CC1, it is likely that regular debate attendees improved their ability to organise their thoughts logically; however, they might have failed to apply the acquired skill to the new circumstances.

Taking into consideration the results of the independent expert's assessment, it might be said that the overall findings seem rather inconclusive and contradictory. This might be accounted for by the fact that participation in debates did not have a considerable impact on each separate criterion within the Coherence and Cohesion criteria of IELTS Writing Task 2; however, it affected the overall impression of an essay's logical structure. We might also suggest that what contributed to the better structured essays in the experimental group was the presence of such elements as warrant (TA3) and backing (TA4) of Toulmin's strong structure (Table 7).

Based on the findings in Coherence and Cohesion, we cannot definitely state that participation in debates improves STEM students' ability to structure their essays well and develop a smooth progression of ideas that can help them to receive a higher score in Coherence and Cohesion. It follows that more research is needed into the influence of debate on the Coherence and Cohesion criteria of IELTS Writing Task 2.

Limitations of the Present Study

We would like to point out a number of limitations to the current study. The first one is the rather small sample. The small size of the sample is explained by the fact that participation in the debates as well as in the study was voluntary. We attempted to increase it by considering the results of the two academic years as one sample. The second limitation is the lack of randomisation during the assigning of participants to the control and experimental groups, which means that we cannot be sure that the students who volunteered to attend debate classes did not have any unknown significant differences from the students of the control group. For example, the students of the experimental group might have been more interested in argumentation theory or have had additional English classes outside the university. Although this limitation is significant and might have affected the results, our sample was rather homogenous in other aspects: age, level of English, educational background, IELTS preparation experience, native language, etc. Furthermore, this limitation was difficult to overcome due to the extracurricular nature of debate classes, students' schedules, and university rules.

Conclusion

In the current paper, we focused on preparing for IELTS Writing Task 2 through debate. We argued that the ability to apply Toulmin's model of argument structure may help STEM students score well in Task Response and Coherence and Cohesion. As a result of the study, the hypothesis concerning the impact of debates on Task Response was provisionally supported whereas the hypothesis regarding Coherence and Cohesion leaves room for speculation, even though the overall score for these criteria has improved.

The findings of the current study may be used in the field of ESL as one of the methods of test preparation, particularly for STEM students, for IELTS Writing Task 2. The use of debates may give STEM students deeper insight into how to structure and develop their ideas and provide extended arguments in order to score more highly in Task Response and Coherence and Cohesion. Apart from that, debate may introduce variety to the exam preparation classes and break the routines.

However, since there are a number of significant limitations to the current study, further research is needed to establish a clearer and more conclusive correlation between the use of Toulmin's model and IELTS scores for Task Response and Coherence and Cohesion. To refine our findings, we believe it would be necessary to carry out another study randomising the sample and expanding its size. Overall, these further studies might ascertain how replicable the currently collected results are and allow for more validity of the method.

Conflict of Interests

The authors declare that they have no conflict of interest.

References

Abdelrahman, O. N. M. B. (2013). The impact of using the word processor to develop EFL learners' writing skill at Al-Imam Mohammad Ibin Saud Islamic University. *IUG Journal of Humanities Research, 21*. 1–26. https://doi.org/10.12816/0013730

Abe, M. (2008). Exploring the role of model essays in the IELTS writing test: A feedback tool [Unpublished Master's thesis]. The University of Queensland.

Alasmari, A., & Ahmed, S. S. (2013). Using debate in EFL classes. *English Language Teaching, 6*(1), 147–151. https://doi.org/10.5539/elt.v6n1p147

Algraini, F. (2014). The Effect of using padlet on enhancing EFL writing performance [Unpublished Master's thesis]. Al-Imam Muhammad Ibn Saud Islamic University.

Allami, H., & Serajfard, H. (2012). Engagement markers: A technique for improving writing skills. *Journal of*

Language, Culture, and Translation, 1, 71–83.

Ananda, R. P., Arsyad, S., & Dharmayana, I. W. (2018). Argumentative features of International English language Testing System (IELTS) essays: A rhetorical analysis on successful exam essays. *International Journal of Language Education, 2*(1), 1–13. https://doi.org/10.26858/ijole.v2i1.4768

Arzhadeeva, D., & Kudinova, N. (2017). Razvitie kommunikativnih umenii kak odnoi iz sostavlyajutschih socialnoj kompetentnosti posredstvom tekhnologii 'debati' [Developing communication skills through debates]. *Philology at Moscow State Institute of International Relations (MGIMO) Journal, 2*(10), 55–64.

Bahgeri, M., & Zare, M. (2009). The role of using IELTS model essays in Improving learners' writing and their awareness of writing features. *Journal of English language Studies, 1*(1), 115–130.

Cheng, X., Xu, K., & Ma, Y. (2007). A survey of engineering student's use of English language learning strategies. *The Journal of Asia TEFL, 4*(2), 123–140.

Coffin, C. (2004). Arguing about how the world is or how the world should be: The role of argument in IELTS tests. *Journal of English for Academic Purposes, 3*, 229–246. https://doi.org/10.1016/j.jeap.2003.11.002

Connor, U., & Mbaye, A. (2002). Discourse approaches to writing assessment. *Annual Review of Applied Linguistics, 22*, 263–278. https://doi.org/10.1017/S0267190502000144

Dickinson, P. (2013). A genre-based Approach to preparing for IELTS and TOEFL essay writing tasks. *Journal of Niigata University of International and Information Studies, 16*, 1–9.

Dupin de Saint-Cyr, F., Bisquert, P., Cayrol, C., & Lagasquie-Schiex, M.-C. (2016). Argumentation update in YALLA (Yet Another Logic Language for Argumentation). *International Journal of Approximate Reasoning, 75*, 57–92. https://doi.org/10.1016/j.ijar.2016.04.003

Frankovskaya, O. (2002). *Razvitie refleksii u shkol'nikov cherez debatnuyu tekhnologiyu* [Developing reflection among schoolchildren through debate]. https://experiment.lv/rus/biblio/debate/1_frankovskaja.htm

Ganji, M. (2009). Teacher-correction, peer-correction and self- correction: Their impacts on Iranian students' IELTS essay writing performance. *The journal of Asia TEFL, 6*(1), 117–139.

Gaumer Erickson, A., & Noonan, P. (2016). Teacher guide college and career competency: adaptability. College and career competency framework. https://www.researchcollaboration.org/uploads/TeacherGuide-Adaptability.pdf

González-Becerra, I. (2019). Language learning among STEM students: Motivational profile and attitudes of undergraduates in a UK institution. *The Language Learning Journal, 47*(3), 385–401. https://doi.org/10.1080/09571736.2017.1320421

Han, T. (2015). Estrategias de aprendizaje de Lenguas Extranjeras en el contexto de la educación STEM [Foreign language learning: Strategies in the context of STEM education]. *GIST Education and Learning Research Journal, 11*, 79–102. https://doi.org/10.26817/16925777.287

Ketabi, S., & Torabi, R. (2012). Teaching academic writing in Iranian EFL classrooms: Teacher-initiated comments or peer-provided feedback? *Research in English Language Pedagogy, 1*(2), 58–65.

Khopiyajnen, O. A. (2015). Tekhnologiya debatov v obuchenii inostrannym yazykam [Debate technology in teaching foreign languages]. *Vestnik YuGU*, *1*(36), 158–160.

Kudinova, N., & Arzhadeeva, D. (2020). Effect of debate on development of adaptability in EFL university classroom. *TESOL Journal, 11*(1), 1–13. https://doi.org/10.1002/tesj.443

Lewthwaite, M. (2007). Teacher and student attitudes to IELTS Writing tasks: Positive or negative washback? *UGRU Journal, 5*, 1–23.

Ocriciano, M. (2016). IELTS Writing: A gamification journey. *Cambridge English: Research Notes, 64*, 31–36.

Ostovar-Namaghi, S. A., & Safaee, S. E. (2017). Exploring techniques of developing writing skill in IELTS preparatory courses: A data-driven study. *English Language Teaching, 10*(3), 74–81. https://doi.org/10.5539/elt.v10n3p74

Sanavi, R., & Nemati, M. (2014). The effect of six different corrective feedback strategies on Iranian English language learners' IELTS writing task 2. *SAGE Open, 4*(2), 1–9. https://doi.org/10.1177/2158244014538271

Sanonguthai, S. (2011). Teaching IELTS writing module through English debate: A case study in Thailand. *Language Testing in Asia, 1*(4), 39–99. https://doi.org/10.1186/2229-0443-1-4-39

Schroeder, H. & Ebert, D. G. (1983) Debates as a business and society teaching technique. *The Journal of Business Education*, *58* (7), 266–269. https://doi.org/10.1080/00219444.1983.10534907

Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20(*4), 286–305. https://doi.org/10.1016/j.jslw.2011.05.010

Smirnova, E. (2017). Using corpora in EFL classrooms: The case study of IELTS preparation. *RELC Journal, 48*(3), 302–310. https://doi.org/10.1177/0033688216684280

Soodmand Afshar, H., Movassagh, H., & Radi Arbabi, H. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *The Language Learning Journal, 45*(4), 1–13.

https://doi.org/10.1080/09571736.2017.1320420

Stab, C., & Gurevych, I. (2014). Identifying argumentative discourse structures in persuasive essays. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 46–56). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1006

Temjakova, T. V., & Ustinova, N. V. (2012). Tekhnologia "Parlamentsckie debaty" v vysshem obrazovanii [Parlimentary debate in tertiary education]. *Konzept*, 5, 48–52.

Toulmin, C., Rieke, R., & Janik, A. (1984). *An introduction to reasoning*. Macmillan Publishers.

Van Eemeren, F. H. (2015). Bingo! Promising developments in argumentation theory. In F. H. Van Eemeren & B. Garssen (Eds.), *Reflections on theoretical issues in argumentation theory* (pp. 3–25). Springer.

Vo, H. X., & Morris, R. L. (2006). Debate as a tool in teaching economics: Rationale, technique, and some evidence. *Journal of Education for Business, 81*(6), 315–320. https://doi.org/10.3200/JOEB.81.6.315-320

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Bin-Hady, W. R. A., Nasser, A. N. A. N., & Al-Kadi, A. T. (2020). A Pre-Experimental Study on a Process-Genre Approach for Teaching Essay Writing. *Journal of Language and Education, 6*(4), 44-54. https://doi.org/10.17323/jle.2020.10347

A Pre-Experimental Study on a Process-Genre Approach for Teaching Essay Writing

Wagdi Rashad Ali Bin-Hady[1], Aref Nassi Abdu Nasser[2], Abdu T. Al-Kadi[3,4]

[1]Hadhrmout University

[2]Abyan University

[3]Ibb University

[4]University of Science & Technology

Correspondence concerning this article should be addressed to Wagdi Bin-Hady, Hadhrmout University, PO Box 50512, Mukalla, Hadhramaut, Yemen. Email: wagdyrashad@gmail.com

This study explored the feasibility of using a *process-genre approach* (PGA) for teaching academic writing from the perspective of EFL undergraduates. The sample consisted of 15 students enrolled in a four-year English program at the College of Education in Socotra, Yemen during the academic year 2018-2019. The study followed a pre-experimental design in which a pretest was given to the sample, and an extensive 30-hours program was pursued using the PGA. Additionally, ten informants were singled out for interviews to explore their opinions about the PGA-based teaching they experienced during the experiment. A Wilcoxon Signed Rank Test was used to calculate the degree of significance in students' improvement on opinion essay writing ($Z=3.408$, $p < 0.05$) between the pretest and posttest in favor of the latter. The findings also revealed that students had positive perceptions towards the PGA that was applied by their instructors. The findings suggest that applying such an approach in writing courses could engage learners in writing practices that they view positively.

***Keywords*:** academic writing, EFL learners, opinion essay, process-genre approach, socotra

Introduction

One of the goals EFL learners aim to achieve is the mastering of academic writing to express themselves proficiently. A well-written text can be a challenging endeavor, not only for second language (L2) learners but also for native speakers. Even contributors to scholarly journals and academic gatherings such as conferences and symposiums subject their papers to peer review to get insights on their writing. This is partially owing to the complex nature of the writing process. In L2 contexts, learners exert efforts to write essays effectively, yet they encounter different problems that limit their writing abilities (Brooks, 2017; Gilmore, 2009; Hyland, 2002; Yuvaypan & Rathert, 2018). In academia, several approaches have been suggested to enhance writing skills. The Process-Genre Approach (hereafter PGA) is one of the prominent approaches in English Language Teaching (ELT). It surfaced as an advantage over the preceding approaches – product approach and the process approach. The product approach, rooted in audio-lingual teaching as an instructor-centered approach, highlights practicing syntactical and rhetorical drills. It views writing as an imitation of certain patterns without an emphasis on the process of writing (Rusinovci, 2015). Later on, concerns about how a text is written became the theme for the process approach. Again, this approach was unsatisfactory for many ELT experts, educators, and researchers (Ajmal, 2005; Arici & Kaldirim, 2015; Raimes, 1983, as cited in Zhang, 2018, Rusinovci, 2017). Another approach (the genre approach) emerged. It lays heavy emphasis on the social context in which writing is produced (Badger & White, 2000). With the passage of time, communicative language teaching paved the way for a more comprehensive approach – the process-genre approach (PGA), which is the thrust of the current study.

It is quite uncontroversial now that the PGA is advantageous over other approaches. It has been widely fostered and it is now flourishing in the field of ELT. Prior to the PGA, many writing instructors intensified their attention on the content of their students' writing, not on the organization of the academic writing process

(Furneaux, 1995). Brooks (2013) affirmed that instructors should focus not only on the sentence level, but also on the use of rhetorical and discourse skills if they wish to enhance their students' writing skills. Viewing writing as a communicative task (Badger & White, 2000), using the PGA for teaching academic writing has been evidently effective. It has been demonstrated in various contexts: the Chinese context (Xu & Li, 2018; Zhang, 2018), Colombia (Arteaga Lara, 2017), Pakistan (Ajmal, 2015), Nigeria (Saputra & Marzulina, 2015), and Indonesia (Eliwarti & Maarof, 2017; Ghufron, 2016), and the present study falls within this landscape of research. In the Arab EFL context, the PGA is largely underrepresented in the body of literature (Assaggaf, 2016). Drawing on Hyland's (2002) and other previous research on the PGA, the study at hand intends to provide more evidence for the viability of this approach from an Arab EFL context. The researchers who conducted this study have taught academic writing at the undergraduate level for years using a variety of approaches; and thus they can add their PGA-based writing experience to the investigation.

This attempt addresses the cogency of the PGA in the domain of English for academic purposes at the College of Education on the island of Socotra, Yemen. More pointedly, the present endeavor addresses the following questions:

1. Is a process-genre approach conducive to improving students' EFL academic essay writing?
2. What are the students' perceptions towards a PGA-based academic writing course?

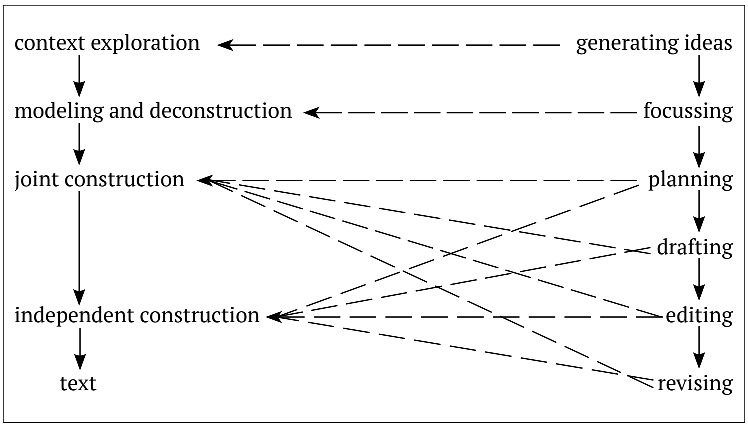
These two questions were used as a point of departure to delineate the significance of the PGA for university students' writing abilities. Besides many other factors, adopting an appropriate teaching approach such as the approach under scrutiny is vital for developing writing syllabi in the EFL program where this study was conducted.

Literature Review

The PGA has gained prominence as an approach subsumed into communicative language teaching, the widely accepted paradigm for teaching an L2. This approach, proposed by Badger and White (2000) is about writing as a communicative activity. Reporting positive perceptions of students and instructors on the PGA in academic writing, numerous studies recommended employing it in writing classes (Ajmal, 2015; Assaggaf, 2016; Eliwarti & Maarof, 2017; Ghufron, 2016). The essence of this approach according to Ghufron (2016) is that it empowers writers to experience every stage of the writing process: prewriting, drafting, and revising in order to deliver messages accurately. As its name suggests, the process-genre approach is a combination of two approaches: the process approach and the genre approach. As the concept articulates, the process-genre approach utilizes the advantages of both the process approach such as planning, editing, and publishing and the properties of the genre approach, i.e. using the language for communication purposes and according to the situation (Assaggaf, 2016; Eliwarti & Maarof, 2017; Saputra & Marzulina, 2015).

The approach in focus involves six criteria that guide students in essay writing. These criteria, according to Refnaldi (2013), include (a) the incipient knowledge to enlarge a paragraph into an essay, (b) writing an introductory paragraph with a hook, (c) background information and creating a thesis sentence, (d) developing the body paragraphs with their elements, (e) topic sentences, (f) and writing a conclusion for the essay. However, these criteria represent the basic elements for maintaining the generic structure of the text. In this respect, these criteria stand for the genre-related part of the hybrid approach, i.e. the GPA. Pertaining to the process-related part, emphasis is given to several processes and strategies starting with generating ideas and ending with revising strategies as encapsulated in Figure 1.

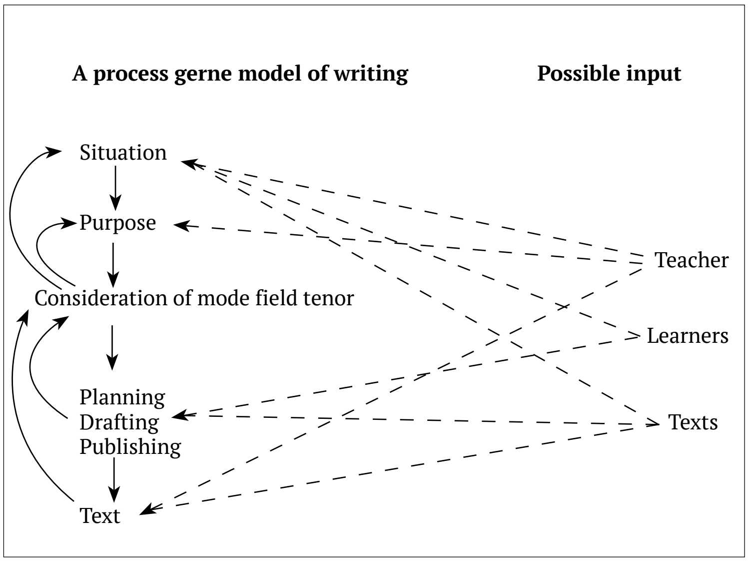
Figure 1
The PGA model adopted by Refnaldi



Note. Adopted from " The process-genre based model for teaching essay writing", by Refnaldi 2013, *Proceeding of the International Seminar on Languages and Arts,* vol. 2, pp. 483. Copyright 2013 by FBS Universitas Negeri Padang.

As can be seen in Figure 1 above, a text is constructed through different strategies and processes. These strategies are represented as generating ideas, focusing, planning, drafting, editing, and revising. Furthermore, these strategies are employed in different stages including context exploration, modeling, joint construction, and independent construction through which texts are constructed. In the process approach, “the instructor primarily facilitates the learners’ writing, and providing input or stimulus is considered to be less important. Like babies and young children who develop, rather than learn, their mother-tongue, second language learners develop, rather than consciously learn, writing skills. Instructors draw out the learners’ potential” (Badger & White, 2000, p.154). According to Badger and White, the instructor, learners, and texts are interconnected throughout the writing process. As Figure 2 illustrates, these three elements are involved from start to finish, and this overlapping/interconnectedness is what gives the approach prominence in the field of academic writing.

Figure 2
Badger & White’s Model of the PGA



Note. Adopted from "A process genre approach to teaching writing", by R. Badger and G. White, G., 2000, *ELT Journal, 45*(2), 153 (http://doi.org/10.1093/elt/54.2.153). Copyright 2000 by ELT Journal.

The PGA has been examined in different contexts. For example, Arteaga Lara (2017) examined the efficiency of the PGA in the Colombian context taking in data from thirteen learners in a bilingual school. The study showed that the PGA helped most of the participants to write well-structured paragraphs; it enabled them to develop an idea and then write on it without deviation from the subject. In the Chinese context, Zhang (2018) studied learners' academic writing using the PGA, recruiting 59 graduates in a Chinese university. A teaching experiment of 14 weeks was designed to test the effectiveness of this approach. The results showed a notable improvement in writing performance. In another research project from China, Xu and Li (2018) conducted action research to examine PGA feasibility for EAP at the doctorate level. This approach, which the authors considered a Western writing convention, helped Chinese EFL learners employ various writing strategies even in an advanced EFL program. In a similar vein, Ghufron (2016) unearthed the usefulness of the PGA in teaching writing, linking it to students' self-esteem. The results of the writing test and questionnaire on self-esteem showed that the PGA was more effective than the product approach, and students with high self-esteem, who were taught by the PGA, had better writing results. More and more results on the usefulness of the PGA are being reported. Quite recently, Alabere and Shapii (2019) explored the value of the PGA in the Malaysian ESL context. The results of pretest-posttest exhibited the outperformance of the experimental group over the control group taught with a traditional (product) method. Prior to the experiment, both groups were homogeneous and performance differences at the end of the intervention were contributable to the efficiency of the PGA. Furthermore, Assaggaf (2016) focused on the PGA in the teaching of English report writing in a computer science department in Yemen. Recruiting 17 students, the study examined the implementation of report writing in five dimensions: form, genre, planning, drafting, and revising. The findings revealed students' positive views on the teaching of academic writing using the PGA. Building on prior research on the PGA, the present study attempts to provide evidence for its value for developing EFL students' academic writing skills. It ascertains that the PGA is viable for Yemeni EFL students' achievement in composing opinion essays through (a) assessing the impact of an EFL program using the PGA on students' writing, and (b) students' perceptions towards an EFL teaching program incorporating the PGA. The study juxtaposed perceptions with experimental data to strengthen the investigation.

Methodology

Sample

The sample consisted of 15 senior students enrolled in the Department of English, Faculty of Education, Hadhramout University on the island of Socotra, during the academic year 2018-2019. The sample comprised all the fourth-year male and female students of English. They were socio-linguistically homogeneous with an average age of 22. At the beginning of the semester, the instructor asked for the students' informed consent regarding their participation in the study. They all agreed to participate in the research and ten of them were chosen randomly to be interviewed. The researcher explained to them the aims of the study, and the research process beginning from the pretest to the treatment, the posttest, and finally the interview. They were also told that their identity would remain anonymous and their information would be kept confidential. During their EFL course of study, this cohort of learners performed several language activities including writing tasks as an essential part of the program. By the end of the PGA-based teaching, ten informants were singled out for interviews to explore their opinions about their experiences.

Design

The inquiry adopted a pre-experimental research design in which a single group of participants was observed after a treatment presumed to cause change. In the experiment, the participants underwent a 30-hour PGA-based teaching program. The experiment continued for ten weeks (three hours a week). Toward the program closure, a posttest was administered to measure students' abilities in essay writing. A particular focus was on elements of content and organization.

Instruments

The study adopted two tools to collect the data from the sample. The first instrument was the writing test. The reliability of the test was calculated using the Spearman-Brown Coefficient. The items were split into two halves. The coefficient between the criteria reached .88, which is strong correlation value. At the beginning of the semester, the instructor gave the students a writing task (a pretest) to write about their opinion on living in the countryside and cities (see Appendix 1). In opinion essays, there are always two contradictory opinions for each topic presented to write about. Students should select one opinion they agreed and try to defend their opinion by providing some facts to persuade the readers. Moreover, they should also think of the counter opinion and try to minimize its importance to the readers.

The assignments were sent to another rater who had been in charge of the same academic writing course for five years. The instructor asked the rater to send his assessments on the students' pretest essays. At the end of the semester, the students were asked to write on *the influence of the internet on students* (see Appendix 2).

The second instrument was an unstructured interview to collect information on students' opinions regarding the academic writing course in general as well as writing through the PGA that the instructor used during the semester. Ten students were interviewed at the end of the PGA-based teaching program. The interviews were conducted by the course instructor. As they were in their last year of study, the students' levels of English were good enough to express their opinions, except few terms that required Arabic-English oral translation during the interview. Responses were audio-recorded. In the analysis, each interviewee was given a number for confidentiality– to avoid using their real names. Coding was used for transcribing the recordings from one to ten. All ten transcriptions were analyzed, but only four of them were quoted to represent the whole sample in the presentation of the data analysis because the other interviews reflected similar positive attitudes toward the academic writing course they experienced for ten weeks.

Classroom Procedures

The researchers followed certain steps for teaching essays to the undergraduates. The steps are similar to the model suggested by Refnaldi (2013) and include the following steps:

(a) *extensive lectures about the elements of paragraphs/essays,*
(b) *extensive lectures about genres (essay types and their aim or purpose),*
(c) *reading essays of the same genre and analyzing them,*
(d) *practicing some of the linguistic areas used in the essay, grammar, and vocabulary,*
(e) *training students on brainstorming and practicing it in the classroom,*
(f) *providing the class with an outline of a specific genre,*
(g) *giving them the elements and purpose of each part in the online - creating a real situation and motiving students to start off writing about it in a group of three students (starting by brainstorming, outlining, drafting).*

The instructor guided the students throughout the process, facilitating the writing tasks and finally providing feedback. Students usually developed their essay working in groups depending on the essay type. For example, the second chapter of the course was on descriptive essays, so after becoming familiar with the writing process, they were prompted to write a three-paragraph essay describing the classroom they were studying in by following the given steps based on format guidelines to check their essay elements and another model for peer feedback. At the end of the lecture, the instructor gave them an assignment to go home and write a three-paragraph descriptive essay on a famous tourist place on their island that tourists usually visit and to submit their writing to their peers for feedback and finally to the instructor for final corrective feedback. This included preparation, modelling and reinforcing, planning, joint constructing, independent constructing, and revising. After assessing the students' essays according to the criteria below, the data were processed to identify statistically significant differences in essay writing performance between the pretest and posttest.

Data Analysis

Adopting Brooks' (2017) benchmark, this study analyzed students' essays with relevance to (a) paragraph structure, (b) location of the main idea, (c) presence or absence of support for the main idea, and (d) internal coherence of the composition. A rubric was used to assess the students' performance on the pretest and the posttest. This rubric was adapted from Savage and Mayer (2012). A scale from 1 into 10 was used to assess the students' writing in each of the seven criteria. It focused on the following criteria:

Criterion 1 (C.1): Stating a clear position in the thesis statement

Criterion 2 (C.2): Stating a clear topic sentence and support for the body paragraphs
Criterion 3 (C.3): Using facts to support opinions
Criterion 4 (C.4): Providing a counter-argument and refutation
Criterion 5 (C.5): Providing a conclusion with convincing suggestions, a prediction, or a warning
Criterion 6 (C.6): Expressing quantifiers for avoiding generalizations
Criterion 7 (C.7): Using connectors

For each of these criteria, the students' performance was scored and the total mark was calculated along with the arithmetic mean of each student's performance. The mean values of the students' performance were calculated. The range between these arithmetic means was also calculated to assess the students' performance with regard to using the PGA in the teaching of academic writing pertaining to the elements determined in the criteria of the above-mentioned rubric. Because of the small size of the sample, the non-parametric test of Wilcoxon Signed Rank Test was used. The data results from the interviews were also analyzed qualitatively.

Results

The results are arranged in two parts. The first part corresponds to the first research question. It verifies statistically significant differences between the students' writing on the pretest and the posttest. The second part depicts data from interviews with the participants.

The pretest and posttest results demonstrate a benchmark for comparing and discussing the results. As the data in Table 1 displays, the performance of students in question improved. After scrutinizing the data, it can be seen that each student's performance, as represented by the arithmetic mean scores of the seven criteria, on the posttest was higher than their performance on the pretest. It can be seen that the general mean of the students' performance on the posttest was 7.46, which was higher than that on the pretest (4.53). Thus, the general mean of the range between the two means was 2.93. Noticeably, the largest improvement in student performance was 6.15, whereas the lowest was 0.26.

Table 1
Participants' performance in the pretest and the posttest with range

S. No.	The mean of the students' performance on the pretest	The mean of the students' performance on the posttest	Range
Student 1	8.14	8.4	0.26
Student 2	5.28	6.57	1.29
Student 3	5.85	8.14	2.29
Student 4	3.42	7.42	4
Student 5	3.57	6	2.43
Student 6	1.85	7.42	5.57
Student 7	2.42	8.57	6.15
Student 8	3	5	2
Student 9	6.71	8.57	1.86
Student 10	4.85	7.14	2.29
Student 11	7.57	8.57	1
Student 12	2.14	7.57	5.43
Student 13	3.14	8	4.86
Student 14	2.28	6	3.72
Student 15	7.71	8.57	0.86
General mean	4.53	7.46	2.93

In order to ascertain the statistical difference between the pretest and posttest results, an additional statistical calculation was used. The Wilcoxon Signed Rank Test was used to calculate the mean scores for the pretest and posttest and the results are shown in Table 2. As the data in the table shows, the mean score of the posttest was higher than that of the pretest (31.7333 vs. 52.2667). Depending on the result of the Wilcoxon Signed Rank Test, there is a statistical difference between the pretest and the posttest. This significant difference in mean values indicates that students' writing improved at Z=-3.408 and Sig.=.001.

Table 2
Wilcoxon Signed Rank Test

	N	Mean	Std. Deviation	Percentiles			Z value	Asymp. Sig (2-tailed
				25th	50th (Median)	75th		
Pretest	15	31.7333	15.46640	17.0000	25.0000	47.0000	-3.408[a]	.001
Posttest	15	52.2667	7.90539	46.0000	53.0000	60.0000		

Note.
a. Based on negative rank

Table 3 shows a comparison based on the students' performance on the pretest and the posttest across the criteria. As indicated in the table, the PGA for teaching academic writing was found to be effective for improving students' performance writing essays. The students' performance on the posttest was statistically higher, as indicated by (2-tailed Sig .018, $p < 0.05$), which is statistically significant for the students' performance on the posttest.

Table 3
Participants' performance on the pretest and posttest with the range across the criteria

The criteria	The general mean scores of students' performance in the pretest	The general mean scores of students' performance in the posttest	Range	Asymp. Sig. (2-tailed)
1. Stating a clear position in the thesis statement	6.4	7.8	1.4	
2. Stating a clear topic sentence and support for body paragraphs	5.53	8.13	2.6	
3. Using facts to support opinions	4.66	7.2	2.54	
4. Providing a counter-argument and refutation	3.2	7.6	4.4	
5. Providing a conclusion with convincing suggestions, a prediction, or a warning	4	7	3	
6. Expressing quantifiers for avoiding generalizations	3.26	7.26	4	
7. Using connectors	4.66	7.26	2.6	
Total	**4.5014**	**7.4643**		**.018**

The second part of the analysis is germane to students' perceptions towards a writing course that incorporated the PGA. After the course, the students were asked to share their perceptions. The responses were centered on two aspects: (a) students' perceptions towards academic writing in general, and (b) students' perceptions towards the instructor's teaching using the PGA. The results show that the students held a positive view towards academic writing as well as the teacher's instruction through the approach in question. The interviewees generally described the course as important and useful for developing their compositions. The extracts below clearly illustrate the interviewees' positive perceptions.

Participant 1:

> In my opinion, academic writing is an important subject because it helps us to write freely about any topic the instructor chose. It gives us to learn and use the new vocabulary.

Participant 5:

> Our instructor gives us some elements of writing such as introduction, thesis statement, and conclusion. He teaches us step-by-step, and we get benefits i*n our writing*.

Participant 8:

> The teaching practice is very enjoyable. The instructor makes it easy to understand through explaining and working in groups.

Participant 9:

> In this course, I had a chance to write freely, step-by- step. I also collaborated with my classmates in writing the essay.

These quotations exemplify that the teaching practice using the PGA was engaging and amiable, reflecting students' confidence and enthusiasm. They also show that the learners had positive perceptions towards academic writing. Further analysis of the interviews revealed that the course that incorporated the PGA helped learners fix their writing mistakes and stimulated their motivation for writing. Taken together, the responses exhibited that academic writing based on the PGA had positive effects on their performance at different levels – getting students to write freely, learning and using new vocabulary, and working collaboratively.

Discussion

The data analysis revealed the viability of using the process-genre approach for teaching academic writing. The results align with findings in the body of literature. For instance, Arteaga Lara (2017), Ajmal (2015), Eliwarti and Maarof (2017), Ghufron (2016), Luu (2011), Saputra and Marzulina (2015), Xu and Li (2018), Yuvaypan and Rathert (2018), and Zhang (2018) all hailed the PGA. In line with this body of prior research, the findings of the study at hand revealed that there was some improvement in students' writing under the auspices of the PGA. This is not surprising as previous studies came up with similar results. For instance, Ajmal (2015) confirmed that the PGA developed students' writing competence in the Pakistani context. Likewise, Gupitasar's (2013) study showed that the PGA developed students' competence writing application letters. Additionally, Alabere and Shapii (2019) reported that in an experimental study, the experimental group outperformed the control group owing to the PGA that the study centered on.

In the present study, however, the greatest improvement was in Criterion 4. The data in Table 3 demonstrates that the highest degree of development in the students' performance was in *providing a counter-argument and refutation* where the mean difference between the pretest and the posttest was 4.4. On the contrary, the lowest improvements occurred in Criteria 1, i.e. *stating a clear position in the thesis statement,* which only increased by 1.4. This was followed by Criteria 2 and 3 (range values of 2.54 and 2.6). Because the thesis statement is the core element of an essay and all the other parts of the essay are directly related to it, students need more practice to acquire this ability before they embark to the main body of an essay. Criteria 5 & 6 also improved. *Expressing quantifiers for avoiding generalizations* was the second most improved criterion with a range of 4. It can also be seen that the students' performance improved in terms of *providing a conclusion with convincing suggestions, a prediction, or a warning* with a range of 3. Overall, the students' improvement was lower for criteria 1, 2, 3, and 7.

The second part of the analysis revolved around students' perceptions toward an *academic writing* course that fostered the approach under scrutiny. The respondents favored learning through the PGA and this validates teaching writing through this approach. These findings are similar to those explored by other researchers in the field, e.g. Assaggaf (2016), Eliwarti and Maarof (2017), who reported positive perceptions from both students and instructors. Likewise, Abdullah (2008) recommended instructors use the PGA for writing instruction as it helped students become independent writers. With regard to the teaching method – defined in this paper as approaching academic writing through the PGA– the interviewees' responses showed positive perceptions towards it. Some interviewees described PGA-based teaching as enjoyable and practical. Such a positive view could be attributable to some of the techniques this approach suggests: teaching writing through discussion, grouping (putting them in groups), and modelling (giving students models to follow). The cohort of learners who undertook the course experienced the difference between the PGA-based writing and the other traditional methods that were in use before the experiment.

Overall, students' perceptions and writing improvement are clear to see. The teaching intervention utilizing the PGA enhanced the participants' writing skills and adjusted their perceptions towards the process of writing that was previously touted as a 'vexing problem' and an unwelcome task. The findings suggest that the PGA instills valuable benefits. Using the PGA boosts learners' experience in the whole process from the beginning to the end of writing tasks. Considering the findings of the present study and similar results in the literature, it is quite reasonable to suggest adhering to this approach in writing classes as it helps students experience a variety of different writing tasks.

Conclusion

The study delved into using the process-genre approach for teaching essay writing at the collegiate level. Although the study was limited to a pre-experimental design with a limited number of informants, the findings– not generalizable though– are vital to build on in further queries. They provide more evidence on the efficacy of the PGA that has been widely accepted in ESL and EFL contexts, and the context of the present study (Yemen) is not an exception. Taken together, the findings imply that learners should be trained on all of the steps of the PGA from start to finish. It is important thus to incorporate this approach when teaching writing to students studying English as a second or foreign language. The study had some limitations due to the research design. For one thing, the pre-experimental design by nature lacked a control group and thus failed to provide cause-effect results. It was also limited to opinion essay writing. Future research may explore other genres such as analytic writing, narrative writing, and the like. One more limitation is that the study attested to the significance of the approach from a trial on a small number of participants. Other researchers may include a larger sample and perhaps other research tools or may approach the effects of the PGA from a different viewpoint.

References

Abdullah, M. F. Y. (2008). A Process-genre approach in the improvement of academic writing for the undergraduates at Aden University [Unpublished doctoral dissertation]. EFL University.

Agibuay, R. M. (2017). Descriptive writing using the process-genre approach. *Asia Pacific Journal of Contemporary Education and Communication Technology, 3*(1), 272–282. https://apiar.org.au/wp-content/uploads/2017/02/24_APJCECT_Feb_BRR7128_EDU-272-282.pdf

Ajmal, A. (2015). Process-genre approach to teaching writing university foundation course students in Pakistan. In *The European Conference on Language Learning 2015* (pp. 351-356). The International Academic Forum.

Alabere R. & Shapii, A (2019). The effects of process-genre approach on academic writing. *Journal of English Education Society, 4*(2), 47-52. https://org.doi:10.21070/jees.v4i2.2598

Arteaga Lara, H. M. (2017). Using the process-genre approach to improve fourth-grade EFL learners' paragraph writing. *Latin American Journal of Content & Language Integrated Learning, 10*(2), 217–244. http://doi.org/10.5294/laclil.2017.10.2.3

Arici, A. F., & Kaldirim, A. (2015). The effect of the process-based writing approach on writing success and anxiety of pre-service teachers. *Anthropologist, 22*(2), 318–327. http://doi.org/10.1080/09720073.2015.11891883

Assaggaf, H. T. (2016). A process genre approach to teaching report writing to Arab EFL computer science students. *International Journal of English Linguistics, 6*(6), 7–18. http://doi.org/10.5539/ijel.v6n6p8

Badger, R., & White, G. (2000). A process genre approach to teaching writing. *ELT Journal, 45*(2), 153–160. http://doi.org/10.1093/elt/54.2.153

Brooks, G. (2013). Assessment and Academic writing: A look at the use of rubrics in the second language writing classroom. *Kwansei Gakuin University Humanities Review, 17,* 227–240. http://hdl.handle.net/10236/10548

Brooks, G. (2017). Teaching Academic writing in content-based classes. In P. Clements, A. Krause, & H. Brown (Eds.), *Transformation in language education* (pp. 1-5). JALT.

Eliwarti, & Maarof, N. (2017). The Effect of process-genre approach on EFL students' writing performance. *Applied Science and Technology, 1*(1), 47–52. https://www.estech.org/index.php/IJSAT/article/view/9

Furneaux, C. (1995). The challenges of teaching academic writing. *BBC English: Teachers' supplement*. EAP

Ghufron, M. A. (2016). Process-genre approach, product approach, and students' self -esteem in teaching writing. *Indonesian EFL Journal: Journal of ELT, Linguistics, and Literature, 2*(1), 37–54. http://ejournal.kopertais4.or.id/index.php/efi

Gilmore, A. (2009). Using online corpora to develop students' writing skills. *ELT Journal, 63*(4), 363-372. http://doi.org/10.1093/elt/ccn056

Gupitasar, H. (2013). The implementation of process-genre approach to teaching writing business letter. *Journal of English and Education, 1*(1), 89–95. https://ejournal.upi.edu/index.php/L-E/article/view/332

Hyland, K. (2002). *Teaching and researching writing*. Longman.

Luu, T. T. (2011). Teaching writing through genre-based approach. *BELT - Brazilian English Language Teaching Journal, 2*(1), 121–136. https://revistaseletronicas.pucrs.br/ojs/index.php/belt/article/view/9361

Refnaldi. (2013). The process-genre based model for teaching essay writing. In *Proceeding of the International Seminar on Languages and Arts* (vol. 2, pp. 479–485). FBS Universitas Negeri Padang. http://ejournal.unp.ac.id/index.php/isla/article/viewFile/4717/3687

Rusinovci, X. (2015). Teaching writing through process-genre based approach. *US-China Education Review, 5*(10), 699–705. http://doi.org/10.17265/2161-623X/2015.10.006

Saputra, H., & Marzulina, L. (2015). Teaching writing by using process genre approach to the eighth grade students of SMP Negeri 22 Palembang. *Edukasi: Jurnal Pendidikan Dan Pengajaran, 2*(1), 1–12. http://jurnal.radenfatah.ac.id/index.php/edukasi/article/view/592

Savage, A., & Mayer, P. (2012). *Effective academic writing 2: The short essay* (2nd ed). Oxford University Press.

Xu, X., & Li, X. (2018). Teaching academic writing through process-genre approach: A pedagogical exploration of an EAP program in China. *TESL-EJ, 22*(2), 1–21. http://www.tesl-ej.org/pdf/ej86/a1.pdf

Yuvaypan, F., & Rathert, S. (2018). Teaching academic writing in ELT postgraduate programmes at Turkish Universities. *International Journal of Language Academy, 6*(2), 49- 64. http://dx.doi.org/10.18033/ijla

Zhang, Y. (2018). Exploring EFL learners' self-efficacy in academic writing based on process-genre approach. *English Language Teaching, 11*(6), 115–124. http://doi.org/10.5539/elt.v11n6p115

Appendix 1

Pretest

Read the opinion in the chart below. Think of the arguments for and against this opinion and make notes in the *Agree* or *Disagree* columns (an example for each was given to you). Consider your arguments: do you agree or disagree with the opinion. Write a three-paragraph opinion essay depending on your agreement.

Agree	Opinion	Disagree
E.g. Good education	Living in big cities is better than living in the countryside.	E.g. the countryside has pure air and water

Appendix 2

Posttest

Read the opinion in the chart below. Think of the arguments for and against this opinion and make notes in the *Agree* or *Disagree* columns (an example for each was given to you). Consider your arguments: do you agree or disagree with the opinion. Write a three-paragraph opinion essay depending on your agreement.

Agree	Opinion	Disagree
E.g. Students lose their study time using the internet	The internet has a negative influence on students.	E.g. Students use the internet in their study

El-Dakhs, D. A. S. (2020). The Art of Saying "No" to University Students: A Pragmatic Analysis of the Speech Act of Refusal in Teacher-Student Role-Plays. *Journal of Language and Education, 6*(4), 55-75. https://doi.org/10.17323/jle.2020.11271

The Art of Saying "No" to University Students: A Pragmatic Analysis of the Speech Act of Refusal in Teacher-Student Role-Plays

Dina Abdel Salam El-Dakhs
Prince Sultan University

Correspondence concerning this article should be addressed to Dina Abdel Salam El-Dakhs, College of Humanities, Prince Sultan University, P.O.Box No. 66833 Rafha Street, Riyadh, 11586, Saudi Arabia.
E-mail: ddakhs@psu.edu.sa

Several studies have examined the speech act of refusal. However, there is a dearth of research on the realization of this face-threatening speech act in the context of teacher-student talk. It is also unclear how the social variables of gender and years of teaching experience could influence the realization of this speech act in teacher-student interaction. In order to address this gap in the literature, the current study investigated how university teachers decline students' requests. To this end, the realization strategies of the speech act of refusal by 60 faculty members at a private Saudi university where English is used as a medium of instruction were examined. Data were collected through role-plays and were coded regarding the types of refusal strategies, adjuncts, and modifiers employed by the teachers. The results showed a clear preference for indirect strategies, a limited use of internal ones, and a minimal influence for gender and the teaching experience on the realization strategies. The results are interpreted in light of the politeness theory, the use of English as a lingua franca, the specific context of teacher-student talk, and the existing literature.

Keywords: speech act, refusal, pragmatics, English as a lingua franca, teacher-student talk, politeness

Introduction

The current study examines the speech act of refusal, which involves the denial of a particular proposition advanced by an interlocutor and constitutes a negative response to other speech acts, including requests, offers, invitations, and suggestions (Gass & Houck, 1999). The refusal is, thus, a response to another speech act that is initiated by another participant and it is highly likely that the refusal threatens the other participant's face (Martínez-Flor & Beltrán-Palanques, 2013). In the current study, I focus on examining how university teachers decline their students' requests. Students make various requests to teachers on a daily basis. Giving consent to the requests will be often well-received by students and is, thus, a relatively easy act to perform. Instead, declining these requests is quite a complicated task since teachers need to carefully consider a number of relevant variables, including the university policies, their prior knowledge of the students, the educational values, the pedagogical goals, and most sensitively, their interpersonal relationship with the students. After all, saying "no" may not be well received by students, will most likely be questioned, and can harm the teacher-student relationship. Hence, teachers are highly likely to pay careful attention to how they decline requests, weighing a number of informational, pedagogic and interpersonal goals simultaneously.

The context of the current study is a private university in Saudi Arabia where English is the medium of instruction (EMI) and where faculty members represent a diversity of nationalities and mother tongues. In this context, it is necessary to use the English language as a medium for everyday communication. Hence, English here is used as a lingua franca (ELF), which can be understood as "any use of English among speakers of different first languages for whom English is the communicative medium of choice and often the only option," (Seidlhofer, 2011, p. 7). The use of English in this context parallels the use of ELF in real life as it has become the language of choice in a variety of international settings (Riekkinen, 2010). Due to its widespread use, ELF has become a hot topic in Applied Linguistics and English studies (Mauranen & Ranta, 2009). The traditional view of ELF as a learner language that may display deviant language use has been largely rejected (e.g., Howatt

& Widdowson, 2004; Riekkinen, 2010; Widdowson, 1994). Competent English language users who use the language to teach at universities, present at international conferences, publish in international journals, etc. cannot be considered language learners who still look to the native speaker as a model to improve their language skills. Instead, there is a general agreement that ELF is a legitimate language variety (e.g., Howatt & Widdowson, 2004; Riekkinen, 2010; Widdowson, 1994) that may have linguistic innovations (Widdowson, 1994). This revised perspective calls for further investigations of the use of ELF in a variety of contexts.

Within the framework of ELF and university teacher-student interactions, the current study is highly significant. First, examining the realization of different speech acts in teacher-student talk is under-researched in the literature. Earlier studies have mainly focused on the speech acts of praising and criticizing (e.g., El-Dakhs, Ambreen, Zaheer, & Gusarova, 2019; Hyland & Hyland, 2001; Lü, 2018) while other speech acts, including the refusal of requests, have not been examined. This is an intriguing gap in the literature since findings in this regard can provide insightful feedback for theoretical models and pedagogical practices in relation to teacher-student interactions. Second, the current study focuses on the realization of the speech act of refusal within an institutionalized academic setting. Earlier studies on the speech act of refusal have generally addressed cross-cultural comparisons (e.g., El-Dakhs, 2018; Moafian, Yazdi, & Sarani, 2019) or interlanguage pragmatics (e.g., Morkus, 2018; Shishavan & Sharifian, 2016). The current study, however, will examine how the realization of this speech act may be different in institutionalized discourse. Third, research into ELF in EMI contexts is highly needed at a time where an increasing number of universities in non-English speaking countries have decided to teach in English. In fact, in 2011, UNESCO reported that English had become the dominant medium of instruction used in higher education internationally (Tilak, 2011). Finally, the current study examines the influence of gender and the years of teaching experience on the teachers' choices of refusal strategies. Gender differences in communication are widely acknowledged in the literature (e.g., Holmes, 2001; Lakoff, 1975; Maltz & Borker, 1982), but the question is: will the effect of gender be equally found in an institutionalized setting such as the one employed in the current study? Additionally, in the education literature, the years of teaching experience is often considered an important variable of study (e.g., Annetta & Minogue, 2004; Graham, White, Cologon & Pianta, 2020; Nixon, Campbell, & Luft, 2016). The current study brings this variable into the domain of pragmatic research and aims to explore its influence on the realization of the speech act of refusal by teachers in a private university. Findings in relation to these two important social variables could be insightful.

The current study draws on the face-saving perspective of politeness by Brown and Levinson (1978, 1987) and a number of other relevant studies on the speech act of refusal and the nature of teacher-student interactions. Hence, the following section will include a brief description of Brown and Levinson's (1978, 1987) politeness theory as well as a brief survey of earlier studies on the speech act of refusal and the realization of other speech acts in teacher-student interactions. This will be followed by defining the research questions, describing the methodology and results, interpreting the results, and drawing final conclusions.

Literature Review

Brown and Levinson's Politeness Theory

At the heart of Brown and Levinson's (1978, 1987) politeness theory is Goffman's (1967) concept of 'face' which is defined as "the positive social value a person effectively claims for himself" (p. 319). Brown and Levinson (1987) consider face to be a person's public self-image "that can be lost, maintained or enhanced and must be constantly attended to in interaction," (p. 66). Due to its considerable importance in their model of politeness, Brown and Levinson (1987) further classify the concept of face into two types; namely, positive face and negative face. Positive face, on the one hand, refers to one's desire to be appreciated and approved of. People generally prefer to be appreciated and to be commended on their behaviour. Negative face, on the other hand, represents people's desire to enjoy freedom without acts of imposition. People generally prefer to have freedom in their choices, actions, and relationships.

According to Brown and Levinson (1987), "certain kinds of acts intrinsically threaten face, namely those acts that by their nature run contrary to the face wants of the addressee and/or the speaker" (p. 70). The speech act of refusal in the current study is considered a face-threatening act (FTA). When a request is declined, the requester’s positive face is threatened because the preference of the requester is not commended or approved of. The requester’s negative face also suffers because the refusal of his/her request entails taking a different course of action than the one preferred by the requester. Due to the sensitivity of FTAs, Brown and Levinson (1987) postulated that the interlocutor performing such acts often attempts to soften the effect of these acts through different redressive strategies. One politeness strategy is to avoid the FTA altogether in order to maintain harmony. In this case, priority is given to the interpersonal relationship between the interlocutors. Another politeness strategy is to express the FTA in a relatively ambiguous and indirect manner through metaphors, irony, hints, rhetorical questions, or understatements, among others. This strategy reduces the interlocutor’s commitment to the FTA and allows room for negotiation. A third strategy would be to cater to the speaker’s positive face. This strategy entails expressing appreciation and approval of the speaker’s behaviour. A fourth strategy would be to address the speaker’s negative face through mitigating the effect of imposition. Allowing the speaker some freedom of action, even if only theoretically, may help maintain face. It is also possible that the FTA is expressed directly or unambiguously for different reasons. This strategy is described by Brown and Levinson (1987) as ‘bald on record’. Certainly, the choice among these strategies will be greatly influenced by several variables, including social distance, social dominance, and cultural preferences.

It is worth noting that Brown and Levinson’s (1978; 1987) politeness theory has received some harsh criticism. According to several researchers, the model disregards the larger linguistic context (Hayashi, 1996), focuses on the self at the expense of social relationships (Spencer-Oatey, 2000), and considers human interaction to be governed by universal principles without paying adequate attention to culture specificity (Baron, 2002). Additionally, the model has been criticized for its inability to account for impoliteness as well as politeness (Eelen, 2001) and its focus on an ideal model of the person instead of accounting for the person in relation to others (Watts, 2003), which constitutes a priority in real life. Despite this criticism, and other contrary perspectives by proponents of postmodern or discursive politeness (e.g., Eelen, 2001; Mills, 2003; Watts, 2003), I adopted Brown and Levinson’s model (1978, 1987) while interpreting the results of the current study because the theory offers an incisive description of linguistic strategies (e.g., Locher, 2006; Pizziconi, 2003) and has proved easy and flexible to use while examining the realization strategies of speech acts.

Studies on the Speech Act of Refusal

Research on the speech act of refusal can be classified into three types; namely, (1) research within one language, (2) research on cross-cultural comparisons, and (3) research into interlanguage pragmatics. Relatively few studies fall under the monolingual category, but they generally aim to explore how the speech act of refusal is realized in a given language considering a number of social variables. For instance, Félix-Brasdefer (2006) investigated the linguistic strategies employed by monolingual native speakers of Mexican Spanish in one community in Mexico in refusal interactions in formal/informal situations. The findings showed that politeness in this community is realized by a set of formulaic/semi-formulaic expressions to negotiate face, and that the negotiation of face is achieved through indirect attempts to (re)negotiate a successful resolution. Interestingly, face needs in this community are oriented towards the group, emphasizing involvement over independence. Another case in point is El-Dakhs (2018). Adopting a variational pragmatic approach, El-Dakhs (2018) compared the realization of the speech acts of giving consent and the refusal of requests in two regional varieties of the Arabic language; namely, Egyptian and Saudi Arabic. The results indicated that both groups preferred the use of indirect over direct refusal strategies. Additionally, the effect of the social variables of distance and dominance was similar across the two groups.

Other studies on the speech act of refusal were cross-cultural in nature. A case in point is Kwon (2004) who compared the variations of refusal expressions between Koreans and Americans. The results showed clear cross-cultural variations in the frequency and content of semantic formulas used by each language group in relation to a number of contextual variables, including social distance and dominance, and the eliciting speech acts, which included requests, invitations, offers, and suggestions. Another example of cross-cultural studies is Moafian, Yazdi, and Sarani (2019) who examined the realization of the refusal speech act in the Persian, English, and Balouchi languages. Comparisons of the refusal realization strategies revealed statistically significant differences among the three groups of speakers concerning both the total frequencies and the frequencies of direct/indirect refusal strategies and adjuncts. Similar to Kwon (2004), social variables seemed to contribute to cross-cultural variations. It is interesting to note that such cross-cultural studies often aim at increasing global understanding and facilitating inter-cultural communication through highlighting cross-cultural variations.

In addition to monolingual and cross-cultural studies, a number of studies have examined the refusal speech act from an interlanguage perspective (e.g., Allami & Naeimi, 2011; Bella, 2011, 2014; Morkus, 2018; Shishavan & Sharifian, 2016; Wannaruk, 2008). These studies mainly examined the pragmatic competence of foreign and second language learners with the aim of identifying potential difficulties and assessing the development of pragmatic competence. There is general agreement in the literature that language learners face great difficulties developing their pragmatic competence and often experience pragmatic interference from the norms of their native language and culture. Additionally, it has often been noted that the development of pragmatic competence is influenced by a number of variables, including language proficiency, learning environment, and type of instruction. Based on the findings of these studies, it is often recommended to integrate the teaching of pragmatics into the language classroom. To this end, several studies (e.g., Da Silva, 2003; Derakhshan & Arabmofrad, 2018; Martínez-Flor & Beltrán-Palanques, 2013; Sa'd & Gholami, 2017) have explored the best practices for teaching pragmatics.

It must be notable by now that the literature has failed to recognize the special status of English when used as a lingua franca. No earlier studies attempted to examine the realization of the speech act of refusal among non-native, yet competent, English language users, who use ELF for intercultural communication. This is an intriguing gap in the literature that the current study aims to fill. This study examines the realization strategies of the refusal speech act as produced by university professors who use English competently and confidently as a lingua franca.

Studies on the Realization of Speech Acts in Teacher-Student Interactions

Research into the realization of speech acts in teacher-student interactions is relatively limited. The existing studies in this direction have mainly targeted the speech acts of criticizing and praising. For instance, Hyland and Hyland (2001) analysed the written comments of two teachers on their language learners' assignments over a complete course. It was found that praise was frequently used, particularly as a softener for criticism and suggestions. Criticism and suggestions were also mitigated with the use of several other strategies, including hedges, questions, and personal attribution. The researchers emphasized the value of mitigating criticism to redress its potential negative influence on the teacher-student relationship. However, they pointed out that mitigation and indirectness can sometimes lead to incomprehension and miscommunication by language learners. In the same vein, Lü (2018) examined the use of emails by Chinese undergraduates to communicate pedagogical criticism to their Western teachers. The results showed that the students expressed their criticism directly and sometimes even used bald critical statements in their emails. The researchers concluded that students view emails as a safe and effective channel to communicate their views. The students' preferences in this computer-mediated form of communication were clearly different from their general preferred patterns of interaction in face-to-face communication.

A few other studies were conducted on the speech act of criticizing in oral teacher-student interactions. Hiraga and Turner (1996) compared tutor-student interactions in British and Japanese academic contexts. It was noted that the Japanese students paid special attention to the face wants of the tutor while the British students tended to be sensitive towards their own face wants. While both British students and tutors tried hard to carefully attend to each other's negative face, this was not a major concern in the Japanese context, whether for the students or tutors. In a later study, Hiraga, Fuji, and Turner (2003) examined the difficulties Japanese students studying in Great Britain face regarding pragmatic understanding in tutorial sessions with British tutors. Whereas the British tutors paid special attention to students' face, students did not show similar sensitivity to their own face wants. The differences between the British and the Japanese participants in the two studies were explained in terms of cross-cultural variation. British tutors generally view university students as members of the academic community and, hence, treat them in an egalitarian manner. The relationship between tutors and students in Japanese universities is rather different since it is more hierarchical and authority-based. In the same vein, Cao (2005) highlighted the role of cultural influence on the realization of the speech act of criticism in the Chinese academic context. The Chinese lecturers' criticism was viewed as justified, and students were expected to accept the criticism and submit to the lecturers' instructions since they are assigned an inferior role in the social hierarchy in the world of academia.

Two other particularly relevant studies are Riekkinen (2010) and El-Dakhs, Ambreen, Zaheer, and Gusarova (2019). Similar to the current study, these two studies focused on the interactions of ELF-speaking university professors in an EMI context. Extending the study of criticism to the domain of doctoral thesis defenses, Riekkinen (2010) compared the use of hedges by university professors who are native speakers of English versus those who are ELF speakers. The results showed that ELF speakers used hedges differently than native speakers of English with respect to the employed expressions and their frequencies. However, these differences did not cause any miscommunication in the professors' interactions with students. El-Dakhs, Ambreen, Zaheer, and Gusarova (2019) studied the realization of the speech act of criticizing among ELF university professors and undergraduate students. It was clear that teachers preferred the use of indirect over direct strategies and used modifiers minimally, particularly internal ones. The influence of gender and years of teaching experience was minimal.

It should be clear by now that there is a lack of studies on the speech act of refusal in the context of teacher-student talk, and a dearth of studies on the effect of important social variables, such as gender and years of teaching experience, on the realization of speech acts in teacher-student interactions. The present study aims to fill this gap through examining the realization strategies of the speech act of refusal in teacher-student role-plays. The current study also considers the effect of gender as an important variable of the study due to the repeated claims of distinctive talking styles for males and females (e.g., Holmes, 2001; Lakoff, 1975; Maltz & Borker, 1982) and the effect of years of teaching experience since the author expected that the differences in years of teaching experience could influence the realization of refusal strategies. Needless to say, the findings of this study will have important educational and pedagogical implications because the face-threatening act of refusal in academic contexts can have a negative impact on the teacher-student relationship as well as the student learning experience and academic achievement. Hence, the current study aims to examine the refusal strategies employed by university teachers in response to students' requests within an ELF context. More specifically, the present study addresses the following research questions:

1. How do university teachers realize the speech act of refusal in their talk with students?
2. How does gender influence university teachers' realization of the speech act of refusal in their talk with students?
3. How do more years of teaching experience influence university teachers' realization of the speech act of refusal in their talk with students?

Methodology

Participants

A total of 60 faculty members were recruited from a private Saudi university that is characterized by its multicultural community. The sample was gender-balanced, consisting of 30 males and 30 females. The female faculty members ranged in age between 28 and 55 (mean= 41.3) and belonged to 11 nationalities. The teaching experience for the female participants varied between 5 and 30 years (Mean= 15.3), while their years of service at the private university ranged between 1 and 14 (Mean= 6.86). The male faculty members ranged in age between 30 and 58 (Mean= 41.9) and belonged to 13 nationalities. Their years of teaching experience ranged between 2 and 30 years (Mean= 14.36) while their years of service at the private university varied between 1 and 20 years (Mean= 5.56). Table (1) shows a summary of the participants' characteristics. The author did not have language proficiency scores for the participants, but the participants were competent language users who completed their university education in English, were teaching in EMI at the university level, and were communicating in English within the academic community (e.g., research papers and conference presentations, among others).

Table 1
The study participants' characteristics

Male	Age	Years of Teaching experience	Years of service at university	Nationality
1	45	20	5	Bangladeshi
2	30	3	1	Egyptian
3	45	18	1	
4	40	12	5	
5	40	15	7	
6	38	10	1	Ethiopian
7	46	20	5	Indian
8	42	15	6	
9	30	2	2	
10	40	13	11	Jordanian
11	43	14	11	
12	58	25	20	
13	39	13	12	
14	55	30	5	Malaysian
15	50	25	4	
16	43	15	3	Nigerian
17	46	20	4	
18	40	15	3	
19	45	15	9	Pakistani
20	48	20	8	
21	32	3	1	
22	40	13	1	
23	45	17	17	Palestinian
24	30	5	5	Saudi
25	35	7	3	
26	45	15	7	Tunisian
27	33	5	1	Turkish
28	58	29	5	Yemeni
29	46	15	2	
30	30	2	2	

Female	Age	Years of Teaching experience	Years of service at university	Nationality
1	45	22	5	Egyptian
2	45	20	12	
3	53	29	10	
4	58	30	5	
5	28	5	5	Indian
6	34	10	10	
7	48	22	14	
8	38	10	8	
9	38	12	5	
10	37	12	9	
11	48	20	2	Indonesian
12	32	5	5	Jordanian
13	43	17	14	
14	39	13	12	Lebanese
15	40	13	9	
16	39	12	7	
17	55	30	6	
18	35	12	2	Malaysian
19	55	25	4	
20	30	5	1	Pakistani
21	38	14	8	
22	48	20	5	
23	50	25	11	
24	50	20	10	Russian
25	30	6	3	Saudi
26	32	6	4	
27	35	8	8	
28	29	6	1	
29	40	11	6	Spanish
30	45	19	5	Sudanese

Data Collection

Data were collected through 10 role-plays (see Appendix A) representing everyday situations that university teachers often encounter with their students. These situations were prepared through a focus group with five teachers at the same university who agreed that these are regular situations they encountered with their students. The role-plays were later reviewed by three Applied Linguistics PhD holders who found them clear and appropriate to use.

The teachers' role in the role-plays was to decline the requests posed by the students including students requesting higher grades, submitting assignments late, and re-scheduling major exams, among others. In order to increase the probability of the teachers being able to naturally decline the students' requests in these role-plays, the researcher included two types of situations. The first type represented violations of university policies (scenarios 1, 3, 4, 7, 8, 9, and 10 in Appendix A), such as a request to change the student’s grade, to cancel a class without a valid reason, or to change the class time without prior coordination with the academic department. The second type of situations included requests that were somehow too difficult to accept (scenarios 2, 5, and 6 in Appendix A). Examples include a request to submit a late assignment after two extensions had already been granted and a request to re-schedule a major exam at an inconvenient time for the majority of the other students in a class without having a valid reason. The participants in the study were asked to read each scenario and then respond to the request as they would do naturally with their students. Since male teachers teach male students and female teachers teach female students in Saudi universities, the teachers imagined talking to students of the same sex. All role-plays were audio-recorded by the researcher and later transcribed manually by the researcher as well.

It is important to note that the researcher decided to collect data through role-plays, not written discourse completion tasks, because the focus of the current study is on oral communication. Role-plays increase the authenticity of teachers’ interaction. Teachers are likely to respond more naturally in role-plays than if they write their responses. The researcher is aware that collecting ethnographic data would have rendered much more authentic data. However, role-plays were more feasible since it was not easy to collect naturally occurring request-refusal sequences of teacher-student talk. Additionally, role-plays allowed for more control of the study variables and comparisons across similar situations. It is also important to note that data collection took place after receiving ethical clearance from the research ethics committee at the Saudi university. Emails were sent to faculty members requesting their participation in the study. Only the faculty members who agreed to participate completed the role-plays.

Data Coding

Data were coded using Beebe, Takahashi, and Uliss-Weltz’s (1990) model of refusal strategies. The model consists of two main categories; namely, direct refusal and indirect refusal. Each of these categories includes a number of sub-categories, such as negative willingness/inability for direct refusal, and statement of regret for indirect refusals. In addition to this classification of refusal strategies, a number of adjuncts to refusals were identified, including statements of positive opinion/empathy, pause/fillers, and expressions of gratitude/appreciation. Beebe et al.’s (1990) model is provided in Appendix B. Participants were also given the choice to opt out or to agree to the requests. However, no faculty member chose to opt out and the few cases when the faculty member chose to agree to the request (N=12 situations) were excluded from the coding. Additionally, it was acknowledged that the participants may soften their refusals by using internal modifiers. Hence, Trosborg’s (1995) model of internal modifiers, which can be found in Appendix C, was also used for data coding.

It must be noted that Beebe et al.’s (1990) model of refusal strategies was used for data coding in the current study because it was widely used in the literature (e.g., Allami & Naeimi, 2011; El-Dakhs, 2018; Bella, 2011, 2014; Kwon, 2004; Morkus, 2018) and because it was easy to use and fit well with the teachers’ responses. In fact, the only adaptation made was to introduce the sub-category “term of endearment” under “adjuncts” because it occurred repeatedly in the data. It is also important to note that another coder who is a PhD holder of Applied Linguistics was hired. After coding the data independently from the author, an inter-coder reliability check was conducted using the Intraclass Correlation Coefficient (ICC) and the outcome was 90%. Hence, the author’s coding was considered reliable and used as the basis for data analysis.

Results

How Do University Teachers Realize the Speech Act of Refusal in Their Talk with Students?

Table (2) shows how frequently the different refusal strategies were used by the university teachers. The teachers strongly preferred to use indirect (75%) rather than direct (25%) refusal strategies. This preference for indirectness can even be noted in their use of the strategies under the direct category. The teachers showed a strong preference to use the strategy of negative willingness/inability (e.g., I can’t cancel the class.), which is the least direct among “direct” strategies. In fact, this strategy was used 274 times out of the total of 332 instances of direct strategies. Regarding the indirect strategies, four main strategies dominated the scene; namely, giving excuses, reasons, and explanation (e.g., The new exam time does not suit other students.) (23%), setting a condition for future/past acceptance (e.g., If you had asked me earlier, I could have sorted it out.) (14%), expressing regret (e.g., I'm sorry.) (13%), and attempting to dissuade the student (e.g., But this move will have serious consequences.) (13%).

Table 2
Percentages of refusal strategies

Strategy	Sum	Percentage
Performative	8	0.6
Non-Performative	324	24.4%
I. Total direct strategies	332	25%
Statement of regret	162	13%
Wish	2	0.2%
Excuse, reason, explanation	302	23%
Statement of alternative	45	3.2%
Set condition for future/ past acceptance	189	14%
Promise of future acceptance	2	0.2%
Statement of principle	99	8%
Statement of philosophy	3	0.2%
Attempt to dissuade interlocutor	171	13%
Acceptance that functions as refusal	1	0.1%
Avoidance	1	0.1%
II. Total indirect strategies	977	75%

Table (3) shows the use of refusal strategies per situation (following the same order of situations in Appendix A). The participants used fewer direct strategies than indirect ones in all situations. The percentage of direct strategies across situations ranged between 18% and 34%. The situations in which the highest number of direct strategies was used involved situation 4 - a request to attend the course in another section than the one in which the student is registered (34%), situation 6 - submitting group work on an individual assignment (30%), and situation 10 - being considered present on the academic portal while not being present in class (29%). The situations in which the lowest number of direct strategies was used included situation 7 - requests to change class time (18%) and situation 2 - submitting an assignment late (22%).

Table 3
Percentage of refusal strategies per situation

Situation	Direct		Indirect	
	Sum	Percentage	Sum	Percentage
1	27	23%	92	77%
2	31	22%	111	78%
3	35	25%	103	75%
4	52	34%	99	66%
5	34	25%	104	75%
6	35	30%	80	70%
7	24	18%	107	82%
8	25	23%	83	77%
9	28	23%	96	77%
10	41	29%	102	71%

Table (4) shows the use of adjuncts and internal modifiers across all situations. Overall, the use of adjuncts and internal modifiers was small (n= 239 for adjuncts and 189 for internal modifiers) considering that each of the 60 participants responded to 10 situations. It is also notable that the participants produced more adjuncts, particularly in the form of terms of endearment (n=87) and giving advice (n= 54), than internal modifiers. The highest numbers for internal modifiers went for the use of intensifiers, mainly “very,” (n= 73), downtoners (n= 40), and cajolers (n= 36).

Table 4
Percentage of adjuncts and internal modifiers

Strategy	Sum	Percentage
Statement of positive opinion	21	5%
Statement of empathy	25	6%
Pause fillers	36	8%
Gratitude/appreciation	2	0.5%
Term of endearment	89	20.5%
Drawing attention	12	3%
Giving advice	54	13%
I. Total adjuncts	239	56%
1. Polite markers	24	6%
2. Understaters	2	0.5%
3. Hedges	2	0.5%
4. Downtoners	40	9%
5. Cajolers	36	8%
6. Subjectivizers	5	1%
Total downgraders	109	25%
Intensifiers	73	17%
Plus commitors	7	2%
Total upgraders	80	19%
II. Total internal modifiers	189	44%

Table (5) shows the use of adjuncts and internal modifiers for every situation (following the same order of situations in Appendix A). The participants’ tendency to use more adjuncts than internal modifiers is noted here again with a higher number of adjuncts in 7 out of the 10 situations. Interestingly, this table also shows that the adjuncts and internal modifiers were used minimally since they ranged between 82 and 28 with a mean of 42.8 per situation.

Table 5
Table 3 Percentage of adjuncts and internal modifiers per situation

Situation	Adjuncts		Internal Modifiers	
	Sum	Percentage	Sum	Percentage
1	53	65%	29	35%
2	30	63%	18	37%
3	26	55%	21	45%
4	21	62%	13	38%
5	15	38%	25	62%
6	18	58%	13	42%
7	30	56%	24	44%
8	19	61%	12	39%
9	15	45%	18	55%
10	12	43%	16	57%

How does gender influence university teachers' realization of the speech act of refusal in their talk with students?

In order to examine the influence of gender on the teachers' refusal strategies, a t-test was run to compare the averages of the two genders. As shown in Table 6, the comparisons reached statistical significance only in the case of one direct strategy and three indirect strategies. While males used significantly more performatives (e.g., *I refuse*.) (P = 0.0003) and letting the interlocutor off the hook (e.g., *Don't worry about it*.) (P= 0.003), females produced significantly more statements of regret (e.g., *I'm sorry*.) (P= 0.001) and statements of principle (e.g., *I never mark absent students as present. This is unethical*.) (P= 0.006). As for adjuncts, females used significantly more terms of endearment (P=0.001), expressions of empathy (e.g., *I understand how you feel*.) (P= 0.025), and expressions of positive opinion (e.g., *That's a good idea*.) (P= 0.015). The results followed a similar pattern with internal modifiers as females produced significantly more subjectivizers (e.g., *I think*) (P= 0.025), intensifiers (e.g., *very*) (P= 0.002), and plus commitors (e.g., *of course*) (P= 0.008). Hence, comparisons reached statistical significance in 4 out of 22 comparisons for direct and indirect strategies, 3 out of 6 for adjuncts, and 3 out of 9 for internal modifiers.

Table 6
T-test results – influence of gender

Strategy	Gender	Mean	SD	T	Sig (2-tailed)
Performative	Female	0.003	0.06	2.140	0.0003
	Male	0.02	0.15		
Total direct strategies	Female	0.58	0.58	1.097	0.273
	Male	0.53	0.61		
Statement of regret	Female	0.35	0.48	4.416	0.001
	Male	0.19	0.39		
Statement of principle	Female	0.21	0.41	2.763	0.006
	Male	0.12	0.33		
Interlocutor off the hook	Female	0.00	0.00	3.041	0.003
	Male	0.03	0.17		
Total indirect strategies	Female	1.78	0.87	4.416	0.001
	Male	1.48	0.77		
Positive opinion	Female	0.05	0.23	2.452	0.015
	Male	0.02	0.13		
Empathy	Female	0.06	0.24	2.253	0.025
	Male	0.02	0.15		
Term of endearment	Female	0.24	0.43	6.594	0.001
	Male	0.05	0.23		
Total adjuncts	Female	0.49	0.71	3.568	0.001
	Male	0.30	0.59		
Subjectivizers	Female	0.02	0.13	2.251	0.025
	Male	0.00	0.00		
Total internal modifiers – Downgraders	Female	0.21	0.47	1.820	0.069
	Male	0.15	0.38		
Intensifiers	Female	0.16	0.37	3.143	0.002
	Male	0.08	0.27		
Plus commitors	Female	0.02	0.15	2.673	0.008
	Male	0.00	0.00		
Total internal modifiers – Upgraders	Female	0.19	0.40	3.829	0.001
	Male	0.08	0.27		

For further examination of the influence of gender on the university teachers’ refusal strategies, a T-tests was run for each situation as well. As Table (7) shows, the majority of situations had one or two statistically significant comparisons. The most notable was the terms of endearment, which were significantly used more often by females in 7 out of the 10 situations. The other comparisons were noted in fewer situations. For example, the request was criticized more often by females in situations 4 (P= 0.043) and 7 (P= 0.047), and by males in situation 2 (0.001). Males produced significantly more conditions in situations 2 (P= 0.047) and 9 (P= 0.009), and more fillers in situations 2 (P= 0.047) and 3 (P= 0.024) while females used significantly more statements of principle in situations 8 (P= 0.023) and 9 (P= 0.012). Finally, females produced significantly more statements of regret in situation 9 (P= 0.006) and positive opinion in situation 3 (P= 0.012), whereas males let the interlocutor off the hook significantly more frequently than females in situation 2 (P= 0.047).

Table 7

T-test results – influence of gender per situation

Situation	Strategy	Gender	Mean	SD	T	Sig.
1	Terms of endearment	Female	0.43	0.50	2.688	0.010
		Male	0.13	0.35		
2	Excuse, reason, & explanation	Female	0.67	0.48	2.112	0.039
		Male	0.40	0.50		
	Setting condition	Female	0.03	0.18	2.047	0.047
		Male	0.20	0.41		
	Criticizing the request	Female	0.00	0.00	3.808	0.001
		Male	0.33	0.48		
	Interlocutor off the hook	Female	0.00	0.00	2.112	0.043
		Male	0.13	0.35		
	Fillers	Female	0.03	0.18	2.047	0.047
		Male	0.20	0.41		
3	Positive opinion	Female	0.20	0.41	2.693	0.012
		Male	0.00	0.00		
	Fillers	Female	0.03	0.18	2.344	0.024
		Male	2.23	0.43		
4	Criticize the request	Female	0.13	0.35	2.112	0.043
		Male	0.00	0.00		
	Terms of endearment	Female	0.27	0.45	2.121	0.039
		Male	0.07	0.25		
5	Terms of endearment	Female	0.23	0.43	2.971	0.006
		Male	0.00	0.00		
6	Terms of endearment	Female	0.23	0.43	2.344	0.024
		Male	0.03	0.18		
7	Criticize the request	Female	0.20	0.41	2.047	0.047
		Male	0.03	0.18		
	Terms of endearment	Female	0.27	0.45	2.121	0.039
		Male	0.07	0.25		
8	Statement of principle	Female	0.17	0.38	2.408	0.023
		Male	0.00	0.00		
	Terms of endearment	Female	0.20	0.41	2.047	0.047
		Male	0.03	0.18		

Situation	Strategy	Gender	Mean	SD	T	Sig.
9	Statement of regret	Female	0.50	0.51	2.878	0.006
		Male	0.17	0.38		
	Setting condition	Female	0.30	0.47	2.699	0.009
		Male	0.63	0.49		
	Statement of principle	Female	0.20	0.41	2.693	0.012
		Male	0.00	0.00		
10	Terms of endearment	Female	0.20	0.41	2.693	0.012
		Male	0.00	0.00		

How Do More Years of Teaching Experience Influence University Teachers' Realization of the Speech Act of Refusal in Their Talk with Students?

In order to address the influence of more years of teaching experience on the realization of refusal among university teachers, a Pearson correlation test was run to measure the statistical relationship, or association, between the two variables. The effect of the years of teaching experience on the realization of the speech act of refusal was minimal. Overall, only one case of positive correlation reached significance. The more years of teaching experience the teachers had, the more likely they were to make use of performatives (e.g., *I refuse*.) (R= 0.116; P= .004). As for the different situations, Table 8 shows that every situation (following the same order of situations in Appendix A) had only one case of statistically significant correlation except for situation 2 (the student requests to submit an assignment late), which has three statistically significant correlations. These numbers are truly minimal since the number of potential refusal strategies, whether direct or indirect, adjuncts, and internal modifiers available to the participants was 39.

Table 8

Pearson correlation test results – influence of the years of teaching experience per situation

Situation	Strategy/Modifier	R	Sig
1	Negative willingness	-0.264	0.042
2	No	0.288	0.026
	Threat/ Consequences	-0.270	0.037
	Criticize the request	-0.264	0.042
3	Not applicable		
4	Intensifier	0.269	0.037
5	Downtoner	0.268	0.038
6	Performative	0.302	0.019
7	Not applicable		
8	Statement of principle	0.307	0.017
9	Not applicable		
10	Giving advice	0.279	0.031

Discussion

The results of the present study show that university teachers prefer to express refusal to students' requests indirectly. This general tendency to use indirect refusals, probably to mitigate the threatening effect of the speech act, has been reported in earlier studies (e.g., El-Dakhs, 2018; Félix-Brasdefer, 2006). This tendency seems to be further enhanced in the current study as these university teachers preferred to use indirect strategies in three-quarters of the situations. This highly frequent use of indirect strategies can be explained in terms of the use of ELF. The users of ELF are generally characterized as being particularly sensitive and cooperative because of their lack of familiarity with their interlocutors' cultures (Mauranen, 2003). When someone is unfamiliar with the addressee’s culture, they tend to be more sensitive and cooperative in order to avoid any misunderstandings or conflicts.

Another contributing factor to the highly frequent use of indirect strategies is the sensitivity of the teacher-student relationship. Teachers often carefully consider their interpersonal relationship with students and give priority to rapport management while addressing FTAs (e.g., El-Dakhs, Ambreen, Zaheer & Gusarova, 2019; Hyland & Hyland, 2001). This is clearly illustrated in the university teachers’ preferences for particular refusal strategies. Under the direct category, the teachers mostly expressed their inability to give consent. They mainly preferred to say “*I can’t do*…” and then explained the relevant university policy or educational principle (e.g., *The university policy does not allow me to cancel classes*). Hence, any arising conflict will not be personal, but will be directed towards the set policies and guidelines. A similar pattern is noted in the use of indirect strategies. The teachers mostly used expressions of regret (e.g., *I’m sorry*), explained the university policy that justified their refusals (e.g., *The lab has a limited number of seats, so we need a limited number of students*), or attempted to dissuade the students by mentioning the consequences of their actions as set by the institution (e.g., *I will have to mark you as absent*).

In terms of politeness, the teachers’ preferences reflect an obvious sensitivity to the students’ face. Their limited use of direct strategies reflects their preference to avoid using bald on record strategies that are highly threatening to students’ face. Instead, they preferred to cater to students’ negative face by using strategies that involve minimal imposition. For instance, they preferred to dissuade students by explaining the consequences of their requests, which may theoretically allow students room to make their own decisions. For example, the teachers explained to students that if they preferred to attend an event during class time, they would be considered absent. It was then the student’s decision to attend class or join the event. Another example is the teachers’ preference for giving reasons and explanations for their refusal. This strategy allows room for discussion and negotiation, even if mainly theoretically. Interestingly, the teachers’ explanations were mainly related to institutional policies, such as the teacher’s inability to change class time without prior approval from the university management, the university stipulation that advisees need to meet with their advisors before being allowed to register for courses, the university policy regarding the maximum number of students allowed in a class, etc. Some teachers even made it clear that they could agree to the students’ requests if the official in charge (e.g., department chair, registrar, etc.) would grant the student permission. Reference to the university policies helps mitigate the imposition of refusals since the primary cause is institutional and applies to all students. Because the imposition of refusals stems from the institutions, not the teachers, the teacher-student interpersonal relationship will not be harmed.

The results of the current study also show that the influence of gender and teaching experience on the teachers’ responses was minimal. In the case of gender, only four comparisons were statistically significant for the direct and indirect strategies of refusal. Males used performatives (e.g., *I refuse*) and let students off the hook (e.g., *Don’t worry about it*) more frequently than females while females produced significantly more expressions of regret (e.g., *I’m sorry*) and statements of principle (e.g., *I never mark absent students as present in the system*). The influence of teaching experience was even smaller with only one statistically significant comparison. Notably, the more teaching experience the teachers had, the more likely they were to produce performatives. The minimal influence of these important social variables on the teachers’ use of refusal strategies can be interpreted in relation to the context of the study. Teacher-student talk in an institutionalized academic setting reflects the values underpinning the educational culture of the discourse (Hiraga & Turner, 1996) and typifies the participants’ actions when engaging in interactions (Araújo, 2012). Teachers’ choices in this particular discourse are governed by similar educational values and workplace guidelines. The similar values and guidelines minimize variation in the teacher talk and, thus, reduce the influence of social variables which may exert stronger influence in other contexts. In other words, whether the teacher is male/female or has few/many years of teaching experience, the policies of the institutions, the teaching guidelines disseminated across academic departments, and the regular teacher training workshops as part of the institutional professional development programs will reduce discrepancies in teachers’ responses to students’ requests and increase the similarity of their preferred strategies.

The university teachers’ use of modifiers in the current study is particularly revealing. One important point in this regard is that the teachers employed an extremely small number of modifiers, whether in the form of downgraders or upgraders. This is in line with Cogo and Dewey’s (2006) argument that ELF speakers are content-oriented and, thus, their language often lacks interactional features such as hedges. The small number of modifiers can also be explained in terms of the fact that students in this study are language learners. As noted in Hyland (2000) and Hyland and Hyland (2001), modifiers may not be easily noticed by language learners and their use may lead to miscomprehension. Aware of their students' relatively limited language competence, university teachers may have preferred to avoid the use of modifiers, particularly internal ones, in order to enhance the clarity of the message.

Another important observation with regard to the use of modifiers is that all the statistically significant comparisons with modifiers were in favour of females. Females produced more terms of endearment, expressions of empathy (e.g., *I understand how you feel*), statements of positive opinion (e.g., *This is a good idea*), subjectivizers (e.g., *I think*), intensifiers (e.g., *very*), and plus commitors (e.g., *of course*). This result comes in line with the frequently reported distinctions between males' and females' talk (e.g., Holmes, 2001; Lakoff, 1975; Maltz & Borker, 1982). Of particular relevance here is Tannen's (1986, 1990, 1994) emphasis that men's talk is different from women's talk. She wrote that the language of females is primarily 'rapport-talk,' which emphasizes establishing connections and promoting solidarity, whereas the language of males is 'report-talk,' which priorities conveying information and exhibiting skill. In the current study, females produced more modifiers that help maintain rapport and mitigate conflicts, such as terms of endearment, expressions of empathy, statements of positive opinion, and subjectivizers. Their use of intensifers also created the same effect since the female teachers in her studies usually added "*very*" to the expression of regret "*I'm sorry*."

Two final observations are worth mentioning. First, the teachers in the current study managed to carefully manipulate the different refusal strategies to convey their messages clearly while effectively managing their rapport with students. The fact that the users of ELF in the present study manipulated the different refusal strategies adequately across different situations supports the call to treat ELF as a legitimate variety of English in its own right (Howatt & Widdowson, 2004; Riekkinen, 2010; Widdowson, 1994). Language learners who use English as a second or foreign language often fail to employ different refusal strategies adequately and reflect poor pragmatic competence (e.g., Allami & Naeimi, 2011; Bella, 2011, 2014). Hence, users of ELF behave differently than other language learners and constitute a legitimate language variety. Second, the teachers' preferences of refusal strategies support the idea that the selection of appropriate strategies is situation-dependent (Brown & Levinson, 1987; Holtgraves, 1992). For example, the teachers used more direct strategies than usual when students asked to attend class in another section than their own, to submit group work on an individual assignment, and to be considered present despite missing class. The severity of these three situations may explain why the teachers resorted to the use of more direct strategies. Attending class in another section not only violates university policy but entails major practical difficulties. It would be difficult to track the students' attendance, record their grades, etc. Similarly, submitting the only individual assignment in the form of group work can seriously impact the course assessment and the achievement of the learning outcomes. Likewise, considering an absent student present is a clear ethical violation in addition to being a breach of institutional policy. Hence, the severity of these situations appears to be an important determinant of refusal strategies.

Conclusion

It is an everyday activity for university teachers to receive requests from their students. Consenting to these requests is a much easier speech act than declining since refusal will threaten a student's face and may, thus, harm the teacher-student relationship. The students are also likely to question the teacher's refusal and engage in follow-up discussions and negotiations. Hence, teachers need to carefully consider the basis of the refusal in terms of educational principles, pedagogic goals, and institutional policies, among others. Despite the fact that declining students' requests is a common activity, and a daunting one at that, the literature lacks relevant studies on the realization of the speech act of refusal in teacher-student talk. It is this gap that the current study aimed to fill through examining how faculty members in a private Saudi university characterized by its use of EMI and its multi-cultural community declined students' requests. The current study also addressed another important gap in the literature, which is the lack of studies on the speech act of refusal among ELF users.

The results showed the teachers' strong preference for the use of indirect rather than direct refusal strategies. This reflects the teachers' sensitivity to the possibility of students losing face by minimizing imposition (Brown & Levinson, 1978, 1987), the teacher's careful attention to their interpersonal relationships with students, and the sensitive and cooperative nature of ELF users (Mauranen, 2003). The results also showed a limited use of modifiers. This comes in line with Cogo and Dewey's (2006) description of ELF as a content-oriented language variety that lacks interactional features. Additionally, it was found that the influence of gender and the years of teaching experience on the teachers' preferred refusal strategies was minimal. This was explained with reference to the institutionalized context of the current study. Talk in institutionalized settings often typifies the actions of the participants who appropriate their choices to the institutional conventions and norms (Araújo, 2012). It should be noted here that some gender influence was noted in the teachers' use of modifiers as females tended to use more modifiers that could support their rapport with students (Tannen, 1990).

Based on these results, two recommendations are proposed. First, it is strongly recommended to consider ELF as a legitimate variety of English that is worthy of investigation on its own. It is not wise, nor feasible, to disregard the widespread use of English as a means of intercultural communication. After all, the number of competent ELF users has surpassed the number of native English speakers in today's world (Statista, 2020). In terms of pragmatic research, empirical descriptions of ELF pragmatics are underrepresented (Seidlhofer, 2011; Taguchi and Ishihara, 2018). Therefore, it is important to conduct more studies on the pragmatic choices and preferences of ELF users. Second, teacher-student talk is an under-researched area in pragmatic research. It would be advisable to conduct more studies on teachers' and students' pragmatic decisions. In these future studies, it is highly recommended to explore other extremely important variables that the current study did not address, such as the effect of the private versus public university context and the influence of a number of individual variables including the teachers' years of service in a country, how long they had been teaching in ELF, and their cultural/linguistic backgrounds. The findings of this research will not only inform pragmatic models and principles, but will also prove useful in teacher-training programs as teachers can be trained on the best practices to ensure the appropriateness of their language while maintaining the clarity of their messages.

Finally, it is important to acknowledge three limitations in the current study. First, the study data represent teachers' responses within role-plays, not within natural interactions with students. Collecting data through ethnographic methods in future studies will add more authenticity to the data and more credibility to the results. Second, the study instrument directed the participants to decline their students' requests to some extent since it required them to explain their reasons why they would have given consent if they were to choose that option. This may have influenced the responses of some teaches. In future studies, it is recommended to allow teachers to choose whether to give consent to or decline requests in order to make the teachers' responses more natural. Finally, the author did not obtain the language proficiency scores of the participants. Although they were competent language users due to their EMI experience as students, teachers, and members of the academic discourse community, it is recommended to obtain the language proficiency scores of participants in the future to support the interpretation of the results.

Acknowledgment

This work was supported by Prince Sultan University under Grant [Applied Linguistics Research Lab- RL-CH-2019/9/1].

References

Allami, H. & Naeimi, A. (2011). A cross-linguistic study of refusals: An analysis of pragmatic competence development in Iranian EFL learners. *Journal of Pragmatics, 43*, 385-406. https://doi.org/10.1016/j.pragma.2010.07.010

Annetta, L. A. & Minogue, J. (2004). The effect of teaching experience has on perceived effectiveness of interactive television as a distance education model for elementary social science teacher's professional development: Another digital divide? *Journal of Science Education and Technology, 13*, 485-494. https://doi.org/10.1007/

s10956-004-1469-8

Araújo, A. D. (2012). Academic genres in university contexts: An investigation of students' book reviews writing as classroom assignments. In C. Bazerman et al. (Eds.), *International advances in writing research: Cultures, places, measures* (pp. 319-333). Parlor Press e WAC Clearinghouse.

Baron, A. (2002). *Acquisition in interlanguage pragmatics: How to do things with words in a study abroad context*. J. Benjamins Pub. Co.

Beebe, L.M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R.C. Scarcella, El Anderson and S.D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55-73). Newbury House.

Bella, S. (2011). Mitigation and politeness in Greek invitation refusals: Effects of length of residence in the target community and intensity of interaction on non-native speakers' performance. *Journal of Pragmatics, 43*, 1718-1740. https://doi.org/10.1016/j.pragma.2010.11.005

Bella, S. (2014). Developing the ability to refuse: A cross-sectional study of Greek FL refusals. *Journal of Pragmatics, 61*, 35-62. https://doi.org/10.1016/j.pragma.2013.11.015

Brown, P., & Levinson, S. (1978). Universals in language usage: Politeness phenomena. In E. Goody (Ed), *Questions and politeness: Strategies in social interaction*. Cambridge University Press.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Cao, J. (2005). *A pragmatic analysis of the speech act of criticizing in primary and junior high school Chinese lecturer-student talk*. [Master's thesis, Northeast Normal University]. https://www.dissertationtopic.net/doc/875784

Cogo, A. & Dewey, M. (2006). Efficiency in ELF communication: From pragmatic motives to lexico-grammatical innovation. *Nordic Journal of English Studies*, *5*(2), 59–93.

Da Silva, A.J.B. (2003). The effects of instruction on pragmatic development: Teaching polite refusals in English. *Second Language Studies, 22*(1), 55-106.

Derakhshan, A. & Arabmofrad, A. (2018). The impact of instruction on the pragmatic comprehension of speech acts of apology, request, and refusal among Iranian intermediate EFL learners. *English Teaching and Learning, 42*, 75-94. https://doi.org/10.1007/s42321-018-0004-6

Eelen, G. (2001). *A critique of politeness theories*. St Jerome Publishing.

El-Dakhs, D. A. S. (2018). Saying "yes" and "no" to requests: Is it the same in Egyptian and Saudi Arabic? *Language & Dialogue*, 8(2), 236–262. https://doi.org/10.1075/ld.00014.eld

El-Dakhs, D. A. S., Ambreen F., Zaheer, M., & Gusarova, Y. (2019). A pragmatic analysis of the speech act of criticizing in university teacher-student talk: The case of English as a lingua franca. *Pragmatics, 29*(4), 493-520. https://doi.org/10.1075/prag.18028.eld

Félix-Brasdefer, J. C. (2006). Linguistic politeness in Mexico: Refusal strategies among male speakers of Mexican Spanish. *Journal of Pragmatics, 38*, 2158-2187. https://doi.org/10.1016/j.pragma.2006.05.004

Gass, S. M. & Houck, N. (1999). *Interlanguage refusals: A cross-cultural study of Japanese -English*. Mouton de Gruyter.

Goffman, E. (1967). *Interactional ritual: Essays on face-to-face behavior*. Double Day.

Graham, L. J., White, S. L. J., Cologon, K., & Pianta, R.C. (2020). Do teachers' years of experience make a difference in the quality of teaching? *Teaching and teacher education, 96*, 103190. https://doi.org/10.1016/j.tate.2020.103190

Hayashi, T. (1996). Politeness in conflict management: A conversation analysis of dispreferred message from a cognitive perspective. *Journal of Pragmatics, 25*, 227-255. https://doi.org/10.1016/0378-2166(94)00080-8

Hiraga, M. K., Fujii, Y., & Turner, J. M. (2003). L2 pragmatics in academic discourse: A case study of tutorials in Britain. *Intercultural Communication Studies*, *12*(3), 19-36.

Hiraga, M. K., & Turner, J. M. (1996). Differing perceptions of face in British and Japanese academic settings. *Language Sciences*, *18*(3-4), 605-627. https://doi.org/10.1016/S0388-0001(96)00037-X

Holmes, J. (2001). *An introduction to sociolinguistics*. Longman.

Holtgraves, T. (1992). The linguistic realization of face management: Implications for language production and comprehension, person perception, and cross-cultural communication. *Social Psychology Quarterly*, *55*(2), 141-159.

Howatt, A. P. R. & Widdowson, H.G. (2004). *A history of English language teaching* (second edition). Oxford University Press.

Hyland, K. (2000). Hedges, boosters and lexical invisibility: Noticing modifiers in academic texts. *Language Awareness*, *9*(4), 179-197. https://doi.org/10.1080/09658410008667145

Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticizing in written feedback. *Journal of Second Language Writing*, *10*, 185-212. https://doi.org/10.1016/S1060-3743(01)00038-8

Kwon, J. (2004). Expressing refusals in Korean and American English. *Multilingua- Journal of Cross-Cultural and Interlanguage Communication, 23*(4), 339-364. https://doi.org/10.1515/mult.2004.23.4.339
Lakoff, R. (1975). *Language and woman's place*. Harper and Row.
Locher, M. (2006). Polite behavior with relational work: The discursive approach to politeness. *Multilingua, 25*(3), 249-267. https://doi.org/10.1515/MULTI.2006.015
Lü, L. (2018). Role of email in intercultural communication of criticizing in a Chinese English curriculum reform context. *English Language Teaching*, *11*(2), 193-207. https://doi.org/10.5539/elt.v11n2p193
Maltz, D. N., & Borker, R. A. (1982). A cultural approach to male-female miscommunication. In J. J. Gumpers (Ed.), *Language and social identity* (pp.196-216). Cambridge University Press.
Martínez-Flor, A. & Beltrán-Palanques, V. (2013). Teaching refusal strategies in the foreign language classroom: A focus on inductive-deductive treatments. *Journal of English Studies, 11*, 41-67. https://doi.org/ 10.18172/jes.2616
Mauranen, A. (2003). The corpus of English as lingua franca in academic settings. *TESOL Quarterly*, *37* (3), 513–26. https://www.jstor.org/stable/3588402
Mauranen, A., & Ranta, E. (Eds.) (2009). *English as a lingua franca: Studies and findings*. Cambridge Scholars Publishing.
Mills, S. (2003). *Gender and politeness*. Cambridge University Press.
Moafian, F., Yazdi, N., & Sarani, A. (2019). The refusal of request speech act in Persian, English, and Balouchi languages: A cross-cultural and cross-linguistic study. *International Review of Applied Linguistics in Language Teaching* (published online ahead of print). https://doi.org/10.1515/iral-2018-0357
Morkus, N. (2018). Negative pragmatic transfer and language proficiency: American learners of Arabic. *The Language Learning Journal*, 1-25. https://doi.org/10.1080/09571736.2018.1474247
Nixon, R.S., Campbell, B.K., & Luft, J.A. (2016). Effects of subject-area degree and classroom experience on new chemistry teachers' subject matter knowledge. *International Journal of Science Education, 38*(10), 1636-1654. https://doi.org/10.1080/09500693.2016.1204482
Pizziconi, B. (2003). Re-examining politeness, face and the Japanese language. *Journal of Pragmatics, 35*, 1471-1506. https://doi.org/10.1016/S0378-2166(02)00200-X
Riekkinen, N. (2010). "This is not criticizing, but...." Softening criticizing: The use of lexical hedges in academic spoken interaction. *Helsinki English Studies*, *6*, 75-87.
Sa'd, S. H. T., & Gholami, J. (2017). Teaching Iranian elementary EFL learners to say "No" politely: An interlanguage pragmatic study. *TESL-EJ, 21*(1), 1-23.
Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford University Press.
Shishavan, H.B. & Sharifan, F. (2016). The refusal speech act in a cross-cultural perspective: A study of Iranian English language learners and Anglo-Australian speakers. *Language and Communication, 47*, 75-88. https://doi.org/10.1016/j.langcom.2016.01.001
Spencer-Oatey, H. (Ed.). (2000). *Culturally speaking: Managing rapport through talk across cultures*. Continuum.
Statista (2020). The most spoken languages worldwide. https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/
Taguchi, N. & Ishihara, N. (2018). The pragmatics of English as a lingua franca: Research and pedagogy in the era of globalization. *Annual Review of Applied Linguistics, 38*, 80-101. https://doi.org/10.1017/S0267190518000028
Tannen, D. (1986). *That's not what I meant*. Ballantine Books.
Tannen, D. (1990). *You just don't understand*. Ballantine Books.
Tannen, D. (1994). *Gender and discourse*. Oxford University Press.
Tilak, J. B. G. (2011). *Trade in higher education: The role of the General Agreement on Trade in Services (GATS). Fundamentals of Educational Planning* (vol. 95, pp. 154). UNESCO.
Trosborg, A. (1995). *Interlanguage pragmatics*. Walter de Gruyter and Co.
Wannaruk, A. (2008). Pragmatic transfer in Thai EFL refusals. *RELC, 39*(3), 318-337. https://doi.org/10.1177/0033688208096844
Watts, R.J. (2003). *Politeness*. Cambridge University Press.
Widdowson, H.G. (1994). The ownership of English. *TESOL Quarterly*, *28*(2), 377–389.

Appendix (A)

Study: Examining university teacher-student talk

Instrument: Role-play

Dear Faculty Member,

Participation in these role-plays is voluntary, so please feel free to refuse participation if you do not wish to take part in the study.

The aim of this study is to examine how university teachers handle students' requests. The data are collected through role-plays that will be recorded and later transcribed. The participants' identities will be kept strictly confidential. No special reward is offered for participation in the study. Your participation will, however, be highly appreciated by the researchers to help advance scientific research.

Procedure

You will read 10 scenarios involving students' requests. Please, decline these requests in natural spoken English as you would do in real life while actually talking to your students. In case you feel that you would not say anything in real life or you would consent to the requests, please say so and explain your reason(s).

Example:

Your student requests that you cancel your class because they want to study for an exam that is scheduled after your class.

You say: Sorry, dear. I would like to help, but I cannot cancel my class. We are already behind with the syllabus.

Now, you will be reading one scenario at a time and then have your response to your student in natural spoken English recorded.

Situation (1)

Your student requests receiving an A+ grade in the course although her performance grants her only a B grade.

Situation (2)

Your student requests submitting an assignment late although you had already extended the submission deadline twice. She has no valid reasons.

Situation (3)

Your student requests cancelling the class in order to participate in another event held on campus. You have not received any instructions from the university management to allow students to attend the event.

Situation (4)

Your student requests attending class in another section than the one he is registered because he has some friends in the other section.

Situation (5)

Your student requests re-scheduling a major exam at a time that is not convenient to the majority of students in class. He fails to provide valid reasons.

Situation (6)

Your student requests submitting an assignment in groups although you had planned this particular assignment to be completed individually. This is the only individual assignment on the syllabus.

Situation (7)

Your student requests changing the class time because the class time at 9.00 a.m. is too early for him. The university does not allow teachers to change class time.

Situation (8)

Your advisee requests that you allow her to register courses for the new semester without meeting you for the advising session required by the university.

Situation (9)

A student requests that you allow her to register in your class although your class has reached the maximum limit for registered students as per the university policy.

Situation (10)

Your student requests that you do not count his absence on the academic portal. He does not give valid reasons why he may need to miss classes.

Appendix B

Request Refusal Strategies

Beebe, Takahashi and Uliss-Weltz, 1990

Strategy	Example
I. Direct	
A. Performative	I refuse.
B. Non-Performative	
B. 1. No	No
B.2. Negative willingness or inability	I can't, I won't, I don't think so.
II. Indirect	
A. Statement of regret	I'm sorry, I feel terrible.
B. Wish	I wish I could help you.
C. Excuse, reason, explanation	My children will be at home that night.
D. Statement of alternative	
D. 1. I can do X instead of Y	I'd rather do – I'd prefer
D.2. Why don't you do X instead of Y?	Why don't you ask somebody else?
E. Set condition for future/ past acceptance	If you had asked me earlier, I would have…
F. Promise of future acceptance	I'll do it next time – I promise I will.
G. Statement of principle	I never do business with friends.
H. Statement of philosophy	One can't be too careful.
I. Attempt to dissuade interlocutor	
I.1. Threat or negative consequences	I won't be any fun tonight.
I.2. Guilt trip	I can't make a living off people who just order coffee.
I.3. Criticize the requester	That's a terrible idea.
I.4. Request help/ empathy.	I hope you understand my difficult situation.
I.5. Let interlocutor off the hook	Don't worry about it- That's okay.
I.6. Self-defense	I'm trying my best – I'm doing all I can.
J. Acceptance that functions as refusal	
J.1. Unspecific or indefinite reply	I don't know when I can give them to you.
J.2. Lack of enthusiasm	I'm not interested in diets.
K. Avoidance	
K.1. Nonverbal	Silence, hesitation, do nothing
K.2. Verbal	Topic switch, joke, postponement
III. Adjuncts	
A. Statement of positive opinion	That's a good idea.
B. Statement of empathy	I realize you are in a difficult situation.
C. Pause fillers	Uhh, well, uh
D. Gratitude/ appreciation	Thank you.
E. Term of endearment	Dear, sweetheart

Appendix C

Coding Scheme for Internal Modifiers –Trosborg (1995)

Type	Characteristic	Example
1. Downgraders		
Polite markers	Words of phrases that express politeness	"*Please*."
Understaters	Expressions that describe or represent (something) as being smaller or less important than it really is	"*I think it's a bit salty for me, the soup*."
Hedges	Mitigating word or construction used to lessen the impact of an utterance	"*You are making kind of a statement with the pants though."*
Downtoners	Words or phrases which reduce the force of another word or phrase	"*Yes, I mean it might be but it still seems to me at the moment that perhaps it's not a good idea."*
Cajolers	Flattery or insincere expressions to persuade someone to do something.	"*you know… you see*"
Subjectivizers	Expressions of subjective opinion that lower the assertive force of an act	"*I think*" "*I feel*" "*I guess*" "*I believe*" "*I suppose*"
2. Upgraders		
2.1. Swear words	Rude or offensive words	*"damned"*, "*bloody"*
2.2. Overstaters	Words or phrases that describe or explain something in a way that makes it seem more important or serious than it really is	*"absolutely"*, "*terribly"*
2.3. Intensifiers	A word, especially an adverb or adjective, that has little meaning itself but is used to add force to another adjective, verb, or adverb	*"very"*, "*so really"*
2.4. Plus commitors	A word or phrase that reflects the speaker's strong commitment to the utterance	*of course*, *certainly*, *I am sure*

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Fajri, M. S. A., Kirana, A. W., & Putri, C. I. K. (2020). Lexical Bundles of L1 and L2 English Professional Scholars: A Contrastive Corpus-Driven Study on Applied Linguistics Research Articles. *Journal of Language and Education, 6*(4), 76-89. https://doi.org/10.17323/jle.2020.11271

Lexical Bundles of L1 and L2 English Professional Scholars: A Contrastive Corpus-Driven Study on Applied Linguistics Research Articles

Muchamad Sholakhuddin Al Fajri, Angkita Wasito Kirana, Celya Intan Kharisma Putri
Airlangga University

Correspondence concerning this article should be addressed to Muchamad Sholakhuddin Al Fajri, Faculty of Vocational Studies, Universitas Airlangga, Jl. Dharmawangsa Dalam No. 28-30, Surabaya, Jawa Timur, Indonesia, 60286. Email: m-sholakhuddin-al-fajri@vokasi.unair.ac.id

The current study examined the structural and functional types of four-word lexical bundles in two different corpora of applied linguistics scientific articles written by L1 English and L1 Indonesian professional writers. The findings show that L2 writers employed a higher number of bundles than L1 writers, but L2 writers underused some of the most typical lexical bundles in L1 English writing. Structurally, unlike previous studies, this study reports the frequent use of prepositional phrase (PP) - based bundles in the articles of L2 writers. However, besides the high frequency of PP-based bundles, L2 authors also used a high number of verbal phrase-based bundles, suggesting that these L2 writers were still acquiring more native-like bundles. In terms of functional types, L2 writers employed fewer *quantification* bundles than their counterparts. This study has potential implications for teaching English for academic writing. Teachers need to raise their students' awareness of the most frequently used lexical bundles in a specific academic discipline and pay attention to the discourse conventions of academic writing, helping L2 students transition from clausal to phrasal styles.

Keywords: lexical bundles, academic writing, corpus linguistics, applied linguistics

Introduction

Lexical bundles (henceforth LBs) are defined as "recurrent expressions that usually co-occur in natural language use, regardless of their idiomaticity and their lexical status" (Biber et al., 1999, p. 990), and they can be "identified empirically by running a computer program in a corpus of language texts" (Cortes, 2015, p. 205). They are identified automatically by using a computer program with frequency and distribution thresholds set by researchers (Hyland, 2012). LBs play a significant role in improving the quality of scientific writing for both native and non-native speakers. LBs are seen as a significant aspect of fluent linguistic production and a noticeable feature of academic written texts (Hyland & Jiang, 2018). Hyland (2012, p, 153) emphasises the importance of LBs for writers and speakers in three points: "(1) their repetition offers users (and particularly students) ready-made sets of words to work with; (2) they help define fluent use and therefore expertise and legitimate disciplinary membership; (3) they reveal the lexico-grammatical community-authorized ways of making-meanings". LBs are therefore seen to be very important in the formulation of texts.

Regarding the second point, LBs could help writers claim their membership in a particular discourse community (Ädel & Erman, 2012). Wray (2006) explains that, when speaking, people choose a specific turn of phrase that they consider to be related to certain values, styles, and groups. In other words, they help registered community members show solidarity with other members (Esfandiari & Barbary, 2017) and build a disciplinary experienced voice (Pang, 2010). Thus, LBs tend to reflect an authentic part of users' communicative experiences (Hyland & Jiang, 2018).

LBs have also drawn the interest of linguists to explore the role of LBs in teaching and learning academic writing. For instance, in English for Academic Purposes (EAP), exposure to multi-word constructions helps students gain a better understanding of the language style of academic textbooks (Wood & Appel, 2014) since LBs make up 21-52.3% of written discourse (Biber et al., 1999). The absence of multiword expressions thus might indicate a writer's lack of expertise in academic contexts (Wray, 2002). In other words, LBs enable us to distinguish novice and expert users of a certain language in different contexts both in oral and written forms, which are seemingly useful in teaching and learning activities especially in enhancing speaking and writing skills. Noticing the significance of those studies on LBs, this article aims to explore the application of LBs by professional writers for whom English is their native language (L1) and those for whom English is their L2, or foreign language, in academic articles within the discipline of applied linguistics.

Structural and Functional Characteristics of Lexical Bundles

Most LBs are incomplete structural units that comprise two or more words, and they can be categorised into different types of structures as they have strong grammatical correlations (Cortes, 2004). Biber et al. (1999) categorised the grammatical structures of LBs into three common forms: verb-phrase bundles which refer to any word combinations with a verb component such as *it is also possible, can be noted that,* and *it is likely that*; noun-phrase fragments which refer to any noun phrases with post-modifier fragments such as *the use of the*, *the nature of the*, and *the way in which*; and prepositional phrase bundles which include any bundles starting with a preposition plus a noun-phrase fragment such as *in addition to the*, *in the context of,* and *at the end of*. Different registers require different grammatical structures. For example, Biber et al. (1999) argued that LBs in a conversation contain mostly clause fragments (60%) and only 15% of them were phrases. In contrast, in academic prose, LBs were mostly in the form of phrases and less than 5% were clausal constructions.

In terms of functions, Biber et al. (2004) divided LBs into three main categories: stance expression (e.g. *the fact that the*), discourse organizers (*as well as the*), and referential expression (e.g. *one of the most*). Like in the LBs structures, they also found a dramatic difference between oral and written registers in their dependence on LBs' functional types. Conversations or spoken registers mostly use stance expression bundles, while academic writing mostly uses referential expressions. On the basis of Biber et al.'s taxonomy, Hyland (2008a) developed a similar functional taxonomy including research-oriented, text-oriented, and participant-oriented bundles. Their taxonomies differ in the sense that Biber et al.'s (2004) classification was based on both written and oral registers that covered various genres, while Hyland's (2008a) taxonomy was far more specific, focusing on written registers only. Therefore, this study used the functional taxonomy proposed by Hyland (2008a).

Previous Studies on Lexical Bundles

To discover the application of LBs, a corpus has been widely employed for analysing various types of texts (written and spoken) in different languages (e.g. Kim, 2009; Ruan, 2017; Wang, 2017; Wright, 2019). Biber, Conrad, and Cortes (2004) investigated LBs of oral and written registers, including conversations, lectures, textbooks, and academic prose. In several other academic writing genres, linguists have investigated LBs in theses, dissertations, and students' academic writing in a range of academic disciplines. For example, Hyland (2008a) examined the forms, structures, and functions of four-word clusters in a corpus of research articles, dissertations, and theses in four academic disciplines: engineering, microbiology, business, and applied linguistics, while Cortes (2004) compared research articles and students' writing within the disciplines of history and biology. Broadening the scope of the field of knowledge, Kwary et al. (2017) analysed the use of LBs in journal articles of four wide academic disciplines: life sciences, physical sciences, health sciences, and social sciences. These previous studies generally suggest that LBs vary in their discourse functions and their use differs from one discipline or register to another.

Three studies have explored the use of LBs in L2 writing at different proficiency levels. Staples et al. (2013) investigated LBs used by non-native English speakers with different proficiency levels in prompted TOEFL writing. Their research shows that learners at lower proficiency level preferred to employ more clusters than those at higher proficiency levels, lending weight to the second language acquisition theory that as students acquire more proficiency in a second language, they have a tendency to use fewer formulaic structures (Ellis, 1996). Chen and Baker (2016) studied second language development by comparing the use of LBs in L2 English (L1 Chinese) rated learner essays across three levels of Common European Framework of Reference (B1, B2 & C1). Their findings indicate that learners' writing at lower levels is likely to share more features with conversation, relying more on colloquial quantifiers, while the discourse of more advance writing has a more impersonal tone, closer to that of academic prose. The findings also show that the CEFR-B2 level seems to be a transition stage in which learners start to recognise the differences between formal and informal writing. In the field of applied linguistics, the LB studies compared English learner writing with professional writing. For instance, Wei and Lei (2011) examined LBs used by advanced Chinese EFL learners and professional writers in the field of applied linguistics, and Qin (2014) compared clusters used by non-native English graduate students at different levels of study and authors of applied linguistics journal articles. These comparisons may make results difficult to interpret because learners and professional writers have different English and writing proficiency levels and specific writing purposes for unique audiences.

Other studies have compared the use of LBs by L1 English and L2 English writers. For instance, Chen and Baker (2010) compared the use of LBs in academic writing by non-native English students with native peer students and native expert writers. They concluded that the structures and functions of LBs in L1 and L2 student writing is similar. However, L2 students tended to underuse some typical bundles in professional academic writing. In a similar vein, Ädel and Erman (2011) analysed the use and the functions of English LBs in advanced undergraduate writing by L1 English and L1 Swedish students. The findings indicated that L1 writers deployed a larger number of LBs with a wider variety, which are generally similar to the results of phraseological analysis tradition in Second Language Acquisition (SLA) (Ädel and Erman, 2011).

There are only a few studies that investigated the use of LBs by L1 and L2 English academic professionals in international journals. Perez-Llantada (2014), for example, analysed the convergent and divergent usage in academic articles from twelve disciplines. However, the wide variety of scientific disciplines is likely to skew the results of the study since registers, genres, and disciplines all affect the structure and function of LBs (Esfandiari & Barbary, 2017). Pan, Reppen, and Biber (2016) compared the use of LBs by L1 and L2 English professional academics in telecommunications articles. Esfandiari and Barbary (2017) analysed the use of LBs by L1 and L2 English professional academics in psychology research articles. However, there is little work that has been devoted to the use of LBs by L1 English and L2 English professional writers, especially L1 Indonesian writers, in the field of applied linguistics. This study tried to fill this gap by analysing and comparing the use, structure, and function of LBs in applied linguistics academic articles written by English professional writers (L1 English) and Indonesian professional writers (L2 English). Examining applied linguistics articles is significant since journal article authors in this field, whose expertise is related to language, seem to be more aware of the use of formulaic language or bundles, which may affect the use of LBs in their writing. This study therefore can contribute to the ongoing discussion regarding the influence of academic disciplines on the use of LBs in professional academic writing.

Methodology

Corpus Construction

The corpora of the present study are a collection of applied linguistic scientific articles published in a three-year period from 2016 to 2018. The decision to include articles from a three-year period intended to mitigate the over-usage of LBs in special issue publications that probably occur in a certain journal, thus avoiding idiosyncrasies of specific issues or topics. The two corpora are the English corpus (EC) comprising articles written by L1 English academics and the Indonesian corpus (IC) consisting of research articles written by L1 Indonesian professional authors.

The EC was collected from scientific research articles published by internationally reputable journals in the field of applied linguistics that have high Impact Factors (IF) and are indexed in the Scopus database and Social Science Citation Index (SSCI). Meanwhile, the IC was taken from research articles published in Indonesian applied linguistics international journals that are indexed by Scopus or Directory of Open Access Journal (DOAJ) and accredited by Indonesian Ministry of Research, Technology and High Education. The selected Indonesian journals only publish articles in English (see Appendix 1 for the journal list). Both corpora have a similar total number of words, with approximately 1,300,000 words in each corpus. The number of words in each corpus was kept equal since LBs are significantly more sensitive to the number of words in a corpus than the number of articles (Cortes, 2004). Therefore, in our corpus there are fewer articles and journals in the EC as the EC articles were typically longer than the IC articles (see Table 1).

Table 1
Distribution of the corpora

Corpus	Number of Journals	Number of Articles	Number of words
English Corpus (EC)	4	158	1,325,986
Indonesian Corpus (IC)	7	274	1,334,752

We considered the selected journals for the IC to be equivalent to the journals for the EC for several reasons. First, the journals for the IC are peer-reviewed journals, following the academic conventions of international journals. Second, the journals in the IC are indexed by international research article databases. Both EC and IC articles consisted of Introduction, Methods, Results/Findings, and Discussion sections (Martinez et al., 2009). Additionally, articles published in Indonesia tend to reflect the language produced by Indonesian authors in Indonesian contexts for international audiences.

To ascertain the first language of the author(s), we followed the method proposed by Wood (2001) which defines L1 English writers as those whose first and last name are considered as typical native English speaker names and those who are affiliated with institutions in countries that use English as their first language. Therefore, L1 Indonesian writers are also categorised as all writers whose first and last names are considered typical Indonesian names and those who are affiliated with Indonesian institutions. We, thus, excluded articles from Indonesian journals where any of the authors did not fulfil both criteria.

Identification of Lexical Bundles (LBs)

This research considered two criteria in identifying LBs, namely frequency and dispersion. For these relatively big corpora, the standardised frequency threshold was set as 40 occurrences per million words to identify bundles that are often considered as characteristics of target texts (Pan et.al., 2016). The cut-off frequency of 40 per million words was equivalent to minimum raw frequency of 53 occurrences for both corpora. This was calculated by multiplying the cut-off frequency by the corpus size and then dividing the result by one million (Wood & Appel, 2014).

The dispersion threshold holds a significant role to avoid individual author idiosyncrasies. Thus, it needs to be clearly determined to guarantee that the bundles are not only used by a handful of authors or texts. Following Chen and Baker (2010) and Hyland (2008a), the current study only included those bundles that occurred in at least 10% of the total texts in each corpus (Hyland, 2008a). The bundles therefore must occur in at least 27 and 16 articles in the IC and EC respectively. Besides, the length of the word sequences (LBs' length) included in the study must also be determined. This study focused on 4-word bundles "because they are far more common than 5-word strings and offer a clearer range of structures and functions than 3-word bundles" (Hyland, 2008b, p.8).

The corpus software AntConc[1] was used to retrieve the bundles. However, context/content dependent bundles such as *teaching and learning process, as a foreign language* or *in the united states* were excluded since "they are not the 'building blocks' which carry a distinct discourse function" (Chen, 2009, p. 58). In addition, overlapping clusters were also checked manually via concordance analyses to avoid inflated results of quantitative analyses (Chen & Baker, 2010). For example, in the IC *it can be seen* and *can be seen that* occurred as a subset of the 5-word sequence *it can be seen that.* Thus, the lower frequency sequence was combined into the higher frequency one: *it can be seen (that)*. After identifying LBs with the above-mentioned criteria, we compared the frequency, structure, and function of our LBs. In analysing LBs' structures, we used Biber et al.'s (1999) classification which includes noun phrase-based (e.g. *the use of the*), prepositional phrase-based (e.g. *in the case of*), and verb phrase-based bundles (e.g. *it can be seen*). For functional analysis, we employed Hyland's classification (2008b) since it is more relevant to academic writing domain. The classification includes research-oriented bundles, which are used to structure writers' experiences (e.g. *the use of the*), text-oriented bundles, which are concerned with the organisation of the text or discourse (e.g. *in addition to the*), and participant-oriented bundles, which are focused on stance and engagement (e.g. *are likely to be*).

To classify the bundles, each author worked independently and the inter-rater reliability was 97% (structure) and 94% (functions). The discrepancies were then discussed to reach 100% agreement based on their contexts. We acknowledged that there were several multi-functional bundles (e.g. *in the present study*), which also have

[1] Anthony, L. (2018). AntConc (3.5.7) [Computer Software]. Waseda University. http://www.laurenceanthony.net/software

been recognised by the previous studies (Güngör & Uysal, 2020; Salazar, 2014). In these cases, the ambiguous bundles were categorised according to their primary function after cross-checking their contexts.

Results and Discussion

Comparison of Frequency

After excluding the content/context dependant and overlapping bundles, we identified 2,700 and 4,874 lexical bundle tokens in the EC and IC respectively, which comprised 31 different bundle types in the EC, and 51 bundle types in the IC (see Appendix 2 for the list of LBs). This finding is congruent with several previous studies including Pan et al.'s (2016) research, which found that L2 professional academic writers in telecommunications used LBs more frequently than their L1 counterparts, and Güngör and Uysal's (2016) study, which showed that L2 English academic authors in educational sciences used a larger number of LBs than LI English writers. This is, however, contrary to previous studies that compared the use of LBs in the corpora of native and non-native student writing (e.g. Adel & Erman, 2012; DeCock, 2004), in that L2 learners employed a lower number of bundles than native English students. One of the reasons for the lower number of LBs in L2 student writing is the learners' incorrect use of articles (e.g. the omission of required definite articles within LBs) (Shin, Cortes, & Yoo, 2018), which may not apply to L2 professional writing since academics are comparatively competent writers. Thus, it seems that when it comes to professional academic writing, L2 writers including Indonesians are likely to employ a substantially higher frequency of bundle types and tokens than L1 authors. Hyland (2008a) points out that both groups of expert writers use the fewest clusters compared to master's and doctoral students. This indicates that L2 expert writers still rely on formulaic expressions to some extent. This greater reliance on LBs might also suggest the comparatively smaller vocabulary of L2 writers, while L1 professional writers might be able to present their arguments in a more flexible manner.

Additionally, a comparative analysis showed that 14 bundles were shared between both groups (e.g. *on the other hand*, *at the same time,* and *as well as the*), indicating that nearly half of the LBs used by native writers (14 of 31) were also employed by non-native writers. Most of the shared bundles (9 of 14) are text-oriented ones, which might not be surprising as text-oriented bundles are common in soft sciences research articles (Hyland, 2008b). However, only two LBs (*on the other hand* and *the results of the*) were shared among the top ten most frequently used bundles. The top two most frequent LBs in the EC (*the extent to which, in the present study*) were not used by IC writers. Qin (2014) also reported the absence and low frequency of *the extent to which* in the master's and doctoral student writing in applied linguistics. This indicates that these L2 advanced writers were still not aware of some typical academic LBs in the field of applied linguistics.

Comparison of Structural Types

Table 2 shows the distribution of structural subcategories of the LBs used by both English and Indonesian writers. Log-likelihood tests comparing the number of tokens in each category were conducted to measure significant differences across the corpora. The results demonstrated that IC writers used considerably more VP-based and PP-based bundle tokens than EC writers. For NP-based bundles, although this category does not indicate substantial differences, the subcategory shows a different pattern. L1 English authors employed significantly more NP with other post-modifier fragments, while L1 Indonesian authors used significantly more NP with *of*-phrase. Examples of NP-based bundles are presented below:

> This paper explores *the extent to which* corpus linguistics can contribute to the study of language ideology in both explicit and implicit forms in news media. (EC: NP-other)
>
> Moreover, *the use of the* nomination strategy also indicates that the Jakarta Post tries to avoid ambiguity. (IC: NP-*of*)

In terms of the structural distribution of LBs (see Table 3), the comparison of the percentages of the main structural categories in both corpora shows that EC writers mostly employed phrasal bundles (NP- and PP-based bundles), accounting for 84% of bundle types and tokens. This finding is congruent with previous studies that point out that the frequency and percentage of phrasal bundles are higher than clausal bundles in English academic prose (Biber et al., 2011, 2013; Biber & Gray, 2011), and L1 English professional writers used considerably more NP- and PP-based bundles than VP-based bundles in academic research articles (Pan et al., 2016).

Table 2
Distribution of structural subcategories

Categories	Subcategories	Types		Tokens		LL
		EC	IC	EC	IC	
NP-based	Noun phrase + of (e.g. *the use of the*)	6	9	464	828	101.57**
	Noun phrase with other post-modifier fragment (e.g. *the extent to which*)	4	1	388	54	286.78**
	Total	10	10	852	882	0.34
PP-based	Prepositional phrase with embedded *of* phrase (e.g. *in the context of*)	10	10	787	1031	31.26**
	Other prepositional phrase fragments (e.g. *in relation to the*)	6	10	635	911	47.74**
	Total	16	20	1422	1942	77.31**
VP-based	copula *be* + noun phrase/prepositional phrase (e.g. *is one of the*)	-	2	-	281	387.70**
	Anticipatory *it* + verb phrase/adjective phrase (e.g. *it can be seen*)	1	7	95	644	453.86**
	Passive verb + prepositional phrase (e.g. *can be seen in*)	2	4	138	362	102.54**
	(Verb/adjective) to-clause fragment (e.g. *to be able to*)	1	3	61	289	159.87**
	(Verb phrase) + *that* clause fragment (e.g. *that there is a*)	-	3	-	268	369.76**
	Total	4	19	294	1844	1241.51**
Others	as well as the	1	2	132	206	15.85**
Total		31	51	2700	4874	-

Table 3
Distribution of structural categories

Categories	Types (%)		Tokens (%)	
	EC	IC	EC	IC
NP-based	32.26	19.61	31.56	19.10
PP-based	51.61	39.22	52.67	39.84
VP-based	12.90	37.25	10.89	37.83
Others	3.23	3.92	4.89	4.23
Total	**100.00**	**100.00**	**100.00**	**100.00**

On the other hand, IC writers predominantly used PP-based and VP-based clusters, accounting for 77% of the types and 78% of the tokens. This result contrasts somewhat with previous studies on writing from other disciplines that revealed that VP-based bundles were more frequent than the other categories (NP and PP) in L2 writing. For example, Pan et al. (2016) found more VP-based bundles (58%) than NP- and PP-based bundles (34%) in L1 Chinese professional writing in telecommunications journals, and Chen and Baker (2010) found more VP-based bundles (52.3%) than NP- and PP-based bundles (47.5%) in L2 learners' writing. With the high frequency of PP-based bundles in the IC corpus, it therefore suggests that L2 professional Indonesian writers in the field of applied linguistics demonstrate relatively higher academic writing proficiency since both L1 and L2 writers will shift from the clausal to phrasal style as their writing skills increase (Bychkovska & Lee, 2017; Pan et al., 2016; Staples et al., 2013). Wei and Lei (2011) also found that advanced Chinese EFL learners in the discipline of applied linguistics used more NP- and PP-based four-word bundles than VP-based formulaic sequences, which is in contrast to previous research on Chinese EFL learner writing in other disciplines (e.g. Bychkovska & Lee, 2017; Pang, 2009). The use of more phrasal bundles may be due to the fact that applied linguistics majors require a more advanced level of English even at the undergraduate level so students and researchers in this field might be more aware of the conventions of English academic register. However, it should be noted that the considerable use of clausal constructions or VP-based bundles in the IC (37% of the types and 38% of the tokens) indicates that these L1 Indonesian expert writers were still in the process of obtaining more appropriate academic English or acquiring more native-like LBs.

Comparison of functional types

As shown in Table 4, the two corpora contained a relatively similar proportion of the three main functional categories. Text-oriented bundles (types and tokens) comprised the largest proportion in both the EC and IC, with similar percentages at 55% and 51% respectively for types, and 57% and 55% respectively for tokens, while participant-oriented bundles constituted the smallest proportion, accounting for 6% of types and tokens in the EC, and 8% of types and 9% of tokens in the IC. Text-oriented bundles are significantly used in applied linguistics journal articles, more generally in the social sciences, to "provide familiar and shorthand ways of engaging with a literature, providing warrants, connecting ideas, directing readers around the text, and specifying limitations", representing the more discursive and evaluative patterns of arguments in the soft sciences (Hyland, 2008b, p. 16). This finding echoes previous related research on LBs used by Persian writers in psychology research articles (Esfandiari & Barbary, 2017), and LBs used by Chinese writers in telecommunications journal articles (Pan et al., 2016). This indicates that L1 and L2 English professional writers do not differ much in the proportion of the main functional distributions of LBs.

Table 4

Distribution of functional types

Categories	Types (%)		Tokens (%)	
	EC	IC	EC	IC
Research-oriented	38.71 (39)	41.18 (41)	36.59 (37)	35.80 (36)
Text-oriented	54.84 (55)	50.98 (51)	57.30 (57)	54.97 (55)
Participant-oriented	6.45 (6)	7.84 (8)	6.11 (6)	9.23 (9)
Total	**100.00**	**100.00**	**100.00**	**100.00**

However, Table 5 shows profound differences in the frequency of bundle tokens across the two corpora in nearly all subcategories of functional types. The results of log-likelihood tests comparing the number of tokens in each category demonstrate that IC (L2) writers used LBs significantly more frequently than EC (L1) writers in most functional subcategories (*procedure, description, transition, resultative, structuring, framing,* and *engagement*), but less recurrently in the subcategories of *location* and *quantification.* While L2 writers used fewer bundle tokens of *location* than L1 writers, the difference in the number of tokens was not significant. *Quantification* was the only subcategory that L2 writers employed bundle tokens significantly less recurrently than L1 writers did. EC writers used four types of *quantification* bundles (*the extent to which, the total number of, a wide range of*, and *the degree to which*), which were not used by IC writers. Cortes (2004, p. 415) found the absence of *quantification* LBs in the learner academic writing in biology. Chen and Baker (2010) also noticed the absence of extent/degree modifiers such as *the extent to which* and *the degree to which* (quantifying bundles) in learner writing. From this, we can assume that L2 writers including professional academic writers pay less attention to the use of quantifying bundles in their academic writing.

Table 5

Distribution of sub-functional types

Categories	**Subcategories**	**Types**		**Tokens**		**LL**
		EC	**IC**	**EC**	**IC**	
Research-oriented	Location (e.g. *at the end of*)	4	4	339	302	2.39
	Procedure (e.g. *the use of the*)	2	9	128	841	582.13**
	Quantification (e.g. *one of the most*)	4	3	356	269	12.73**
	Description (e.g. *the ways in which*)	2	5	165	333	56.70**
	Total	12	21	988	1745	207.49**
Text-oriented	Transition signals (e.g. *on the other hand*)	2	4	276	474	51.60**
	Resultative signals (e.g. *the results of the*)	2	7	194	695	296.30**
	Structuring signals (e.g. *as shown in figure*)	5	6	437	509	5.02*
	Framing signals (e.g. *in the case of*)	8	9	640	1001	77.71**
	Total	17	26	1547	2679	299.55**

Categories	Subcategories	Types		Tokens		LL
		EC	IC	EC	IC	
Participant-oriented	Stance features (e.g. *it is possible that*)	-	-	-	-	
	Engagement features (e.g. *it should be noted*)	2	4	165	450	135.39**
	Total	2	4	165	450	135.39**
Total		31	51	2700	4874	-

Framing is the subcategory that makes up the largest proportion of bundles in both corpora, which is congruent with the findings of Hyland's (2008a) and Hyland and Jiang's (2018) studies on applied linguistics research articles. This subcategory is used to elaborate arguments by specifying cases (*in the case of*) and pointing to limitations (*with the exception of*) (Hyland, 2008a). For example:

> Moreover, students needed to draw on similar skills and knowledge to those required for their disciplinary assignments, especially *in the case of* students from Applied Linguistics/ TESOL and Education. (EC)
>
> It will also benefit those teaching Computer-Assisted Language Learning (CALL) courses in EFL contexts, particularly *in the context of* higher education in Indonesia. (IC)

However, unlike previous studies that suggest that L2 expert writers employed four-word framing bundles significantly less frequently than their counterparts (e.g. Esfandiari & Barbary, 2017; Pan et al., 2016), this study shows the opposite. This may be due to the higher frequency of PP-based bundles in the IC corpus, since framing clusters consist mainly of PP-based bundles (Pan et al., 2016).

In the participant-oriented category, we did not find any *stance* bundles in either corpora. *Stance* refers to how writers explicitly convey their attitudes, epistemic and affective judgments, and evaluations (e.g. *it is obvious that* and *are likely to be*) (Hyland, 2008a, 2008b). This finding might lend weigh to Hyland and Jiang's (2018) analysis that showed a dramatic decline (-38.2%) in the *stance* bundle tokens of applied linguistics journal articles over a 50-year period from 1965 to 2015. Engagement bundles also experienced a decrease, but not significantly, around -9.2% (Hyland & Jiang, 2018). In the present study, similar to Esfandiari and Barbary's (2017) and Pan et al.'s (2016) studies, we found that L2 English writers used relatively more *engagement* bundle types and tokens than L1 English writers to engage readers and guide them to particular interpretations. For instance:

> *It is important to* note that this study uses the term reader knowledge rather than reader characteristics to specifically refer to linguistic knowledge in terms of grammatical knowledge. (IC).
>
> From the teachers' responses, *it can be seen* that teachers in this study share similar beliefs in grading the students although they come from different schools. . . (IC)

It is important to note that three of the four *engagement* clusters used by IC writers are in the form of anticipatory *it* and passive constructions (*it can be seen, it can be concluded*, *it can be said*), which indicate an impersonal tone. This might be influenced by the writers' preference for impersonality in their academic writing. Like in Hong Kong (Hyland, 2008a) and mainland China (Wei & Lei, 2011), impersonality devices such as passive patterns in academic writing are also suggested in Indonesian universities, accounting for the overuse of passive structure bundles. This might support Li, Franken, and Wu's (2018) study, which argued that one of the reasons for the different bundle selections of L2 learners is classroom learning.

Conclusion

In the present study, we compared the frequency, structure, and function of four-word LBs of L1 and L2 professional writers in research articles published in applied linguistics journals. In terms of frequency, L2 writers employed a higher number of bundles than L1 writers, but L2 writers underused some of the most typical LBs in native writing, such as *the extent to which*. Structurally, unlike previous studies (e.g. Chen & Baker, 2010; Pan et al., 2016), in this study, L2 writers, more specifically Indonesian writers, predominantly used PP-based clusters, which may indicate that L2 professional writers in the field of applied linguistics demonstrated relatively high academic writing proficiency. However, despite the high frequency of PP-based bundles, L2 authors still used a significant number of VP-based bundles, suggesting that these Indonesian professional writers (L2) were still acquiring more native-like LBs, such as NP- and PP-based formulaic sequences. Functionally, L1 English and L2 English professional writers did not differ much in the proportion of the main functional distributions of LBs, but showed marked differences in nearly all subcategories of functional types. L2 writers employed *quantification* bundles significantly less frequently than their counterparts, which also had been reported by the previous studies (Chen & Baker, 2010; Cortes, 2004). These findings show that L2 English journal article authors in the discipline of applied linguistics are more aware of the use of LBs, compared to other disciplines. This lends weight to the argument that academic disciplines influence the use of LBs in professional academic writing.

It is important to point out the two limitations of this research. First, the method of determining L1 and L2 writers may not be a satisfactory approach since it may not be accurate to identify the first language of the author(s) by means of their names and institutions. Second, log-likelihood tests have recently been suggested as problematic for measuring lexical variations by Bestgen (2017). However, the test is still commonly used by many lexical bundle researchers including Hyland and Jiang (2018) so we chose this test together with the calculation of percentages for easier comparisons with other bundle studies. Therefore, future studies can employ different methods of determining L1 and L2 writers, such as by looking at their profile and education background, and different lexical variation measures, such as proportions. Also, examining crosslinguistic influence of L1 Indonesian on the use of L2 English bundles can be conducted.

Despite these limitations, the findings of this study have potential implications for academic writing pedagogy. Teachers are suggested to raise learners' awareness of the most frequently used LBs in a specific academic discipline and to pay attention to the discourse style of academic writing, helping L2 students shift from clausal to phrasal bundle use. Focusing on the most typical bundles in native writing that are underused by L2 writers, such as *quantification* bundles, is also suggested.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81-92. http://dx.doi.rg/10.1016/j.esp.2011.08.004

Bestgen, Y. (2017). Getting rid of the Chi-square and Log-likelihood tests for analysing vocabulary differences between corpora. *Quaderns de Filologia: Estudis Lingüístics*, *22*, 33-56. http://dx.doi.rg/10.7203/qf.22.11299

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371-405. http://dx.doi.rg/10.1093/applin/25.3.371 %J Applied Linguistics

Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: The influence of written language use. *English Language and Linguistics, 15*(2), 223-250. http://dx.doi.rg/10.1017/S1360674311000025

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35. http://dx.doi.rg/10.5054/tq.2011.244483

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond t-units—A response to WeiWei Yang. *TESOL Quarterly, 47*(1), 192-201. http://dx.doi.rg/10.1002/tesq.84

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education Limited.

Bychkovska, T., & Lee, J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30*, 38-52. http://dx.doi.rg/10.1016/j.jeap.2017.10.008

Chen, Y.-H. (2009). *Lexical bundles across learner writing development* [Unpublished doctoral dissertation].

Lancaster University.

Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30-49. http://dx.doi.rg/10125/44213

Chen, Y.-H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics, 37*(6), 849-880. http://dx.doi.rg/10.1093/applin/amu065 %J Applied Linguistics

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*(4), 397-423. http://dx.doi.rg/10.1016/j.esp.2003.12.001

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora, 3*(1), 43-57. http://dx.doi.rg/10.3366/E1749503208000063

Cortes, V. (2015). Situating lexical bundles in the formulaic language spectrum: origins and functional analysis development. In V. Cortes & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber* (pp. 197-218). John Benjamins.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures, 2*(2), 225-246.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91-126. http://dx.doi.rg/10.1017/S0272263100014698

Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes, 29*, 21-42. http://dx.doi.rg/10.1016/j.jeap.2017.09.002

Güngör, F. & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching, 9*(6), 176-188. http://dx.doi.rg/10.5539/elt.v9n6p176

Gungor, F. & Uysal, H. H. (2020). Lexical bundle use and crosslinguistic influence in academic texts. *LINGUA, 242*, 1-22. http://dx.doi.rg/10.1016/j.lingua.2020.102859

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41-62. http://dx.doi.rg/10.1111/j.1473-4192.2008.00178.x

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4-21. http://dx.doi.rg/10.1016/j.esp.2007.06.001

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150-169. http://dx.doi.rg/10.1017/S0267190512000037

Hyland, K., & Jiang, F. (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus Linguistics, 23*(4), 383-407. http://dx.doi.rg/10.1075/ijcl.17080.hyl

Kim, Y. (2009). Korean lexical bundles in conversation and academic texts. *Corpora, 4*(2), 135-165. http://dx.doi.rg/10.3366/E1749503209000288

Kwary, D. A., Ratri, D., & Artha, A. F. (2017). Lexical bundles in journal articles across academic disciplines. *Indonesian Journal of Applied Linguistics, 7*(1), 131-140. http://dx.doi.rg/10.17509/ijal.v7i1.6866

Li, L., Franken, M., & Wu, S. (2019). Chinese postgraduates' explanation of the Sources of sentence initial bundles in their thesis writing. *RELC Journal*, *50*(1), 37-52. http://dx.doi.rg/10.1177/0033688217750641

Martínez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes, 28*(3), 183-198. http://dx.doi.rg/10.1016/j.esp.2009.04.003

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes, 21*, 60-71. http://dx.doi.rg/10.1016/j.jeap.2015.11.003

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal, 47*(1), 10-11.

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes, 14*, 84-94. http://dx.doi.rg/0.1016/j.jeap.2014.01.002

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System, 42*, 220-231. http://dx.doi.rg/10.1016/j.system.2013.12.003

Ruan, Z. (2017). Lexical bundles in Chinese undergraduate academic writing at an English medium university. *RELC Journal, 48*(3), 327-340. http://dx.doi.rg/10.1177/0033688216631218

Salazar, D. (2014). Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching (Vol. 65). Amsterdam: John Benjamins Publishing Company. http://dx.doi.rg/10.1075/scl.65

Shin, Y. K., Cortes, V., & Yoo, I. W. (2018). Using lexical bundles as a tool to analyze definite article use in L2 academic writing: An exploratory study. *Journal of Second Language Writing*, *39*, 29-41. http://dx.doi.

rg/10.1016/j.jslw.2017.09.004

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL IBT writing section. *Journal of English for Academic Purposes, 12*(3), 214-225. http://dx.doi.rg/10.1016/j.jeap.2013.05.002

Wang, Y. (2017). Lexical bundles in spoken academic ELF. *International Journal of Corpus Linguistics, 22*(2), 187-211. http://dx.doi.rg/10.1075/ijcl.22.2.02wan

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal, 42*(2), 155-166. http://dx.doi.rg/10.1177/0033688211407295

Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 71-83). Cambridge University Press.

Wood, D. C., & Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes, 15*, 1-13. http://dx.doi.rg/10.1016/j.jeap.2014.03.002

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wray, A. (2006). Formulaic language. In E. K. Brown & A. Anderson (Eds.), *Encyclopedia of language and linguistics* (pp. 590-597). Elsevier.

Wright, H. R. (2019). Lexical bundles in stand-alone literature reviews: Sections, frequencies, and functions. *English for Specific Purposes, 54*, 1-14. http://dx.doi.rg/10.1016/j.esp.2018.09.001

Appendix 1

Journal list

No	English Journals	Indonesian Journals
1	Annual Review of Applied Linguistics	Indonesian Journal of Applied Linguistics
2	Applied Linguistics	TEFLIN journal
3	Journal of Second Language Writing	PAROLE: Journal of Linguistics and Education
4	Studies in Second Language Acquisition	Celt: A Journal of Culture, English Language Teaching & Literature
5		IJELTAL (Indonesian Journal of English Language Teaching and Applied Linguistics)
6		Lingual: Journal of Language and Culture
7		TELL-US Journal

Appendix 2

Lexical bundle lists (shared bundles are bolded)

No	Bundles in the English corpus		Bundles in the Indonesian Corpus	
	Bundles	Frequency	Bundles	Frequency
1	the extent to which	178	**in the form of**	300
2	in the present study	153	**on the other hand**	245
3	**on the other hand**	144	the result of the	186
4	**as well as the**	132	it can be seen + (that)	162
5	in the case of	123	(this) + is in line with + (the)	147
6	**at the same time**	118	to be able to	145
7	**the results of the**	115	it was found that + (the)	143
8	**in the context of**	107	(it) + can be concluded that (the)	139
9	in the current study	100	is one of the	134
10	the ways in which	99	**the results of the**	133
11	**in terms of the**	97	can be seen in + (the)	121
12	**it is important to**	95	**the use of the**	117
13	**the end of the**	92	**in the context of**	112
14	**on the basis of**	82	that the use of	107
15	**as a result of**	79	can be seen from (the)	105
16	**at the end of**	74	in the process of	105
17	as can be seen	70	**as well as the**	99
18	can be seen in	68	in this study the	99
19	**the use of the**	67	on the use of	91
20	the nature of the	66	**in terms of the**	90
21	the total number of	66	in this case the	85
22	participants were asked to	61	**at the same time**	84
23	with regard to the	61	to find out the	83
24	of the present study	59	**the end of the**	82
25	a wide range of	58	as one of the	79
26	in the field of	58	it can be said + (that)	79
27	**in the use of**	58	the meaning of the	79
28	the fact that the	57	**at the end of**	77
29	at the beginning of	55	in other words the	76
30	**in the form of**	54	can be used to	75
31	the degree to which	54	in relation to the	73
32			in accordance with the	72
33			that there is a	72
34			it means that the	71
35			**it is important to**	70
36			the findings of the	65
37			**in the use of**	64
38			of this study is	62
39			can be found in	61
40			in order to make	61

No	Bundles in the English corpus		Bundles in the Indonesian Corpus	
	Bundles	**Frequency**	**Bundles**	**Frequency**
41			in this research the	61
42			it is found that	60
43			above it can be	59
44			**on the basis of**	58
45			that there is no	57
46			the implementation of the	57
47			one of the most	56
48			**as a result of**	55
49			an important role in	54
50			in addition to the	54
51			the findings of this	53

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Hashemnezhad, H. (2020). Applying Freire's Critical Pedagogy to Iranian EFL Bilingual and Monolingual Speaking Performance. *Journal of Language and Education, 6*(4), 90-104. https://doi.org/10.17323/jle.2020.10349

Applying Freire's Critical Pedagogy to Iranian EFL Bilingual and Monolingual Speaking Performance

Hossein Hashemnezhad
Islamic Azad University

Correspondence concerning this article should be addressed to Hossein Hashemnezhad, Department of English Language Teaching, Khoy Branch, Islamic Azad University, Khoy, Iran.
Email: h_hashemnezhad2000@yahoo.com

The purpose of the present pre-experimental study is to examine the extent to how Critical Pedagogy (CP) may function in EFL teaching in Iran. Compared with the growing but far from conclusive body of research, virtually few studies have been covered comparatively among monolinguals and bilinguals. Also, no studies have examined Freire's CP among monolinguals and bilinguals especially in Iran, which considered as two privileged and less privileged groups respectively. Therefore, this study was done among sixty Iranian monolingual and bilingual university sophomores to know if CP affects them differently. The study is done under two available classes in bilingual and two other classes in monolingual context. The first groups received problem-posing and the second groups were exposed to banking model. The scoring procedure of participants' performance was based on IELTS speaking band descriptors. Findings reveal that applying problem-posing model cause improvement in speaking performance of both monolingual and bilingual learners than banking model. More importantly, it is concluded that there is no significant difference between monolinguals and bilinguals in terms of problem-posing model, perhaps because the standards of educational justice have been partially observed among both communities. Finally, implications were drawn for EFL teachers and syllabus designers.

Keywords: critical pedagogy, Freire, monolingual, bilingual, speaking

Introduction

CP has been discussed and used in education for decades, but there have not been many studies in English language teaching on implementing the basic principles of Brazilian theorist Paulo Freire's philosophy of education and critical pedagogy. Van Canh (2018) considered Pennycook's (1990a) perspective on CP as an approach that "seeks to understand and critique the historical and sociopolitical context of schooling and to develop pedagogical practices that aim not only to change the nature of schooling, but also the wider society" (p.113).

The purpose and process of education are central aspects discussed by critical pedagogy. Principally, education illuminates the relation among knowledge, authority, and power (Giroux, 1994; as cited in Steven, 2002). The origin of critical pedagogy dates back to the Frankfurt School, yet, it is through the work of Freire (1970), which infuses critical pedagogy into present-day education. Freire, the pioneering figure of CP, mainly discussed two aspects of power and social injustice and put forward critical pedagogy as an approach to confront with these concepts (Baladi, 2007).

Freedom and justice in education are two central concepts which are sometimes overlooked by societies in general and teachers and policymakers in particular. This leads to creation of a bipolar and oppressed society, which as Ramos (2005) stated, includes two major groups of oppressors and oppressed. The first group owning the maximum power and authority and the second group with lack of enough autonomy cause an imbalance of democracy in society. The oppressed can neither help nor able to liberate themselves. Freire (1970) introduced his 'pedagogy of oppressed' especially for such societies. He stated that to resolve the oppressor-oppressed contradiction, we need a critical intervention. His pedagogy grew to theorize this critical intervention. This intervention requires "liberated pedagogues, fully conscientized in such pedagogy" (Esteva, Stuchul, & Prakash, 2005. p. 14). This pedagogy was first considered suitable for the oppressed, "it would then become a pedagogy for all people in the process of permanent liberation, a pedagogy of humankind" (Esteva, Stuchul, & Prakash, 2005, p. 14). To apply liberation through critical intervention, the teacher should attempt to change the reality of the oppressed by teaching some necessary virtuous and political moralities as well as some specific instruments.

In describing the power relations embedded in society, Auerbach (1995) mentioned that power is unevenly and unfairly distributed in society and the dominant class exercises power through coercion and through consent. For these reasons, the oppressor and the oppressed will always exist. The dominant monolingual group (Persians who live in Tehran) possesses some privileges due to power and facility they own in the capital of Iran (Tehran) while the minority monolingual group (Azeris who live in Tabriz) lacks the fair power and privileges. Iran's population is estimated at 79926270 (2016 estimate, Vice Presidency Plan and Budget Organization, Statistical Center of Iran). It is divided equally between men and women. Campo (2009) reports that Persians (the majority monolingual group in the present study) make up 51 percent of the population. Azeris, the Turkic people, are the largest non-Persian minority and constitute 24 percent of the population (the minority bilingual group in this study). In Iran, minority groups have the right to use their language and respect ethnic languages, but the national language is still dominant. Minority groups are struggling to learn because they lack the power and influence of the majority language.

Tehran, the capital of Iran and Tehran province, is the most populous city in Iran with 8,737,510 populations (2016 estimate, Vice Presidency Plan and Budget Organization, Statistical Center of Iran). Persian is the native language and official language of people. Tabriz, the capital of East Azerbaijan province, is the sixth populous city of Iran with 1,773,033 populations (2016 estimate, Vice Presidency Plan and Budget Organization, Statistical Center of Iran). The native language spoken in Tabriz is Azerbaijani language which is a Turkic language and their official language is Persian.

Although there is a general view that bilingualism has many benefits in learning, it can take a different perspective in terms of the concept of justice, especially educational justice. Justice is a necessary condition of development in any country, and equality refers to provide equal chances for every individual. The monolinguals living in Tehran seem to have more privileges than those in the non-capital cities because of many facilities. According to Shahraki-Sanavi, Nasseri, Shahraki-Sanavi, and Salehiniya (2019), social justice in general and educational justice in specific can include equal input, equal process, and equal output. They described the results of some research in addressing educational challenges. Some challenges include:

> Professors' fair expectations of students and their fair behavior with students, having access to good professors, … university's disciplinary procedures, having access to culturally based learning resources, attaining appropriate language support, having access to programs and resources tailored to individual abilities, disabilities, interests, talents, special needs, access to technology and sports facilities and programs, participation in educational administration, policymaking, bulletin boards, and advisory bodies (Shahraki-Sanavi, Nasseri, Shahraki-Sanavi, & Salehiniya, 2019, p. 2).

Due to the aforementioned facilities, the monolingual learners of the capital city in the present study seem to enjoy more advantages of educational justice than those bilinguals who live in a non-capital city and enjoy less educational justice. One of the tenets of CP is to establish justice among all learners and educators. As Haque (2007) argues, critical pedagogy tries to provide a framework in the manner that it sanctions inequalities and injustice. Also, some theorists in CP field such as Nortan and Toohey (2004) took the sophisticated political and socio-historical aspects of language learning and teaching into account and regarded them as vital essence of language learning and teaching.

To operationalize the two poles of the contradiction, Freire (1970) introduced *banking education* as an instrument for oppression and *problem-posing* education as a symbol or tool for liberation. *Banking education* develops when teachers transfer the contents and knowledge to those of the students (Bartolome, 1994) and *problem-posing* occurs when education proceeds through meaningful and real-life dialogues between teacher and student with the aim of acting on the world to improve it and support students' political and personal developments. In his later years, Freire (1997) wrote a book entitled ‘pedagogy of freedom’. He further emphasized the two poles of contradiction and considered universal love for others as the universal ethic of the human being.

Perumal (2016) implied that one manifestation of critical pedagogy in disadvantaged societies is democracy. She added that in societies such as South Africa, “the materiality of democracy is limited to idealist slogans” (p. 762). She further adds that lack of democracy, political, social, and structural inequalities, poverty, and violence in communities proved that in societies where the necessary foundation to CP has not been provided, the implementation of CP will not be effective. In relation to having a free attitude and free expression of ideas in CP, Joseph Jeyaraj (2020) conducted a qualitative study using interviews with Malaysian students to explore their perceptions of critical pedagogy and social justice. Most of the students expressed that CP was neglected and teachers were the only source of authorized knowledge. They also believed that talking about critical issues could be problematic to them and perhaps endanger their safety. They stated that the imbalance of power in the relationship between teacher and learners and the fear of expressing their opinions freely caused most learners to be cautious in their critical behavior and unwilling to share their opinions and knowledge. Of course, despite some limitations, it seems that there is a good desire for CP in Malaysia.

Correa (2015) defined flipped foreign language classroom and compared it with critical pedagogy. She introduced the basic tenets and advantages of each. Then the fundamental principles of flipped foreign language classrooms were presented and compared to the basic tenets of critical pedagogy. She found that flipped language classes and critical pedagogies are almost identical and seem to be two sides of the same coin, especially when it comes to shifting responsibility of students, scaffolding and rejecting the banking model of education. Recently, Mitsikopoulou (2020) introduced the genre approach and combined Freire’s critical pedagogy with Halliday’s Systemic Functional Grammar. She sought to turn learners into critical text analysts by encouraging them to critically evaluate different texts, to understand different meanings in different social contexts, and to discover the relationship between text and context.

Some researchers (Norton & Toohey, 2004; Pennycook, 1999; 2001) stated that applying critical pedagogy is essential to English language teaching. It affects the general process and outcomes of English language learning and teaching. In another study, Brown (2013) carried out a collaborative action research to explore the perception of teachers implementing Freire’s theory of problem solving pedagogy into middle school classroom. During the study, teachers compared and contrasted Freire’s problem-posing pedagogy with the banking notion of education. The main findings grouped under opportunities and barriers towards applying problem-posing pedagogy. ‘Increased depth of student learning’ was one out of the three major opportunities towards using problem-posing pedagogy. In EFL context, especially with young learners, Moorhouse (2014) examined the effect of critical pedagogy with young EFL learners in Hong Kong primary school. The findings proved that although using CP may be challenging, it can be interesting and raise their awareness. It was further concluded that applying CP might empower students, especially those marginalized by society. In another EFL study, Benites (2012, as cited in Kamali Sarvestani & Yamini, 2016) examined the role of CP in teaching and learning ESL/EFL. It was concluded that applying CP in ESL and EFL settings led to better and successful bilingual learners. Hong and White (2012) investigated the impact of CP in Chinese EFL writing programs. According to the holistic approach, all language skills are interrelated, and the development of one skill may lead to the development of another. This is why the effect of CP on other skills than speaking is reported here. They concluded that critical pedagogy could empower students or writers and increase their awareness. Regarding the impact of critical pedagogy on foreign language education, especially the challenges and benefits of using this model, Yulianto (2020) conducted a study in Subang. The results showed that there were 23 activities necessary to use this type of education in schools and 4 main conditions had to be provided for the implementation of critical education. Finally, the challenges and benefits of using this pedagogy were expressed in the form of 3 challenges and 4 benefits, respectively. The three types of challenges were (1) students’ limited English proficiency, (2) cultural problems, and (3) pre-determined curriculum and the four kinds of benefits were (1) new experience, (2) new knowledge acquisition, (3) increased language awareness, and (4) improved English proficiency. In another similar study conducted in Punjab, Pakistan, Kalsoom, Kalsoom, and Mallick (2020) addressed the problems of using critical pedagogy in Pakistan through a mixed-method study using interviews with students. The results showed that the four main reasons for using the banking model instead of the critical model were the large number of students in each class, lack of resources and trained staff, lengthy and fixed syllabi and finally lack of motivation and interest among students. The results also indicated that for the application and formation of critical pedagogy in the country, the necessary knowledge about the benefits of critical pedagogy should be given to teachers and students and the necessary resources and facilities for better learning and education should be provided to both teachers and learners. In terms of examining the impact of CP on emotional and social dimensions, researchers in other disciplines have also studied the impact of this model. Manca, Gormley, Johnston, and Hart (2020) have also studied the impact of CP in medical sciences and have concluded that this model facilitates professional power and improves social accountability and socio-cultural responsibility of medical staff due to increasing awareness and critical consciousness of medical staff and ultimately leads to foster empathy and compassion among doctors and medical staff.

In this area, another study was conducted by Barjesteh, Alipour, and Vaseghi (2013) in Iranian context. They studied the impact of critical pedagogy strategies on reading comprehension performance of Iranian EFL learners. Findings revealed that critical pedagogy strategies had a significant effect on reading comprehension ability of EFL learners. Abdollahzadeh and Haddad Narafshan (2016) conducted another similar study to examine the effect of critical pedagogy on Iranian EFL learners' motivation. The findings, as previous study proved, revealed that applying critical pedagogy led to an increase in learners' motivation. In bilingual settings, Hones (2002) examined the impact of critical pedagogy, especially the value of a dialogical process, on the bilingual students in the USA. He concluded that dialogic pedagogy would benefit young bilingual students' academic and social aspects.

Despite research on critical pedagogy, few studies have been conducted in this area. In Iran, no comparative research has been conducted between monolinguals and bilinguals. In this regard, the present study has sought to compare the effectiveness of Freire's critical pedagogy among Iranian monolingual and bilingual EFL learners. Certainly one of the effects of critical education is the establishment of educational justice among all learners, regardless of linguistic, economic, and social discrimination. As one of the primary tasks of any researcher and teacher is responsibility for his/her audience, the researcher in the present study with many years of teaching experience in both bilingual and monolingual settings has come to this conclusion that education in such environments faces some problems. One of these problems is the lack of educational justice, especially in bilingual environments, which can be partially addressed by providing scientific and practical solutions and by introducing critical education so that EFL teachers can achieve better results in educational justice in the future.

Applying CP, especially Freire's critical pedagogy in both monolingual and bilingual settings, is very effective. This may lead to the conclusion that the minority group may favor this method much more than the other group and show much more progress in their career. It is worth mentioning that English is the foreign language for both monolingual and bilingual groups and Persian is the native language for the monolingual group and the second/official language for the bilingual group whose native language is Azeri (Turkish).

The main purpose of the present study is to examine the extent to how CP may function in EFL teaching and learning in Iran. The majority, high privileged and monolingual group who live in the Capital is compared with the minority, less privileged and bilingual group who live in a non-capital city. Using the critical pedagogy, the researcher aims to find which group reacts better to CP.

Research Questions

The following research questions were addressed and investigated in this study:

1. Is there any statistical difference between problem-posing education and banking education of Iranian EFL monolingual speaking performance?
2. Is there any statistical difference between problem-posing education and banking education of Iranian EFL bilingual speaking performance?
3. Is there any statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of banking education?
4. Is there any statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of problem-posing education?

Methodology

Participants' (Subject) Characteristics

The participants were all male and female Teaching English as Foreign Language (TEFL) sophomores and their ages ranged from 20 to 24 (with mean age of 22). The number of participants was the same in both groups (thirty students in the bilingual group and thirty in the monolingual group). Sixty percent was female and the remaining forty percent was male. The students were at the same proficiency level in that they had been exposed to English Lab Course 1 (listening and speaking course 1) as prerequisite to English Lab Course 2 (listening and speaking course 2). They were all at the elementary level of English.

The teacher for both groups were different but with identical demographic characteristics. They were both male bilingual teachers, got Ph. D. in TEFL, and aged 38-40. They had 17-19 years' experience in teaching EFL in university and had good command of English. More importantly, they had enough knowledge, interest, and familiarity with the subject under study.

Instruments

Textbook

The researchers introduced Top Notch 1 by Saslow and Ascher (2015). It is a dynamic communicative course which covers eight top goals and achievement-based lessons. The textbook was designed to take beginner university students in English as a second/foreign language context. It builds confidence for successful oral expression through an expanded speaking pedagogy and prepares students for academic work through the development of critical thinking skills and reading and listening strategies. The point is that the textbooks were the same for both classes, but the teacher's teaching methods were different. In the first group, the teacher played a more active role and encouraged the students to participate more, which made the class more interesting for the students, but in the second class the teacher did not play a very active role and the students were not very interested in the participation. This made the class boring (for more information about the banking and problem-posing classes, see Table 1).

Learning Approaches (Banking Model vs. Problem-posing Pedagogy)

As it was discussed earlier, in banking education the teacher merely transfers the knowledge but in problem-posing education, real-life aspects are presented to the students through dialogues and their political and personal developments are supported. The characteristics of Freire's (1970) banking and problem-posing education are summarized in Table 1.

Table 1
Characteristics of Freire's (1970) Banking and Problem-posing Education

	Banking Education	Problem-posing Education
Aims	Depositmaking	Increasing critical consciousness
Role of Teacher	Lecturer, mere teacher, and problem solver	Problem poser and coordinator of discussions and dialogues.
	Teacher is indifferent to the political and especially sociocultural climate/ context	Teacher is interested and sensitive to the political and especially sociocultural context
Role of students	Mere learner, they are passive	Teacher and students are learning together, they are active
Tasks and materials	They are boring and monotonous	They are motivating and engaging
Teacher education	Static	Dynamic
	Education as the exercise of domination	Education as the exercise of liberation and practice of freedom
		Deep and critical understanding
	Parrot like memorization	Pattern of education is in the form of horizontal relationship between teacher and students (dialoguers)
	Pattern of education is vertically established between teacher and students	

Pre-Test and Post-Test

The sample was distributed into four groups. In all four classes, the students received one sample Cambridge IELTS 9 speaking test as pre-test and post-test. Each speaking test included three parts. In part 1, the examiners asked the participants some general questions about themselves and a range of familiar topics, such as home, family, work, studies and interests. This part lasted four or five minutes. In part 2, the participants were given a card which asked them to talk about a particular topic. They had one minute to prepare before speaking for up to two minutes. The examiners then asked one or two questions on the same topic. In part 3, they were asked some further questions about the topic in Part 2. The questions gave the participants the opportunity to discuss more abstract ideas and issues. This part of the test lasted four or five minutes.

The textbook used in the present study, as it is mentioned earlier, is suitable for beginner university students in English as a second/foreign language context. The sample tests as pre-test and post-test are taken from IELTS sample test. As it is clear, IELTS is an international (standardized) test of English language proficiency for non-native language speakers and it is one of the major English language tests in the world. This is the major reason for the researcher's use of IELTS sample test as pre-test and post-test. Table 2 compares IELTS scores with the European scale of English levels.

Table 2

Comparison of IELTS Scores with the European Scale of English Levels

Levels	IELTS scores
Proficiency	8.0 - 9.0
Advanced	7.0 - 7.5
Upper-Intermediate	6.0 - 6.5
Intermediate	5.0 - 5.5
Pre-Intermediate	4.0 - 4.5
Elementary	3.5 and below

IELTS Speaking Band Descriptors (Public Version)

The purpose of this speaking grading criterion was to assess and score students' oral ability. This checklist included four major criteria of *fluency and coherence*, *pronunciation*, *lexical resource*, and *grammatical range and accuracy*. Each of the major sections was divided into nine band scores, and the raters chose one that corresponded to their understanding of examinees' knowledge and improvement in oral ability. The nine band scores as mentioned, ranged between band score 0 (does not attend) to band score 9 (full competent and advanced). All four criteria were judged throughout a three-part speaking test. The mentioned nine band scores describe the speaking performance as follows: Band 0, did not attempt the test; Band 1, non-user; Band 2, intermittent user; Band 3, extremely limited user; Band 4, limited user; Band 5, modest user; Band 6, competent user; Band 7, good user; Band 8, very good user; Band 9, expert user. The inter-rater reliability of IELTS band scores is calculated and was equal to 0.75.

Procedure

The present pre-experimental study is done under four groups. Two available classes in bilingual context, Islamic Azad University, Tabriz (17 and 18 respectively) and two other classes in monolingual context, Islamic Azad University, Tehran (19 and 19) are introduced as the research participants. In order to have equal number of students in all groups, those students who got the lowest score in the last term final exam English Lab Course 1 (listening and speaking course 1) were excluded. The research sample became 15 in each group.

In order to follow the principle of ethics, all participants, even those excluded from the research received the normal teaching/testing process. The researcher informed the students that they were participating in a study designed by the researcher and written informed consent was obtained from all participants. Also, in order to observe the ethical principles of the research and in order for the banking group not to be deprived of the benefits of the problem-posing intervention, the banking group was taught the problem-posing intervention in the form of delayed interventions a few sessions after the research.

Fourteen sessions were devoted to conduct the present study and to administer the treatment. The first experimental groups in both monolingual and bilingual contexts received problem-posing model of teaching and the second groups were exposed to banking model.

Banking model is like the traditional autocratic methods. Following Freire's (1970) banking model, the teacher was more active and the processor of knowledge and the students were just passive recipients of knowledge. No emphasis was given on learners' capabilities, interest and needs. The teacher acted as depositors and students patiently received, memorized and repeated. The teacher dictated and selected the content and did the teacher correction process but the students did not have any role in selecting the content and evaluation process. Another key point in this model was that the teacher was the subject and the students were the objects of the education and learning process. In short, the students did not find any evidence of critical thinking from the surrounding context

In problem-posing group, the teacher initially assigned several sessions to work on the students' spiritual, emotional, and cultural dimensions, and provided the necessary context for the students to realize that good thinking requires better learning. The teacher also encouraged the students not to be afraid of asking questions. The teaching model was more of a humanistic model in which students could use social interactions to construct their own environment. To apply this model, a dialogic approach was used to increase the critical thinking power of learners. The teacher used the reflection-action cycle to change the learners' personal assumptions, actions, and the world around them. Instead of having the role of knowledge transmitter, he played the role of the problem poser. He encouraged the students to think by asking a series of critical questions. By challenging them, he tried to make the students understand the obvious reasons for those obvious issues. Also, to answer all the students' questions logically and correctly, he tried to be aware of all the social, political and cultural issues of the society and played the role of a learner. He even tried to learn new things by interacting with individual students One of the necessary tools for critical pedagogy was the use of dialogue, and by making these dialogues even briefly, the teacher tried to create constructive student-student and student-teacher interactions. Brainstorming was used during dialogues in class. Needs analysis was also used to address the students' actual interests and needs so that they could be more motivated and interested in learning. Another tool used in the classroom was applying dynamic assessment, which is itself a teaching model. Also, the teacher used alternatives in assessment tools such as portfolio and dialogue journals. This method first helped to create a free and democratic atmosphere that both facilitated the learning process and created an atmosphere where students could think, ask questions, and do all the class work freely. Another important point was the use of self-assessment method. It was an empowering tool that helped the learners assess their learning and reflect on their own actions as a group.

The scoring procedure of participants' performance in pre-test and post-tests writing was based on IELTS speaking band descriptors. The participants' performance was recorded by two raters and scored during pre and post-test. The inter-rater reliability between the two raters was 75 percent. The results were finally analyzed using Statistical Package for Social Sciences (SPSS) software, version 23.

Research Design

The present study is a pre-experimental study with one group pre-post design (no control group). The independent variable was different models of pedagogy (problem-posing as a model of CP and banking education as a dominant uncritical system of education) and dependent variable included speaking performance of EFL learners in monolingual and bilingual contexts.

In order to have equal number of students in all four groups, those students who got the lowest score in last term final exam were excluded. In order to analyze all independent groups with quantitative variables and to compare the effect of critical pedagogy on EFL speaking performance of monolingual and bilingual learners, an independent sample t-test was used.

Data Analysis

The main concern of the present study was to examine the effect of Freire's critical pedagogy on Iranian EFL bilingual and monolinguals' speaking performance. This study attempted to find out if there was any difference between Freire's problem-posing and banking education. Furthermore, this study aimed to explore the probable statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of problem-posing education.

Descriptive Statistics for Hypotheses

In order to analyze the scores in all four groups, an independent t-test was utilized which was preceded by descriptive statistics to give a comprehensive manifestation of the results. The results of descriptive statistics for the participants' performance in all four groups are shown in Table 3.

Table 3

Descriptive Statistics for all four Hypotheses

	Group	N	Mean	Std. Deviation	Std. Error Meas.
(POST-TEST) H1. statistical difference between problem- posing education and banking education in monolingual group	PmG	15	3.2000	.56061	.14475
	BmG	15	2.0667	.25820	.06667
(POST-TEST) H2. statistical difference between problem- posing education and banking education in bilingual group	PbG	15	3.3333	.61721	.15936
	BbG	15	2.1333	.51640	.13333
(POST-TEST) H3. statistical difference between monolingual and bilingual groups in terms of banking education	BmG	15	2.0667	.25820	.06667
	BbG	15	2.1333	.51640	.13333
(POST-TEST) H4. statistical difference between monolingual and bilingual speaking performance in terms of problem-posing education	PmG	15	3.2000	.56061	.14475
	PbG	15	3.3333	.61721	.15936

Note: PmG: Problem-posing (education) in monolingual Group; BmG: Banking (education) in monolingual Group; PbG: Problem-posing (education) in bilingual Group; BbG: Banking (education) in bilingual Group

Table 3 illustrates the descriptive statistics of four hypotheses. The mean score and standard deviation of both problem-posing and banking groups are given respectively for all four hypotheses. Based on the mean scores in problem-posing and banking education classes in all four groups, the researcher believes that problem-posing instruction results in higher score than banking education in both hypotheses 1 and 2 but it does not result in higher score in hypotheses 3 and 4.

The Assumption of Normality of Scores

The purpose of using this statistical analysis is to determine whether the distribution of scores has normality. To this end, in each case, the significance value was determined and the results were shown in Table 4.

Table 4

Kolmogorov-Smirnov Test for Normal Distribution of Scores in all Four Groups

		Monolingual Group		Bilingual Group	
		Problem- Posing	Banking	Problem- Posing	Banking
N		15	15	15	15
Normal Parameters[a]	Mean	3.2000	2.0667	3.3333	2.1333
	Std. Deviation	.56061	.25820	.61721	.51640
Most Extreme Differences	Absolute	.211	.328	.241	.282
	Positive	.166	.323	.225	.244
	Negative	-.211	-.328	-.241	-.282
Kolmogorov-Smirnov Z		1.138	1.326	1.442	1.161
Asymp. Sig. (2-tailed)		.328	.203	.158	.301

[a] Test distribution is Normal.

According to the results which are shown in Table 4, the significance value in each case is greater than critical value ($p > 0.05$) and z-scores are less than 1.96 which indicate the normality of the scores in all four groups prior to applying t-test.

Statistical Analysis for Hypothesis 1.

There is a statistical difference between problem-posing education and banking education of Iranian EFL monolingual speaking performance. The researcher used a t-test to provide statistical evidence of whether the difference between the two groups' performance in post-test scores is significant. In other words, to see whether the difference between problem- posing and banking educations' means of post-test scores is statistically significant or not, an independent sample t-test is used.

As Table 5 shows, there is a significant difference between the post-test scores in both problem-posing and banking groups, t=7.112, p = .008 (*p < 0.05), that is, there is a significant difference between two sets of means as the obtained p-value is less than critical value. Consequently, problem-posing education outperformed banking group in monolingual setting and it had significant effect on Iranian EFL learners' speaking performance. In other words, the directional hypothesis of the study (There is a statistical difference between problem-posing education and banking education of Iranian EFL onolingual speaking performance) is confirmed.

Table 5

Independent Sample t-test for both Problem-posing and Banking Groups' Post-tests in Monolingual Setting

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
H1	Equal variances assumed	8.092	.008	7.112	28	.000	1.13333	.15936	.80689	1.45978
	Equal variances not assumed			7.112	19.684	.000	1.13333	.15936	.80056	1.46610

Statistical Analysis for Hypothesis 2.

There is a statistical difference between problem-posing education and banking education of Iranian EFL bilingual speaking performance.

In order to provide statistical evidence of whether the difference between the two groups' performance in post-test scores is significant, another independent sample t-test is used. In other words, to check whether the difference between problem-posing and banking educations' means of post-test scores in Iranian bilingual setting is statistically significant or not, an independent sample t-test is used.

As Table 6 illustrates, there is a significant statistical difference between the post-test scores in both problem-posing and banking groups, t=5.775, p = .045 (*p < 0.05), that is, there is a significant difference between two sets of means as the obtained p-value is less than critical value. Consequently, problem-posing education, as hypothesis one, outperformed banking group in bilingual setting and it had significant effect on Iranian EFL learners' speaking performance. In other words, the second directional hypothesis of the study (some statistical difference between problem-posing education and banking education of Iranian EFL bilingual speaking performance) is supported.

Table 6

Independent Sample t-test for both Problem-posing and Banking Groups' Post-tests in Monolingual Setting

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
H2	Equal variances assumed	4.389	.045	5.775	28	.000	1.20000	.20778	.77437	1.62563
	Equal variances not assumed			5.775	27.154	.000	1.20000	.20778	.77377	1.62623

Statistical Analysis for Hypothesis 3:
There is a statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of banking education.

The third independent sample t-test is used to provide statistical evidence of whether the difference between the two groups' performance in post-test scores is significant. In other words, to see whether the difference between post-test scores of two banking groups in Iranian monolingual and bilingual settings is statistically significant or not, an independent sample t-test is applied.

Table 7 indicates that there is not any significant difference between the two different banking groups' post-tests in both monolingual and bilingual settings, t=-.447, p = .344 (p > 0.05), that is, there is not any significant difference between two sets of means as the obtained p-value is more than critical value. Consequently, banking education in both bilingual and monolingual settings does not differ. In other words, the third directional hypothesis of the study (these is a statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of banking education) is rejected.

Table 7
Independent Sample t-test for two Different Banking Groups' Post-tests in both Monolingual and Bilingual Settings

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
H2	Equal variances assumed	.928	.344	-.447	28	.658	-.06667	.14907	-.37203	.23869
	Equal variances not assumed			-.447	20.588	.659	-.06667	.14907	-.37706	.24372

Statistical Analysis for Hypothesis 4.
There is a statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of problem-posing education.

In order to provide statistical evidence of whether the difference between the two groups' performance in post-test scores is significant, another independent sample t-test is used. In other words, to check to see whether the difference between two problem-posing groups' means of post-test scores in Iranian monolingual and bilingual settings is statistically significant or not, an independent sample t-test is used.

Table 8 illustrates that there is not any significant difference between the two different problem-posing groups' post-tests in both monolingual and bilingual settings t= -.619, p = .358 (p > 0.05), that is, there is not any significant difference between two sets of means as the obtained p-value is more than critical value. Consequently, problem-posing education in both bilingual and monolingual settings does not differ. In other words, the fourth directional hypothesis of the study (there is a statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of problem-posing education) is rejected.

Table 8
Independent Sample t-test for two Different Problem-posing Groups' Post-tests in both Monolingual and Bilingual Settings

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
H2	Equal variances assumed	.873	.358	-.619	28	.541	-.13333	.21529	-.57433	.30767
	Equal variances not assumed			-.619	27.745	.541	-.13333	.21529	-.57451	.30785

Discussion

The purpose of this study was to explore the application of Freire's critical pedagogy in Iranian EFL bilingual and monolingual speaking performance. To address the objective of the research, a pre-experimental study was set out. The findings indicated that there is no significant difference between Freire's problem-posing and banking education in Iranian monolingual and bilingual settings.

Pennycook (1990b) stated that CP is informed by critical social theory. It is aimed to critique the historical and sociopolitical context of schooling and is directed to change the wider society in general and the nature of schooling in particular. Lack of enough freedom and justice in education in some societies may lead to creation of a bipolar and oppressed society, which as Ramos (2005) stated, includes two major groups of oppressor and oppressed. In order to overcome the problem of injustice, especially educational injustice, Freire (1970) introduced the problem-solving method as an alternative to the banking education for undemocratic societies. As mentioned earlier, in the banking model the teacher tries to transfer mere knowledge to the students but in the problem-posing education the teacher tries to transfer knowledge through real-life and meaningful situations on the one hand and to support the personal and political development of the students on the other. In Iranian context, as similar to other bilingual and multilingual societies, it is sometimes stated by some that the dominant monolingual group possesses some privileges due to power and facility they own in the capital while the minority bilingual group lacks the fair power and privileges. However, controversies over the effectiveness of Freire's (1970) critical pedagogy especially on monolingual and bilingual contexts urged the researcher to scrutinize two different models of critical pedagogy, banking and problem-posing, among Iranian EFL monolingual and bilingual EFL learners' speaking ability. To solve the oppressor and oppressed positions, an experimental study was done among monolinguals and bilinguals to know if they react to CP differently.

Regarding the first research hypothesis, that is, the statistical difference between problem-posing education and banking education of Iranian EFL monolingual speaking performance, the results of an independent sample t-test showed that problem-posing education outperformed banking group in monolingual setting and it had significant effect on Iranian EFL learners' speaking performance.

Concerning the second research hypothesis, that is, the statistical difference between problem-posing education and banking education of Iranian EFL bilingual speaking performance, the results of another independent sample t-test indicated that problem-posing education outperformed banking group in bilingual setting and it had significant effect on Iranian EFL learners' speaking performance.

Regarding the third research hypothesis, that is, the statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of banking education, the results of another independent sample t-test demonstrated that banking education in both bilingual and monolingual settings does not differ.

The fourth research hypothesis proposed a statistical difference between Iranian EFL monolingual and bilingual speaking performance in terms of problem-posing education. To achive this goal, a final independent sample t-test was used and demonstrated that problem-posing education in both bilingual and monolingual settings does not significantly differ.

The aim of the present study, as it is mentioned earlier, was to investigate the effect of Freire's critical pedagogy on Iranian EFL bilingual and monolingual speaking performance. Of course, the reason for comparing these two groups is not because of their language types, but as previously explained in the introduction, it is related to the concepts of monolingualism and bilingualism, the existence of different social and educational advantages of the city in which these two groups lived and it also seems to be related to the different expectations that society has of these two groups. In order to be more specific, Freire's problem-posing and banking models were compared in different monolingual and bilingual contexts. The analysis of data revealed that applying problem-posing model caused improvement in speaking performance of both monolingual and bilingual learners. That is, problem-posing education outperformed banking group in both monolingual and bilingual settings and it had significant effect on Iranian EFL learners' speaking performance. The finding of this study is in line with different empirical studies which have done to investigate the effect of critical pedagogy, especially problem-posing model, on EFL learners' skill performances. For example, the finding of the present study is in line with the findings of (Norton & Toohey, 2004; Pennycook, 1999; 2001). They stated that applying critical pedagogy affects the general process and outcomes of English language learning and teaching. Also, the findings are in line with Joseph Jeyaraj (2020). He conducted a qualitative study using interviews with Malaysian students to explore their perceptions of critical pedagogy and social justice. He concluded that despite some limitations, it seems that there is a good desire for this model in this country. Furthermore, the finding of the present study is in line with the findings of Brown (2013) who concluded that increased depth of student learning was one out of the three major opportunities towards using problem-posing pedagogy. The finding of the present study also supports the findings of Hong and White (2012) who concluded that critical pedagogy could empower student writers and increases their awareness. Finally, the findings are in line with Manca et al. (2020) who came to the conclusion that CP facilitates social power, critical consciousness, and fosters empathy and compassion among medical staff.

Hypotheses three and four deal with the probable differences between Iranian monolingual and bilingual EFL learners in terms of banking and problem-posing models. The findings revealed that there is no significant difference between monolingual and bilingual EFL learners neither in terms of banking nor problem-posing models. The findings of this part of research are not in line with Hones (2002). He examined the impact of critical pedagogy, especially the value of a dialogical process on the bilingual students in the USA. He concluded that dialogic pedagogy would benefit young bilingual students' academic and social aspects. The findings of the present study also reject Moorhouse (2014). He examined the effect of critical pedagogy with young EFL learners in Hong Kong primary school and came to this conclusion that applying CP may empower students, especially those marginalized by society (bilingual or multilingual learners).

Finally, the findings of the present study, specially hypotheses three and four, do not support the findings of Benites (2012, as cited in Kamali Sarvestani & Yamini, 2016) who concluded that applying CP in ESL and EFL settings leads to better and successful bilingual learners.

Conclusion

Speaking is probably one of the most significant skills in the world of information where people feel a growing need for international and intercultural communication. This makes learners enthusiastically attend language classes to improve their speaking ability. Freire's critical pedagogy is an approach which might help learners overcome the challenges that learners face in EFL context.

There are two explanations regarding the practical applications of CP in Iranian context. The first is concerned with priority of the problem-posing model of CP over the banking model in Iranian bilingual and monolingual contexts. The second refers to the different impact of the problem-posing model of CP in Iranian bilingual and monolingual contexts. Since this method is primarily aimed at addressing educational, cultural, and environmental injustices, it has been mistakenly assumed that bilingual environments in Iran are more affected by this method and respond better to it because of some shortcomings, especially in terms of educational and environmental opportunities. But the findings of the third and fourth hypotheses indicated that neither the banking nor the problem-posing were significantly different in Iranian monolingual and bilingual settings. Therefore, it is recommended to apply problem-posing model of CP in both Iranian bilingual and monolingual contexts. As the main concept of critical pedagogy suggests that the application of this pedagogy will cause users to change their attitudes toward free and critical thinking, the results of this research are likely to lead Iranian students think and act democratically, critically, and actively.

Last but not least, the findings should be interpreted in light of some limitations. First, the small sample size decreased the generalizability of the present study. Another limitation is the availability of sample which did not allow random selection of participants. Finally, the participants of the study were chosen from only EFL university students. Therefore, further studies can be done using random sampling techniques and with a larger number of participants. Also, other researches can be done with the participation of students from other disciplines.

Findings of the present study can have pedagogical implications for EFL teachers and syllabus designers.

The teachers should know how to apply critical pedagogy and distinguish between banking and problem-posing models. They should propose problem-posing as an alternative to the banking model. This alternative model, as Freire (1970) stated, uncovers reality and leads to critical consciousness. This consciousness lets students take the required actions to improve their life conditions and motivate students to have the right to ask questions. Also, in order to resolve the oppressor-oppressed contradiction, the teacher should introduce critical pedagogy as an alternative critical intervention. In short, freedom and justice which are sometimes overlooked by societies are motivated by introducing CP in general and problem- posing education in specific.

Teachers and students expect the curriculum to help them teach, learn, and communicate freely. This type of curriculum should provide a democratic context in which teachers feel relaxed to take changes and do risk taking activities. Also, syllabus designers should provide adequate resources which motivate problem-posing model in particular and critical pedagogy in general.

One of the main goals of education is to bring all people together and provide the necessary ground for access to human knowledge, which is one of the main goals of CP in both theory and practice. One solution that can be suggested to those involved in education and policy makers is to provide online learning, of course, if the tools and facilities of this technology are easily available to everyone. Souranta (2020) believes that this type of learning is in fact a new stage of Freire's critical model and can be considered "a globally distributed pedagogy of all people in the process of permanent liberation" (1129). In this type of learning, everyone can learn, participate, be creative, and evaluate each other's learning process.

Finally, critical pedagogy should improve both teachers and learners' points of view and enable them to develop their speaking performance by exposing them to authentic materials and highlighting the real life problems and situations.

Acknowledgments

First, I express my thanks to the Islamic Azad University, Khoy branch, Iran, for its support in carrying out this study.

I am grateful to the EFL departments in Islamic Azad University, Iran, for giving me the time and opportunity to conduct the study.

My special thanks go to the EFL students and teachers who took part in this study. Without them this study would not have been possible. Their eager participation and cooperation were essential in carrying out this study.

References

Abdollahzadeh, H., & Haddad Narafshan, M. (2016). The practicality of critical pedagogy: A case study of Iranian EFL learners' motivation. *Studies in English Language Teaching*, *4*, 196-207. http://dx.doi.org. 10.22158/selt.v4n2p196

Auerbach, E.R. (1995). The politics of the ESL classroom: Issues of power in pedagogical choices. In Tollefson (Ed). *Power and inequality in language education* (pp. 9-33). Cambridge University Press.

Baladi, N. (2007). *Critical pedagogy in the ELT industry: Can a socially responsible curriculum find its place in a corporate culture?* [Unpublished master's thesis]. McGill University. http://digitool.library.mcgill.ca/thesisfile18457.pdf

Barjesteh, H., Alipour, B., & Vaseghi, P. (2013). Critical pedagogy: Improving Iranian EFL learners' reading comprehension ability through CP strategies. *Indian Journal of Fundamental and Applied Life Sciences*, 3, 335-341. http://www.cibtech.org/jls.htm

Bartolome, L. (1994). Beyond the methods fetish: Toward a humanizing pedagogy. *Harvard Educational Review*, 64, 173-194. http://dx.doi.org.doi: 10.17763/haer.64.2.58q5m5744t325730

Brown, P. M. (2013). *An examination of Freire's notion of problem-posing pedagogy: The experiences of three middle school teachers implementing theory into practice* [Unpublished doctoral thesis]. The university of Georgia. https://getd.libs.uga.edu/pdfs/brown_philip_m_201308_phd.pdf

Campo, J. E. (2009). *Encyclopedia of Islam* (2nd ed.). Facts on File.

Correa, M. (2015). Flipping the foreign language classroom and critical pedagogies: A (New) old trend. *Higher Education for the Future*, 2, 114-125. http://dx.doi.org.10.1177/2347631115584122

Esteva, G., Stuchul, D. L., & Prakash, M. S. (2005). From a pedagogy for liberation to liberation from pedagogy. In C. A. Browers & F. Apffel-Marglin (Eds.), *Rethinking Freire: Globalization and the environmental crisis* (pp. 13-30). LEA Publishers. http://dx.doi.org. 10.4324/9781410611741

Freire, P. (1970). *Pedagogy of the oppressed*. Continuum.

Freire, P. (1997). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Rowman & Littlefield.

Hanafiyeh, M., Afghari, A., & Koosha, M. (2019). The role of HEXACO personality traits in predicting the speaking ability of male and female EFL learners. *Iranian Journal of Learning and Memory*, *2*, 1-7. http://dx.doi.org.10.22034/iepa.2019.90174

Haque, E. (2007). Critical pedagogy in English for academic purposes and the possibility for tactics of resistance. *Pedagogy, Culture & Society*, 15(1), 83-106. http://dx.doi.org.10.1080/14681360601162311

Hones, D. F. (2002). In quest of freedom: Towards critical pedagogy in the education of bilingual youth. *Teachers College Record Volume*, *104*(6), 1163-1186. http://dx.doi.org.10.1111/1467-9620.00200

Hong, Y., & White, J. (2012). Writing without fear: Creativity, and critical pedagogy in Chinese EFL writing programs. *Journal of Asian Critical Education*, *1*, 8-20. https://www.researchgate.net/publication/236885673_The_Journal_of_Asian_Critical_Education

Joseph Jeyaraj, J. (2020). Possibilities for critical pedagogy engagement in higher education: Exploring students' openness and acceptance. *Asia Pacific Education Review*, *21*(1), 27-38. http://dx.doi.org. 10.1007/s12564-019-09605-0

Kalsoom, S., Kalsoom, N., & Mallick, R. J. (2020). From banking model to critical pedagogy: Challenges and constraints in the university classrooms. *UMT Education Review*, *3*(1), 25-44. http://dx.doi.org.10.32350/user.31.02

Kamali Sarvestani, N., & Yamini, M. (2016). Iranian EFL teachers' attitudes towards critical pedagogy. *International Journal of English Linguistics*, *6*(1), 110-117. http://dx.doi.org. 10.5539/ijel.v6n1p110

Manca, A., Gormley, G.J., Johnston, J.L., & Hart, N. D. (2020). Honoring medicine's social contract: A scoping review of critical consciousness in medical education. *Academic Medicine*, *95*(6), 958-967. http://dx.doi.org.10.1097/ACM.0000000000003059

Mitsikopoulou, B. (2020). Genre instruction and critical literacy in teacher education: Features of a critical foreign language pedagogy in a university curriculum. *L2 Journal*, *12*(2), 94-109. http://dx.doi.org.10.5070/L212245951

Moorhouse, B. L. (2014). Using critical pedagogies with young EFL learners in a Hong Kong primary school. *International Journal of Bilingual and Multilingual Teachers of English*, *12*, 79-90. http://dx.doi.org.10.12785/IJBMTE/020201

Norton, B., & Toohey, K. (2004). Critical pedagogies and language learning: An introduction. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 1-17). Cambridge University Press.

Pennycook, A. (1990a). Critical pedagogy and second language education. *System*, *18*, 303-314. http://dx.doi.org.10.1016/0346-251X(90)90003-N

Pennycook, A. (1990b). Towards a critical applied linguistics for the 1990s. *Issues in Applied Linguistics*, *1*, 8-28.

Pennycook, A. (1999). Introduction: Critical approaches to TESOL. *TESOL Quarterly*, *33*, 329-348. http://dx.doi.rg.10.2307/3587668

Pennycook, A. (2001). *Critical applied linguistics: A critical introduction*. Lawrence Erlbaum Associates.

Perumal, J. (2016). Enacting critical pedagogy in an emerging South African democracy: Narratives of pleasure and pain. *Education and Urban Society*, *48*, 743-766. http://dx.doi.org.10.1177/0013124514541466

Ramos, M.B. (2005). *Paulo Freire: Pedagogy of the oppressed*. Continuum.

Shahraki-Sanavi, F., Nasseri, N., Shahraki-Sanavi, F., & Salehiniya, H. (2019) Investigating the observance of educational justice in Iranian universities: From students' viewpoints. *Journal of Education and Health Promotion*, 8, 1-4. http://dx.doi.org.10.4103/jehp.jehp_456_18

Steven, CH. (2002). *Critical pedagogy on the web: Key terms and concepts related to critical pedagogy and educational theory and practice*. http://mingo.info- science.uiowa.edu/~stevens/critped/terms.htm

Saslow, J., & Ascher, A. (2015). *Top notch 1* (3 rd ed.). Pearson.

Suoranta, J. (2020). Critical pedagogy and Wikilearning. In S. Steinberg & B. Down (Eds.), *The SAGE handbook of* *critical pedagogies* (pp. 1126–1138). Sage. http://dx.doi.org. 10.4135/9781526486455.n102

Van Canh, L. (2018). A critical analysis of moral values in Vietnam-produced EFL textbooks for upper secondary schools. In H. P. Widodo, M. R. Perfecto, L. Van Canh & A. Buripakdi (Eds.), *Situating moral and cultural values in ELT materials* (vol. 9, pp. 111–129). Springer. http://dx.doi.org. 10.1007/978-3-319-63677-1_7

Yulianto, S. W. Y. (2020). The use of critical pedagogy principles in teaching English as a foreign language (EFL) for senior high school students in Subang. *Biormatika: Jurnal ilmiah fakultas keguruan dan ilmu pendidikan*, *6*(1), 98-106. http://dx.doi.org.10.35569/biormatika.v6i1.698

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Lavrova, N., & Nikulina, E. (2020). Advanced Russian EFL Learners' Awareness of Idiomatic Synonymy, Antonymy, and Polysemy. *Journal of Language and Education, 6*(4), 105-120. https://doi.org/10.17323/jle.2020.9689

Advanced Russian EFL Learners' Awareness of Idiomatic Synonymy, Antonymy, and Polysemy

Nataliya Lavrova, Elena Nikulina
Moscow State Pedagogical University

Correspondence concerning this article should be addressed to Nataliya Lavrova, Department of Foreign Languages, Phonetics and Lexicology, Moscow State Pedagogical University, 1/1 M. Pirogovskaya Str.,Moscow, 119991, Russian Federation. E-mail: na.lavrova@mpgu.su

Foreign language acquisition is notoriously constrained by learners' lack of awareness of the systemic relations that are obtained among stable multiple-unit lexical items. This results in learners' inability to variegate their performance (both written and oral) with idioms that stand in complementary (synonymy) or contrastive (antonymy) distribution to one another. Nor are learners typically able to distinguish between the multiple senses of English idioms. Given these impedimenta, the present research investigates the degree of entrenchment of idiomatic synonymy, antonymy, and polysemy and, on the back of it, sets the agenda for partial revision of the practice of exposing learners to English idioms. Data were collected to investigate the knowledge of idiomatic synonymy, antonymy, and polysemy amongst Russian EFL learners. The results of the ANOVA analysis revealed that the degree of awareness of the major types of idiomatic paradigmatic relations significantly differed between the groups, with learners being more aware of synonymy and polysemy than antonymy. The findings suggest that current EFL materials and dictionaries need to be updated and revisited with a view to exposing foreign learners to an extended network of paradigmatic idiomatic relations.

Keywords: idioms, paradigmatic relations, synonymy, antonymy, polysemy

Introduction

'*Speak idiomatically unless there is special reason not to*'
J. Searle (1993)

According to the most conservative estimates (Deignan, 2005; Siyanova-Chanturia, 2017; Zykova, 2016), every fifth uttered expression is idiomatic, a finding which suggests that speakers heavily rely on prefabricated items stored in long-term memory.

The investigation of systemic relations among English idioms is an under-researched area; most scholars are interested in separate or more specific aspects of idioms, such as polysemy (Baranov & Dobrovol'skij, 2014[1]; Moon, 1998), synonymy (Dobrovol'skij & Piirainen, 2018), speed of comprehension (Gibbs, 1986; Glucksberg, 2001; McGlone et al., 1994), variation (Langlotz, 2006), style (Altenberg, 1998), the role of idioms in second and foreign language acquisition (Conklin & Carrol, 2019; Türker, 2016), specific subclasses of idioms (Mel'čuk, 2015), functional idiomaticity (Pamies-Bertrán, 2017), etc. Previous findings suggest that learners associate one form of a set expression with one meaning (Nikulina, 2015), have difficulty paraphrasing a set expression idiomatically (Macis & Schmitt, 2017), and can rarely come up with an antonymic periphrasis to a set expression (Ivanova, 2017)[2]. To the best of our awareness, there has been no systemic or comprehensive research into a network of ramified systemic relations that exist among idioms and into how well non-native speakers of English are aware of these relations.

[1] Baranov, A. N., & Dobrovol'skij, D.O. (2014). Ocnovui frazeologii [Foundation course in English phraseology]. Flinta.
[2] Ivanova, E. V. (2017). Leksikologoya i frazeologiya sovremennogo anglijskogo yazuika [English lexicology and phraseology]. Academiya.

Given the existing gap, the main aim of the present research is to establish how well Russian learners of English are aware of idiomatic synonymy, antonymy, and polysemy. The three research questions are thus how well are Russian learners aware of (1) idiomatic synonymy, (2) antonymy, and (3) polysemy?

Idiomatic Polysemy, Synonymy, and Antonymy

Research on idiomatic polysemy either reveals the speakers' lack of awareness or treatment of the semantically connected meanings of an idiom as separate linguistic items (Baranov & Dobrovol'skij, 2014). Alternatively, some speakers rely on an idiom's underlying image as its primary meaning, misinterpreting the secondary idiomatic meaning altogether or deducing it directly from its literal counterpart. This could be explained by the default strategy most speakers utilise when faced with an unfamiliar expression. As most of the secondary meanings of a polysemous idiom seem to be closely connected, speakers tend to collapse two or three meanings together rather than teasing them apart. Compared to other researchers, who never mention idioms' functions as a possible cause of their polysemy, Fedulenkova argues that it is idiomatic pragmatic-functional malleability that seems to be one of the reasons why idioms develop secondary meanings (Fedulenkova, 2019). Contrasted with research on idiomatic polysemy, there are more papers devoted to research on idiomatic synonymy. This is because both FLT writers and language instructors are directly involved in the ways of enhancing the teaching of formulaic language. Research reveals that the dearth of synonyms among idioms is mostly due to the psychological salience of the composite idiomatic image and subsequent associations that arise out of the literal meaning of idiomatic components (Dobrovol'skij & Piirainen, 2018; Teliya, 1998). Since this image never peters out, absolute synonymy is non-existent among English idioms. Structurally and semantically isomorphic synonymic idioms are more readily acquired by learners and are easier to access than idiomatic synonyms whose structure and underlying image are drastically different from each other (Türker, 2016).

There is still inconclusive evidence as to the differences between the so-called idiomatic synonyms and idiomatic variants and whether there is any difference in their processing, with scholars making rather arbitrary distinctions between the two categories (Cserép, 2017; Langlotz, 2016). Finally, there is an appreciable dearth of research on idiomatic antonymy, with most scholars registering learners' lack of awareness or observing that proverbs and sayings are more prone to have antonyms than non-predicative idiomatic expressions (Dronov[3], 2015; Ivanova, 2017).

Idioms in Second Language Acquisition and Cross-Language Influence

Research on the role of idioms in second language acquisition is burgeoning, as awareness of idiomatic paradigmatic relations is definitely of benefit to second language learners for at least two reasons. First, the ability to use idiomatic synonyms is one of the main foci of international examinations: they put a premium on learners' extended idiomatic vocabulary. Recognition of idiomatic polysemy is also tested in the Reading and Use of English part of CAE (Certificate in Advanced English) and CPE (Certificate of Proficiency in English). Given this, raising learners' awareness of a set of paradigmatic relations among English idioms should be the primary focus of advanced and proficient second language learning and teaching (Choonkyong, 2015). Second, most research demonstrates an impact of the L1 on the interpretation of set expressions in an L2 (Conklin & Carrol, 2019). It seems that second-language congruence is a factor to be reckoned with when processing different structural types of idioms. Apparently, the more congruent idioms from L1 and L2 are, the less time that is spent on interpretation and less cognitive effort required from learners to identify the meaning of an idiom correctly. The degree of accurate interpretation thus correlates with the degree of cross-linguistic congruence of set expressions (see Nguyen & Webb, 2017; Wolter & Gyllstad, 2011; Wolter & Yamashita, 2015, 2018). The findings also suggest that while interpreting the meanings of synonymous idioms, non-native speakers rely more heavily on the underlying image of set expressions, which frequently leads them astray due to the known gap between the literal and figurative meanings of an idiom (Wray, Bell, & Jones, 2016).

The study by Tavakoli and Uchihara (2020) is a step forward in the examination of the relationship between oral fluency and use of multiword sequences across four proficiency levels. The obtained data show that idiomatic frequency correlates with a higher and quicker articulation rate. Lower-proficiency speakers use idioms in contexts that mostly fit idioms' underlying image, while higher-proficiency test-takers use idioms more competently in a variety of contexts, most of which fit the idioms' current figurative meaning. The

[3] Dronov, P.S. (2015). Obschaya leksikologiya [General lexicology]. Yazuiki slav'anskoj kul'turui.

authors' research is valuable in that it provides conclusive evidence of the correlation between the degree of learners' proficiency and their ability to recognize and efficiently use English idioms.

The data obtained by Katsyuba, Ismailova, & Bondareva (2020) have a direct bearing on our research in that they demonstrate that there is a two-way link between idioms and language proficiency; not only does the learners' degree of proficiency increase with the number of internalized idioms, they are also much better at interpreting idiomatic synonyms and antonyms if their degree of proficiency is relatively high. This means that it makes sense to enlist the help of proficient or nearly proficient learners of English in experimental research, which we do in our research, in order to obtain valid data.

Idioms, Language Change, and Natural Language Processing

The findings by Buerki (2019) are well-nigh groundbreaking in that he demonstrates that changes in idiomatic language proceed apace and that learners need to be aware of idiomatic synonymy and antonymy, since new idioms, which are the synonyms or antonyms of current ones, appear on a regular basis, especially in colloquial language. This testifies to the need for a more optimal way of exposing learners to formulaic language.

Apparently, multiword expressions pose challenges to Natural Language Processing (Inurrieta et al., 2020). Verbal idioms seem to be harder to identify and interpret. They are also less likely to be associated with idiomatic synonyms and antonyms and to be recognized by learners as polysemous. Verbal idioms are thus predicted to be interpreted with a greater margin of error than other types. This implies that when selecting idioms for interpretation, it won't do to have an abundance of verbal idioms at the expense of other types.

Dictionary Practice

Analysis of dictionary practices demonstrates that if and when idiomatic synonyms and antonyms are featured in a dictionary entry, for some reason, they either fall under the radar of learners, who seem surprised when their attention is drawn to their presence (Arsenteva & Ayupova, 2013), or else dictionaries simply do not enter cross-references to idiomatic synonyms and antonyms (Perevezentseva, 2015). Although polysemy is featured slightly more frequently, learners seem to pick up only one meaning from a dictionary entry, the one that is either first or seems to be more familiar to them, probably, by virtue of its higher frequency of occurrence.

Given this paucity of idiomatic relations in most dictionaries, there is a need for specific guidelines on how to register different types of idiomatic relations so that learners are aware of their existence in the first place. Lončar and Valero (2020) believe that online digital tools endowed with various search modalities could expose learners of English to a ramified network of paradigmatic relations among idioms. Without doubt, an online dictionary is a digital tool that can provide access to set expressions with which the target idiom is somehow connected. This, we believe, will result in learners' better awareness of idiomatic synonymy, antonymy, and polysemy.

Dal Maso's proposal (2020) is thus to put forward a possible strategy for the lemmatization of lexical variation and synonymy in dictionaries of idioms. The data can be taken on board by English dictionary compilers, who should develop a means of recording idiomatic variation and synonymy. This is crucial, because learners frequently confuse the two categories and rarely think of them as semantically connected.

The brief analysis of research on various aspects of systemic relations among English idioms has demonstrated an appreciable lack of empirical research into the extent to which learners are aware of these relations. There are several possible reasons for this, such as suboptimal teaching and dictionary practice, the vivid image at the basis of idioms, and the lack of conclusive evidence about which type of idiomatic relation is entrenched the most. Drawing on some of the preliminary results obtained by the above-mentioned authors and using their findings as the point of departure, we intend to present a more integrated picture of systemic paradigmatic relations among English idioms.

Research Hypothesis

The working hypothesis for this research was that synonymic relations would be easiest for learners to come up with; partly due to the current practice of teaching idioms (synonyms are mentioned more often than antonyms), and partly because learners themselves typically aim to variegate their speech through the use of synonyms, rather than antonyms, which are not associated with learners who have a better command of English. As far as polysemy is concerned, we hypothesized that learners would be more aware of polysemy than antonymy, since a close, pre-research analysis of various dictionaries of idioms revealed that the lexico-semantic idiomatic variants are typically closely connected to each other, and the second or, possibly, the third meaning can be figured out provided that the first meaning is understood correctly.

Methodology

Participants

The participants that took part in the present research are 50 Russian speakers of English of comparable age, socio-economic status, and educational level: students from Moscow State Pedagogical University, aged 20-22, whose level of English is no lower than C1 in accordance with the CEFR.

Materials

The final sample used in the experimental research comprises 104 idioms (one idiom was used in two different tasks) selected through the method of random sampling from the following dictionaries of idioms: The American Heritage Dictionary of Idioms (2003), The Oxford Dictionary of Idioms (2004), McGraw-Hill Dictionary of American Idioms and Phrasal Verbs (2006), Longman Dictionary of Contemporary English (2009), The Farlex Idioms and Slang Dictionary (2017), and Idioms and Phrasal verbs (2011).

Given the numerous classifications and subdivisions of idioms in relevant outlets, only the most prototypical idioms were selected; collocations, formulae, proverbs and sayings, and grammatical idioms were excluded from the sample. Collocations are semi-set word combinations in which one of the words is used in its direct meaning: cf. *overwhelming majority, mental acuity, academic prowess*, etc. Formulae are set expressions used as speech acts in their own right: *That figures! You don't say so*. Finally, grammatical idioms are those that either consist of only functional parts of speech or contain only one non-functional word: *in accordance with*, *in comparison with*, *on balance*, etc. None of these are easily associated in modern scholarship with idioms or else they have qualifications like *grammatical*, *discursive*, etc. The same goes for proverbs and sayings, which are structurally different from lexical idioms in that they are predicative units. Hence, these were also excluded from the final sample. Some of the idioms selected for Task 1 also have idiomatic antonyms, which is a predictable result of random sampling. However, we were not interested in whether the selected synonyms had antonyms and therefore did not control for this factor. The selected idioms also vary in frequency of usage, which is again something that random sampling can account for. The idiom *at the eleventh hour*, for example, occurred 105 times in COCA compared to the idiom *the icing on the cake*, which occurred 289 times. However, most of the dictionaries from which idioms were selected relied on frequency bands and adhered to the practice of including idioms that are relatively frequent in modern English. Most of the idioms selected for Task 3 have two or three meanings. This is because most polysemous idioms tend to have two meanings, with 4 or 5 occurring extremely rarely. Given all of the above, the idioms selected for all the tasks were *relatively* homogeneous, with the exception of some outliers such as *at the end of one's tether*, which is infrequent in the corpora. Finally, we believe that learners are likely to come across a range of idioms that differ on several parameters, including register (something we did not control for). Since an educated learner of English may regularly see or hear idioms from a spectrum of registers, frequencies, lengths, etc., it is our genuine belief that it is necessary for them to be aware of their meanings and of the systemic relations in which they enter with other idioms. In a nutshell, it was not part of the research objective to control for a multiplicity of parameters on which idioms may vary, but to demonstrate learners' relatively low awareness of idiomatic synonymy, antonymy, and polysemy, and to spell out the reasons why it is necessary to enhance this knowledge.

Procedure and Data Analysis

Regarding the steps of dividing the treatment into three groups, those idioms were selected that contain cross references, i.e. references to possible synonyms or antonyms, indicated by one of the strategies by dictionary compilers, which were outlined in the literature review section. As for polysemous idioms, those whose number of meanings exceeded one were selected. Idioms with technical meanings (cf. *on an even keel*) or idioms with very closely related meanings were excluded from the final sample. The former group was excluded on the grounds that terms and idioms are different entities; the second group was sidelined because, as was also mentioned in the literature review section, the split of meanings does not seem to be justified from the point of view of the storage of linguistic items in the mental lexicon, which means that students are unlikely to tease two almost identical meanings apart (cf. *between the devil and the deep blue sea* 1) 'caught in a dilemma'. 2) 'trapped between two equally dangerous situations'). Finally, idioms whose first meaning seemed to be non-metaphorical were also excluded (cf. *from the dead* 1) 'from the state of death'. 2) 'from a period of obscurity or inactivity').

All of the participants were presented with three lists of 35 idioms each. This relatively small number of idioms for each task was justified by the cognitive load placed on participants: a greater number would result in much greater mental exertion being required and would prohibitively prolong the experiment, whose results might thus have been compromised. This number also seems to be the mean number of linguistic items typically suggested for interpretation in an experimental format. Learners were given 130 minutes to complete all three tasks. Each idiomatic expression was thus given slightly more than a minute, which is enough time to come up with the target item if the learner is aware of its existence. In Task 1, learners were asked to paraphrase given idioms with at least one synonymic set expression. To facilitate the completion of the task, one word from the synonymic idiom was mentioned with the rider that in some cases only one component of the idiom may be changed, in others a synonymic idiom had a completely different structure. At this stage, we deliberately did not differentiate between idiomatic synonyms and variants, since this is a theoretical question with which most students are unlikely to be familiar. Eight idioms contained only one word that needed to be paraphrased and were thus regarded as variants, while the rest were idiomatic synonyms. In Task 2 learners were required to come up with at least one antonym to each idiom. To facilitate the completion of the task, one word from the antonymic idiom was mentioned. It was also explained to the participants that in some cases only one component of the idiom may be changed, in others antonymic idioms are lexically completely different. In Task 3, potentially polysemous idioms were presented for interpretation, and learners had to come up with at least one more meaning of each idiom. In two tasks (Task 1 and Task 2), one prompt word was given after each idiom in order to facilitate the completion of the task and to 'nudge' students in the right direction. This 'nudging', however, should not be regarded as an invalidation of the whole procedure: similar or comparable tasks in proficiency courses are always provided with examples and some prompting word(s) from the target expression.

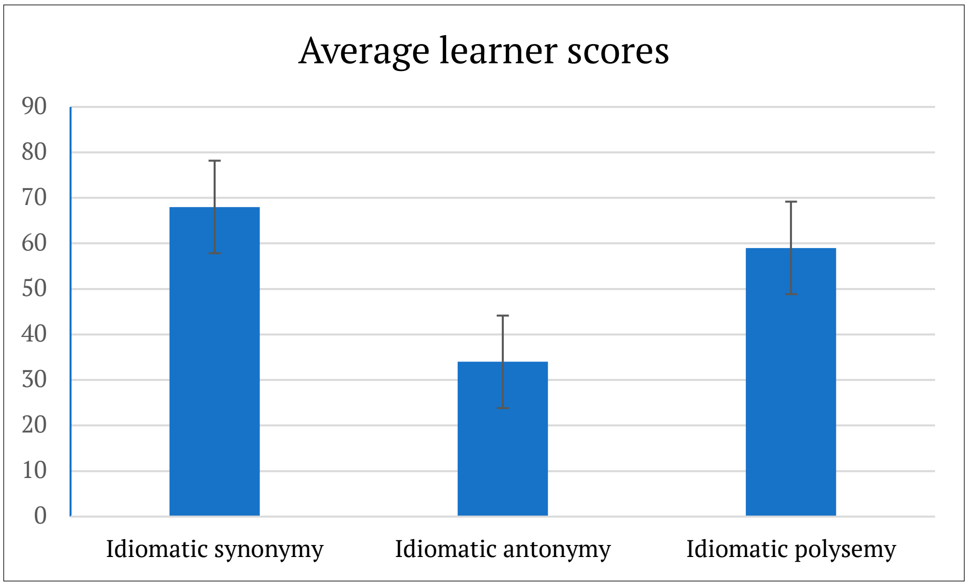
Apart from an explanation of the task, an example of its possible completion was provided. This detailed presentation of the task is in accordance with the standards required by the CAE and CPE examinations. An unstructured, post-hoc interview was conducted with participants to find out what difficulties they had when completing the tasks.

The ANOVA analysis was used to figure out the within-group variance, the between-group variance and to compare all three groups, which comprise synonymic idioms (analysed in Task 1), antonymic idioms (analysed in Task 2), and polysemous idioms (offered for analysis in Task 3). To establish the different degrees of awareness of the three types of paradigmatic idiomatic relations, the averages scores on each task were also calculated.

Results

As we anticipated, learners demonstrated the least awareness of idiomatic antonymy; polysemy came second, while awareness of idiomatic synonymy was the highest. Figure 1 illustrates the average learner scores on all the three types of learner awareness and the standard deviations.

Figure 1
Average learner scores on all the three tasks (in percentage) and standard deviations



As shown in Table 1, the difference between groups was significant ($F = 3.772$; $p = .035$). Given the degrees of freedom (2 and 30) and the probability level of 0.05, the critical F-value was 3.31582950. The null hypothesis can thus be rejected, since F > F critical, i.e. the means of the populations are not all equal.

Table 1
Results of One-way Statistical Analysis of the Respondents' Answers by Groups of Idioms

Source	SS	df	MS	
Between-treatments	39.8182	2	19.9091	F = 3.77153
Within-treatments	158.3636	30	5.2788	
Total	198.1818	32		

Note. The p-value is 0.034584. The result is significant at $^*p < .05$

The ANOVA analysis was followed up with post-hoc comparisons to see where the critical differences were, i.e. each task was compared to establish which ones were significantly different from each other. We conducted Tukey's Honestly Significant Difference Post-Hoc Test and came up with the value of HSD being 1.44. Having compared the means in each treatment, we established that the honestly significant difference was meaningful (2.46>1.44) between the groups of synonyms (mean = 4.09, SD = 2.8) and antonyms (mean = 1.63, SD = 2.1) as well as the groups of synonyms and polysemous idioms (mean = 1.9, SD = 1.86), their difference being 2.19>1.44. However, the difference between students' awareness of antonyms and polysemous idioms turned out to be statistically insignificant (1.9-1.63=0.27), with HSD = 1.44.

The analysis was done by items, i.e. we were primarily interested in how many idioms from each group were interpreted correctly. Although potentially possible, partly or 'semi-correct' answers were conspicuous by their absence. Prior to the experiment, we had anticipated this problem. For example, how would we mark an answer which was only partially correct? (cf. *to let the cat out of the *sack*). For such answers, we were ready to give students half a point. Only one lexical word and, possibly, one grammatical word, such as an article, were allowed to be produced incorrectly to merit half a point. However, as it turned out, such answers were statistically dismissible: only two were given, in response to item №6 in Task 1 (**to throw the gasket*) and to item 15 in Task 2 (**to take years of smb.*). Potentially, the production of more than one secondary meaning of a polysemous idiom could earn participants one extra point, i.e. the number of points was directly linked with the number of correctly supplied meanings: if a learner had supplied two more meanings of a polysemous idiom, two points would have been earned. However, not a single participant gave more than one extra meaning of a polysemous idiom. The marking was done by both the authors, and inter-rater reliability was ensured by analysing all the answers first separately, then together, and by additionally consulting each other on potentially problematic cases.

In Task 1, most of the subjects failed to supply a periphrastic synonymic alternative to the following idioms: *to have egg on your face, to carry a torch for smn., at the end of your tether, at the eleventh hour, thick and fast*. Among some of the idioms that were paraphrased correctly by most of the participants were *the icing on the cake, a weight off your shoulders*, *a bundle of nerves*, and *to laugh one's head off*. Among antonymic idioms, the following ones caused a great deal of difficulty: *to be going strong*, *to spare one's feelings*, *a load of rubbish*, and *of your own free will*. In contrast, only a small number of learners failed to supply an antonym for the idioms *to put years on smn.*, *to be under an obligation to smn.*, *to keep your options open*, and *to have money to burn.* Finally, most learners failed to supply a second or secondary meaning for the following idioms in Task 3: *On your bike!*, *on the button*, *to put smn. out of their misery*, *to come of age*, *with your eyes shut*, and *to cut both ways*. In Task 3, few subjects failed to supply a second meaning to the idioms *to blow smn. away*, *to have nothing on smn.*, and *to be on the run* (see Table 2).

Discussion

Possible Causes of Difficulty in Interpreting Some of the Idioms

The idioms that have been paraphrased most adequately and without effort were those that are congruent: cf. *the icing on the cake* (R. *vishenka na torte*)*, a weight off your shoulders* (R. *gruz s plech*), *a bundle of nerves* (*R. komok nervov*). The findings that congruent items were processed more quickly and more correctly are in line with the studies reported in Conklin and Carrol (2019).

The participants misinterpreted some of the idioms as they may not have been aware of their meanings in Task 1 and Task 2. Apparently, in order to be able to come up with a synonymous or an antonymic idiom, learners should be aware of the original meanings of the idioms. Although maximal objectivity was ensured due to the bias-free selection of the three subsamples, some of the resultative idioms may have caused learners' greater or lesser interpretative difficulty. The reasons for failing to supply a synonymous alternative to some of the idioms in Task 1 may have been quite diverse, ranging from rarely used words in the idiomatic structure, such as *tether*, through a somewhat obscure metaphor (cf. *to carry a torch for smn*.) to Biblical allusions, such as *at the eleventh hour*, which requires the activation of learners' background knowledge and their general awareness of cultural practice, precedents, and allusions. Although many native speakers may not have the background knowledge and still know the meaning of an intertextual idiom, background knowledge is usually a facilitating factor for *non-native* speakers. Since many idiomatic allusions are traced back to a common source, such as the Bible, knowledge of the original text (the so-called 'prototext') in which a particular idiom was first mentioned helps to decode and interpret the meaning of an unfamiliar idiom in a foreign language. However, the parable about the 'eleventh hour' is unfamiliar to most non-religious Russian speakers, nor is there a comparable expression in Russian, hence, its form may have seemed somewhat obscure.

Failure to come up with an antonymic alternative to some of the idioms in Task 2 may have been caused by the key words suggested for paraphrasing, such as *duress* and *doldrums*, for instance, which are rarely, if ever, used outside their idiomatic expressions. Finally, the subjects' failure to suggest another meaning for idioms in Task 3 may have been due to the fact that some of these idioms are either rarely used in their second meaning or have two or three closely related meanings. Still others contain what are known as 'enantiosemic semantic components', i.e. diametrically opposed or poorly compatible meanings within idioms' semantic structures (cf. *to be downhill all the way* 1. 'be easy in comparison with what came before', 2. 'become worse or less successful').

Research Question 1. How Well are Russian Learners Aware of Idiomatic Synonymy?

Awareness of idiomatic synonymy was as high as 68%. More synonymic idioms seemed to be recognized when they clustered around a specific thematic field and thus expressed a salient concept. This pertains to idioms that deal with *madness/obsession*, *anger*, and *inebriation.* Apart from the idioms tested in Task 1, for illustrative purposes Table 2 also contains idioms that reflect some of the most salient negative concepts in modern English.

Table 2
The Most Frequent Concepts that Engender Idiomatic Synonyms

Frequent concepts expressed by English idioms	
Madness/obsession	nutty as a fruitcake, not play with a full deck, not have both oars in the water, lights are on but nobody's home, (the) elevator doesn't go to the top floor, to be half a bubble of plumb, to have bats in the belfry, to have a bee in one's bonnet, etc.
Anger	to fly off the handle, to flip one's wig, to blow a fuse, to blow one's top, to have a conniption, to have a hissy, to blow a gasket, to throw a fit, to go haywire, go ballistic, to fly into a rage (a rampage, a rant), etc.
Inebriation	drunk as a skunk/as a lord, three sheets to the wind, rolling/screeching/raring drunk, etc.

Although we did not set out to differentiate between synonymic and variable idioms, the results of the research revealed that synonymic idioms were interpreted more easily than idiomatic variants, i.e. set expressions that differ in one or two components (cf. *the icing/cherry on the cake*). This is probably because most Russian learners are exposed to such practice of English teaching when variability in the structure of stable multi-word units is not welcome or considered to be impeding the process of second language acquisition. In the Russian tradition of education, for instance, learners are typically asked to come up with a no-alternative idiomatic expression, and variants are typically penalized by teachers, who genuinely believe that these may puzzle and mislead students, who have substantial difficulty in mastering the form and usage of one specific idiom.

Despite the relative paucity of idiomatic synonymy *per se*, idiomatic synonyms turned out to be relatively easy to come up with for a number or reasons: first, idiomatic synonyms are more frequently included in current EFL course-books; second, it is common practice among language instructors to variegate their speech using alternative expressions, as this is strongly associated with vocabulary expansion.

Research Question 2. How Well Are Russian Learners Aware of Idiomatic Antonymy?
Idiomatic antonymy turned out to be the least entrenched type of paradigmatic relation among English idioms (34%), a finding which is seemingly in contradiction with the data obtained by Ivanova (2017). However, this may partly be explained by the subject-matter of Ivanova's research, which mostly focused on proverbs and sayings, i.e. predicative idioms that differ on a number of parameters from non-predicative stable multi-word units. The learners' low awareness of idiomatic antonymy may also be explained by the prevalent teaching practice through which speech variability is attained by learning a number of alternative synonymic expressions rather than contrastive words and meanings. Speech variability is also frequently tested via international examinations, such as the CAE and CPE, hence most proficiency textbooks are focused on periphrastic (synonymic) set expressions rather than antonymic ones. That said, it probably makes sense to also activate learners' antonymic links with idioms, not least because, however indirectly, doing so may also contribute to speech variability, since a synonymic meaning may frequently be expressed via negation of its opposite: cf. *to take a turn for the better – not to take a turn for the worse, to call a spade a spade – not to beat around the bush.*

Research Question 3. How Well Are Russian Learners Aware of Idiomatic Polysemy?
The research revealed that polysemy among idiomatic expressions was not as regular or productive as synonymy (59%). This finding is in accordance with the data obtained by Boyarskaya and Zabotkina (2017), Dobrovol'skij and Piirainen (2018), and Ivanova (2017). The authors indicated that idioms' inherently evaluative and expressive nature, as well their extended lexical make-up prevented them from developing a ramified network of paradigmatic relations. This finding also supports Nikulina's research (2015), which showed that native speakers failed to recognise a set-expression's dual or polysemous nature. Apparently, advanced and proficient non-native speakers also tend to associate one form of a set expression with only one meaning. Whether this finding can be extrapolated to speakers of languages other than Russian remains to be seen, although we have a strong suspicion that this might be the case.

The lower productivity of idiomatic polysemy may partly be explained by the activation of an idiom's literal meaning in the process of its comprehension and interpretation. Although this activation does not necessarily mislead speakers and learners of English if *they are familiar* with the idiom's figurative meaning, if the meaning seems to be *unfamiliar*, speakers primarily rely on the literal meaning to figure out the idiom's possible figurative meaning, presupposing that it is the only one and tends to be closely connected with the underlying literal image. Apart from potentially taking extra time and thus preventing learners from finishing the task on time, this reliance on the summative meaning derived directly from an idiom's constituent elements is far from an optimal strategy because of the well-known law in the development of linguistic signs, according to which, over time, meaning and form tend to diverge, i.e. there is no reliable or predictable connection between the exponent of a linguistic sign and its current meaning (in synchrony). This finding aligns with research by Cieślicka (2006), who advanced the proposition that literal meanings of idiom constituents have processing priority over their figurative interpretations. This suggestion forms the core of the literal-salience resonant model of L2 idiom comprehension.

The commitment of modern dictionary compilers to split rather than lump together idiomatic meanings may partly explain learners' relative awareness of idiomatic polysemy, compared to antonymy. That said, the post-hoc interviews with the participants revealed that they guessed rather than knew the second meaning of a polysemous idiom. While guessing is a useful strategy for figuring out polysemous words' and expressions' meanings (Boyarskaya & Zabotkina, 2017), this may still result in a significant margin of error and may lead to appreciable communicative mistakes, which is inadmissible if much is at stake, not least the interpreter's reputation and/or integrity. Given this, we believe that a more rigorous methodical tool for singling out idioms' meanings should be proposed, such as the 'cognitive trace' criterion, discussed in Glucksberg (2001), or the context of usage, which may be definitive in teasing idiomatic meanings apart (Naciscione, 2010). Idioms whose second meaning was far removed from the first, such as *to have a go at*, *to bring smb. out of their misery, to be downhill all the way*, etc. were hardly ever interpreted correctly. This is despite the fact that these idioms are relatively common in corpora.

Another reason for learners' being less aware of idiomatic polysemy than synonymy is that the meanings of lexico-semantic variants of a polysemous idiom differ in several aspects. First, one of the meanings is naturally primary, while the other is secondary. This seemingly trivial observation accounts for differences in the frequency, usage, and psychological salience of different meanings of an idiomatic expression. Second, many idioms' first meanings are derived from a technical sense of a set expression, i.e. a term or profession jargon which is registered as such in some idiomatic dictionaries. However, learners seem to store the etymologically linked meanings of a set expression separately, and only one meaning seems to be activated for learners in interpretative contexts, either terminological or metaphorical. This finding is in accordance with the data obtained by Nikulina (2015), whose experiment revealed that most learners are not aware of the second meaning of the English idiom *to bring home the bacon*, being only familiar with the more current meaning 'to earn enough money to provide for the maintenance of a household'. This is because the second meaning – 'to be successful in sports' – is more technical and restricted.

Given that some of the recurrent idioms, such as *to bring home the bacon, to dig deep, to put one's foot down, fair and square*, etc. are polysemous, we believe that the lack of awareness of idiomatic polysemy may lead to embarrassing communicative breakdowns, especially for those learners who work or intend to work as simultaneous translators or interpreters. An awareness of polysemy and the ability to differentiate between two distinct meanings of an idiom is crucial when these two meanings are antonymic, for example, in the case of the idiom *to be downhill all the way*, whose meanings are mutually exclusive: 1) 'to be easy in comparison with what went before', 2) 'to become worse or less successful'.

Some of the appreciable research limitations are that we did not control for as many parameters as potentially were possible when researching the degree of awareness of idiomatic synonymy, antonymy, and polysemy. This may partly explain why the difference between learners' awareness of idiomatic antonymy and polysemy turned out to be statistically insignificant. Register, frequency, the number of meanings of polysemous idioms, and the potential presence of idiomatic variants as well as idiomatic synonyms are some of the factors that could have an impact on the number of correctly/incorrectly interpreted idioms. However, this was neither the aim nor the objective of the present research, which is only the first step in what we hope could be a set of experiments aimed at analysing not only the degree of learners' awareness, but also possible causes of the suboptimal knowledge of various aspects connected with formulaic language.

Conclusion

The initial hypothesis that Russian learners of English are not equally aware of idiomatic synonymy, antonymy, and polysemy was validated, with idiomatic synonymy being most entrenched, while idiomatic antonymy was the least so. The difference between Russian learners' awareness of the three types of relations between English idioms proved to be statistically significant. Congruent idioms were interpreted more quickly and with a less margin of error in all the three tasks. The activation of an idiom's literal meaning is a factor that prevents Russian learners from interpreting its meaning correctly in the first place and, hence, diminishes the chances of accomplishing the task correctly and on time. The low awareness of idiomatic antonymy could be explained by an idiom's unique underlying image, which never peters out and which impedes learners in finding a diametrically opposed match for a particular set expression. The inequality of the meanings of a polysemous idiom on such parameters as frequency of usage, degree of semantic connection, and distance between an idiom's direct and figurative meaning have a bearing on Russian learners' awareness of idiomatic polysemy. Last but not least, enantiosemic meanings of a polysemous idiom are never produced by learners, because this contradicts learners' expectations that one and the same set expression can have incompatible meanings.

With certain qualifications, the obtained data testify to rather tenuous links between paradigmatically connected idioms in Russian learners' mental lexicon, which is partly explained by the suboptimal methodology of teaching formulaic English to FL learners. Since idiomatic idioms are paid more attention than idiomatic antonyms or polysemous idioms, it is only explicable that Russian learners of English are more aware of idiomatic synonyms than other types of relations. Polysemous idioms either fall under the radar of learners or are experienced as monosemous items. This is because idiomatic meanings are either too closely connected or because one of the meanings is infrequent or terminological.

In light of what has been established in this experimental research, it is our genuine belief that the current practice of presenting idioms in idiomatic dictionaries, reference sources, and textbooks should be revisited in the direction of a more explicit demonstration of the systemic paradigmatic elations that exist between English idioms. Not only will this raise FL learners' awareness of their existence, but will also paint a more veritable picture of formulaic language and of its complexity and interconnectedness.

There are, obviously, further directions in which research on relations between English idioms could develop. First, we believe that there is a need for the development of consistent and systematic criteria for distinguishing between idiomatic synonymy and idiomatic variation and that experimental research that could establish what is better retained by FL learners – idiomatic synonyms or variants – is called forth. Dictionary writers should distinguish more clearly and unambiguously between these two groups of idioms, as this might lead to their better retention by learners as well as to a more seamless and error-free usage of both variants and synonyms. It is also important to analyse the factors that make learners of English perceive and interpret some semantically related idioms as such while others seem to be stored in the mental lexicon as separate, disconnected items. Apparently, psycholinguistic experimentation, involving primed lexical decision tasks, is required to accomplish such a daunting task. Other types of idiomatic relations should also be systematically investigated, such as paronymy and blending, with the aim of clarifying whether these arise as a result of performance errors or errors that are due to an infirm grasp of the English language system. It would also be interesting to see whether similar or comparable results could be obtained with regard to paradigmatic idiomatic relations in other languages and whether the results of the present research are more relevant to typologically or etymologically related languages.

Conflict of interests

The authors declare that they have no conflict of interest.

References

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 101-122). Clarendon Press.

Arsenteva, E., & Ayupova, R. (2013). Bilingual and multilingual Russian-English phraseological dictionaries. *Yearbook of Phraseology*, *4*(1), 65-78. http://dx.doi.org/10.1515/phras-2013-0005

Boyarskaya, E. L., & Zabotkina, V. I. (2017). On the challenge of polysemy in contemporary cognitive research: what is conscious and what is unconscious. *Psychology in Russia: State of the Art*, *10*(3), 28-39. http://dx.doi.org/10.11621/pir.2017.0302

Buerki, A. (2019). Furiously Fast: On the speed of change in formulaic language. *Yearbook of Phraseology*, *10*(1), 5-38. https://doi.org/10.1515/phras-2019-0003

Choonkyong, K. (2015). L2 learners' recognition of unfamiliar idioms composed of familiar words. *Language Awareness: Achievements and Challenges*, *25*(1-2), 89-109. http://doi.org/10.1080/09658416.2015.1122025

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learner, *Second Language Research*, *22*(2), 115-144. https://doi.org/10.1191/0267658306sr263oa

Conklin, K., & Carrol, G. (2019). First language influence on the processing of formulaic language in a second Language. In A. Siyanova-Chanturia, & A. Pellicer-Sanches (Eds.), *Understanding formulaic language. A second-language acquisition perspective* (pp.62-77). Routledge. https://doi.org/10.4324/9781315206615-4

Cserép, A. (2017a). Idiom variation and decomposability Part I: Verbal variation. *Yearbook of Phraseology*, *8*(1), 95–122. http://dx.doi.org/10.1515/phras-2017-0006

Cserép, A. (2017b). Idiom variation and decomposability Part II: Variation in the noun phrase. *Yearbook of Phraseology*, *8*(1), 123–144. http://dx.doi.org/10.1515/phras-2017-0007

Dal Maso, E. (2020). Synonymy and lexical variation in Spanish and Italian phraseology: Proposal for an online bilingual bidirectional dictionary. *Circulo de Linguistica Aplicada a la Comunicacion*, *82*, 27-40. https://dx.doi.org/10.5209/clac.689561

Deignan, A. (2005). *Metaphor and Corpus Linguistics*. John Benjamins.

Dobrovol'skij, D., & Piirainen, E. (2018). Conventional figurative language theory and idiom motivation. *Yearbook of Phraseology*, *9*(1), 5-30. http://dx.doi.org/10.1515/phras-2018-0003

Fedulenkova, T. N. (2019). Pragmatic functions of modern English phraseology. Научный результат. Серия. Вопросы теоретической и прикладной лингвистики, *5*(2), 74-83.

Gibbs, R. W. (1986). Skating on thin ice. Literal meaning and understanding idioms in conversation. *Memory and Cognition*, *8*, 149-156. https://doi.org/10.1080/01638538609544629

Glucksberg, S. (2001). *Understanding figurative language*. Oxford University Press.

Inurrieta, U. E., Aduriz, I., de Ilarraza, A. D., Labaka, G., & Sarasola, K. (2020). Learning about phraseology from corpora: A linguistically motivated approach for multiword expression identification. *PLos ONE*, *15*(8), 1-18. https://doi.org/10.1371/journal.pone.0237767

Katsyuba, L., Ismailova, K., & Bondareva, O.V. (2020). Place of phraseology in the study of foreign languages. *XLinguae, 13*(2), 92-101. 10.18355/XL.2020.13.02.07

Langlotz, A. (2006). *Idiomatic creativity. A cognitive-linguistic model of idiom-representation and idiom-variation in English*. John Benjamins Publishing.

Lončar, I., & Valero, P. (2020). Proposals on the representation of phraseology in online dictionaries. *Circulo de Linguistica Aplicada a la Comunicacion, 82*, 9-26. https://dx.doi.org/10.5209/ clac.68960

Macis, M., & Schmitt, N. (2017). Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research*, *21*(3), 321-340. https://doi.org/10.1177/1362168816645957

Mel'čuk, I. (2015). Clichés, an understudied subclass of phrasemes. *Yearbook of Phraseology*, *6*(1), 55–86. https://doi.org/10.1515/phras-2015-0005

Moon, R. (1998a). *Fixed expressions and idioms in English*. Oxford University Press.

Moon, R. (1998b). Phraseology in written and spoken corpora. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 79-100). Clarendon Press.

Naciscione, A. (2010). *Stylistic use of phraseological units in discourse*. John Benjamins Publishing Company.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *21*(3), 298-320. https://doi.org/10.1177/1362168816639619

Nikulina, E. A. (2015). English Phraseology: Integration with terminology science. *Journal of Language and Education*, *1*(2), 41-45. https://doi.org/10.17323/2411-7390-2015-1-2-41-45

Nunberg, G., Sag, I. A., & Wasow, T. (1994). Idioms. *Language*, *70*(3), 491-538

Pamies-Bertrán, A. (2017). Grammatical metaphor and functional idiomaticity. *Yearbook of Phraseology*, *8*(1), 69–104. https://doi.org/10.1515/phras-2017-0005

Perevezentseva, Yu. S. (2015). Towards lexicographic description of idioms in the English language: Comparative analysis. *Journal of Siberian Federal University. Humanities & Social Sciences*, *11*(8), 2572-2581. http://dx.doi.org/10.17516/1997-1370-2015-8-11-2572-2581

Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching*. University of Oxford.

Searle, J. (1993). Metaphor. In A. Ortony (Ed.), *Metaphor and thought* (pp. 83-111). Cambridge University Press.

Siyanova-Chanturia, A. (2017). Researching the teaching and learning of multi-word expressions. *Language Teaching Research, 21*(3), 289-297. http://dx.doi.org/10.1177/1362168817706842

Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency. *Language Learning, 70*(2), 506-547. https://doi.org/10.1111/lang.12384

Teliya, V. N. et al. (1998). Phraseology as a language of culture: Its role in the representation of a collective mentality. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 55-75). Clarendon Press.

Türker, E. (2016). Idiom acquisition by second language learners: The influence of cross-linguistic similarity and context. *The Language Learning Journal*, *47*(2), 1-12. https://doi.org/10.1080/09571736.2016.1221441

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430-449. https://doi.org/10.1093/applin/amr011

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation. *Applied Psycholinguistics*, *36*(5), 1193-1221. https://doi.org/10.1017/S0142716414000113

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition*, *40*(2), 395-416. https://doi.org/10.1017/S0272263117000237

Wray, A., Bell H., & Jones, K. (2016). How native and non-native speakers of English interpret unfamiliar formulaic sequences. *European Journal of English Studies*, *20*(1), 47-63. https://doi.org/10.1080/13825577.2015.1136163

Zykova, I. V. (2016). Linguo-cultural studies of phraseologisms in Russia: Past and present. *Yearbook of Phraseology*, *7*(1), 127-148. https://doi.org/10.1515/phras-2016-0007

Appendix

The Three Tasks Offered to Learners to Establish the Degree of Entrenchment of the Main Types of Paradigmatic Relations among English Idioms

Task 1. Below is a list of 35 English idioms, all of which have synonyms that are also idioms. Look through the list and give at least one synonym to each idiom. In brackets, one word from the synonymic idiom is mentioned. Remember that in some cases only one component of the idiom may be changed, in others a synonymic idiom has a completely different structure.

Example:

not to turn a hair (eyelid)	syn. not to bat an eyelid
to go down the tubes (drain)	syn. to go down the drain
the icing on the cake (cherry)	
to spill the beans (cat)	
to get egg on your face (turn)	
to have money to burn (a hole)	
to carry a torch for smb. (heels)	
to blow a fuse (gasket)	
an unknown quantity (horse)	
flesh and blood (kin)	
at the end of your tether (wits)	
on the breadline (mouth)	
miles from anywhere (beyond)	
pull your socks up (act)	
at the eleventh hour (nick)	
to go bust (wall)	
a weight off your shoulders (mind)	
to wear yourself out (tire)	
to be a bundle of nerves (bag)	
to be locked in battle (loggerheads)	
to come to a standstill (grind)	
thick and fast (leaps)	
to get smb.'s back up (hackles)	
to make no odds (difference)	
to hold your cards close to your chest (play)	
to talk at cross purposes (wires)	
at a pinch (push)	
to be streets ahead of smb. (shoulders)	
to pull out all the stops (stone)	
to give the game away (cat)	
I wasn't born yesterday (pull)	
a close shave (call)	
to laugh your head off (stitches)	
to have a bee in one's bonnet (bats)	
under the influence (wind)	
to keep one's distance (berth)	
on edge (tenterhooks)	

Task 2. Below is a list of 35 English idioms, all of which have antonyms that are also idioms. Look through the list and give at least one antonym to each idiom. In brackets, one word from the antonymic idiom is mentioned. Remember that in some cases only one component of the idiom may be changed, in others two idioms are lexically completely different.

Example:

to keep trim (shape)	ant. to be out of shape
to take a turn for the better (worse)	ant. to take a turn for the worse
to put all your eggs in one basket (bets)	
to have money to burn (hand)	
to toss and turn (log)	
to take a turn for the better (worse)	
to have time on one's hands (bee)	
to call a spade a spade (bush)	
in the red (black)	
on the surface (down)	
on top of the world (dumps)	
to pour you heart out (bottle)	
keep you options open (eggs)	
to be going strong (doldrums)	
the high point of smth. (low)	
under an obligation to do smth. (no)	
to put years on smb. (take)	
to shape smth. out (phase)	
to nip smth. in the bud (blind)	
to stand up for smb. (turn)	
slow off the mark (quick)	
to bode well for smb. (bode)	
to tread carefully (headlong)	
to miss the point (hit)	
in dribs and drabs (bounds)	
to feel at home (fish)	
to the detriment of smb./smth. (benefit)	
to spare smb.'s feelings (mince)	
to be in one's element (depth)	
a turn up for the books (home)	
to push your luck (err)	
what a load of rubbish (figures)	
to set the seal on smth. (touch)	
of your own free will (duress)	
to do smth. by the book (corners)	
as bright as a button (brush)	
many hands make light work (cooks)	

Task 3. Below is a list of 35 English idioms, all of which have (at least) two meanings. Look through the list and give two meanings of each idiom.

Example:

0	under your belt	1. (of food or drink) consumed safely 2. satisfactorily achieved or acquired___________________
	to put smb. out of his/her misery	1. to kill 2. ____
	to be on the run	1 while running 2 ____
	to cast someone adrift	1 (of a boat or its passengers) floating without being either moored or steered 2 ____
	to come of age	1 (of a person) to reach adult status 2 ____
	under the banner of	1 claiming to support a particular cause or set of ideas 2 ____
	to have nothing on someone	1 to be inferior to smn. in a particular aspect 2 ____
	a bit on the side	1 a person with whom one is unfaithful to one's partner 2 ____
	in black and white	1 in writing or print 2 ____
	to blow someone away	1 to kill someone using a firearm 2 ____
	to bring home the bacon	1 to supply material support 2 ____
	to catch the sun	1 to be in a sunny position 2 ____
	a back-seat driver	1 a passenger in a car who gives the driver unwanted advice 2 ____
	On your bike!	1 (British, informal) to go away (used as an expression of annoyance) 2 ____
	close to the bone	1 (of a remark) penetrating and accurate to the point of causing discomfort 2 ____
	to burn the candle at both ends	1 to go to bed late and get up early 2 ____
	on the button	1 punctually 2 ____
	on the carpet	1 (of a topic or problem) under discussion 2 ____
	in the clear	1 no longer in danger or under suspicion 2 ____
	to knock smth. into a cocked hat	1 to put a definitive end to something 2 ____
	to go crackers	1 insane 2 ____
	to cut loose	1 to distance or free oneself from a person, group, or system 2 ____
	to dig deep	1 to give money or other resources generously 2 ____
	to come a cropper	1 to fall heavily 2 ____
	crowning glory	1 the best and most notable aspect of something 2 ____

	to cut both ways	1 (of a point or statement) to serve both sides of an argument 2 ______________________________________
	to be downhill all the way	1 be easy in comparison with what came before 2 ______________________________________
	with your eyes shut	1 with one's eyes shut (or closed) 2 ______________________________________
	fair and square	1 with absolute accuracy 2 ______________________________________
	to find one's feet	1 to stand up and become able to walk 2 ______________________________________
	to follow your nose	1 to trust one's instincts. 2 ______________________________________
	give and take	1 mutual concessions and compromises 2 ______________________________________
	to fall from grace	1 a loss of favour or a position of power or honour 2 ______________________________________
	to give up the ghost	1 to die 2 ______________________________________
	to have a go	1 to make an attempt 2 ______________________________________
	in the frame	1 to be (or not to be) eligible 2 ______________________________________

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Lin, W. Y., Ang, L. H., Chan, M. Y., & Paramasivam, S. (2020). Analysing Cultural Elements in L2 Mandarin Textbooks for Malaysian Learners. *Journal of Language and Education, 6*(4), 121-137. https://doi.org/10.17323/jle.2020.10332

Analysing Cultural Elements in L2 Mandarin Textbooks for Malaysian Learners

Wen Yue Lin, Lay Hoon Ang, Mei Yuit Chan, Shamala Paramasivam
Universiti Putra Malaysia

Correspondence concerning this article should be addressed to Lay Hoon Ang, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.
Email: hlang@upm.edu.my

Culture is an important aspect of foreign or second language education as the teaching of foreign languages straddles two languages, the learner's first language and the target/foreign language, and the different cultures associated with them. Textbooks for the teaching of foreign languages must inevitably orient to cultural elements from at least two cultural practices and environments. In this study, cultural elements in four Mandarin as a second language textbooks written for Malaysian learners were examined using content analysis. Cultural elements and categorisations of types of culture were employed to investigate the extent to which cultural elements (knowledge-culture or communicative-culture) and types of culture (source, target, international cultures, or intercultural interaction) were represented in these textbooks. The study revealed that both knowledge-culture and communicative-culture were present in the sampled textbooks. Communicative-culture was found to be explicitly represented in the textbooks. For knowledge-culture, however, there was an imbalance in the representation of the Big 'C' and small 'c' elements. Further, the sampled textbooks placed more emphasis on target culture, and feature localisation rather than internationalisation. Regarding diversity, there was only a slight representation of learners' diverse cultural backgrounds. The findings of the study provide insight into how cultural elements are incorporated into L2 Mandarin textbooks for Malaysian learners and may contribute to wider issues concerning the use of appropriate learning materials for developing cultural competence for L2 learners in a multicultural context.

Keywords: cultural elements, types of culture, Mandarin textbook, teaching Mandarin as a second language, multicultural context, materials development, applied linguistics

Introduction

Culture has become an important aspect of foreign/second language (FL/L2) instruction (Brown, 2000). Culture cannot simply be regarded as an 'expandable' fifth skill tacked on to the teaching of listening, speaking, writing, and reading (Kramsch, 1993). Instead, if language is seen as social practice, culture becomes the core of language teaching to the extent that cultural awareness needs to be considered for enabling language proficiency (Kramsch, 1993, p. 8). As described by Baker (2011, p. 65), cultural awareness is "a conscious understanding of the role culture plays in language learning and communication (in both a first and foreign language)". Besides that, teaching about culture can "enhance students' overall language learning experience" (Vernier et al., 2008, p. 268). It is necessary for learners to foster what Kumaravadivelu (2008) calls global cultural consciousness through which learners learn to interact effectively with people from different cultural backgrounds. FL teachers, therefore, need to pay attention to how culture is presented and imparted to learners in language classes.

The importance of culture in L2 teaching has also been acknowledged in Chinese as a foreign or second language teaching (TCFL or TCSL). As stated by Hanban (2014, p. 5) in the International Curriculum for Chinese Language Education report, the overall goal of TCSL is to equip learners with comprehensive language skills. Comprehensive language skills consist of four components: linguistic knowledge, linguistic skills, strategies, and cultural competence. Cultural competence covers cultural knowledge, cultural understanding, intercultural competence, and international perspectives. Cultural competence is seen to be an essential factor for enabling learners to use Chinese language more appropriately. This overall goal of TCSL stresses the significant role of culture in teaching and learning Chinese as a second language.

Textbooks in FL education, as Cortazzi and Jin (1999) assert, play the role of a teacher, a map, a resource, a trainer, an authority, a restrictor, and an ideology. Without a textbook, it is difficult for teachers to teach systematically (Ajideh & Panahi, 2016). Further, the textbook is an important input resource for learners to learn language and the cultural aspects of language use. In light of the importance of culture in FL instruction, there is a necessity to pay close attention to how culture is integrated in teaching materials such as textbooks.

Malaysia is a multi-cultural and multi-ethnic country that consists of Malays, Chinese, and Indians as the major ethnic groups, as well as the Dayaks (in the Borneo states) and various other minority groups. The Chinese language, or Mandarin, has been taught and learned as a second language since the 1960s in the country (Hoe, 2014a). All of the 20 Malaysian public universities offer Mandarin as a second language (L2 Mandarin) course for students (Hoe, 2014a; Ye, 2011). Based on statistics from 2011, there were 21,693 students enrolled in L2 Mandarin courses in each semester in Malaysian public universities (Hoe, 2014b, as cited in Chan et al., 2018, p. 67). However, there is no uniform teaching syllabus, assessment standard, or textbooks for the courses (Pan, 2011, as cited in Hoe, 2014a, p. 102). In Malaysian public universities, 63.2% of L2 Mandarin courses utilise textbooks that are written by Malaysian authors (Hoe, 2014a). The local L2 Mandarin textbooks are designed based on the authors' own teaching experience or adapted based on existing textbooks from China and Malaysia (Hoe, 2014a).

With regard to the local L2 Mandarin textbooks, most of the books focus mainly on developing learners' linguistic skills (Chan et al., 2018). Studies have shown that a range of locally grounded words and expressions were found in the textbooks, for example, 'Ma liu jia马六甲 (Melaka)', 'Yin du miao印度庙 (Hindu temple)', 'Basha巴刹 (market)', 'Ling ji令吉 (RM)', and 'Chong liang冲凉 (take a shower)' (Lam & Hoe, 2013; Yang, 2019; Ye & Wu, 2010). However, the majority of the textbooks still lack the integration of cultural information (Chan et al., 2018; Hoe, 2016, as cited in Chan et al., 2018). Cultural elements in past studies have been shown to be reflected mainly in illustrations (Ye & Wu, 2010) and lexis (Yang, 2019; Ye & Wu, 2010), while not much information can be found on the infusion of cultural elements in texts such as dialogues. Cultural elements in textbooks can be considered resources for cultural awareness-raising and reflection among FL learners (Tajeddin & Teimournezhad, 2014). Among past works on cultural elements in FL learning materials, Chan et al.'s (2018) study appears to be the only one that has attempted to give detailed information on the integration of cultural elements related to the culture of the target language and intercultural elements in textbooks. The authors conclude that there is a need for more cultural and intercultural elements to be integrated into local L2 Mandarin textbooks. However, as in all of the previous studies conducted on culture in L2 Mandarin textbooks, attention on cultural elements in the learners' own culture and international culture appear to be lacking. In the unique multicultural and multi-ethnic context of Malaysia, language learners' cultural backgrounds are not monolithic and including references to the learners' cultures poses a huge challenge for textbook writers. In Chan et al.'s (2018) study, the multicultural character of learners' in Malaysia's culturally diverse context was not addressed.

The past studies discussed above have shed some light on cultural elements that are integrated into local L2 Mandarin textbooks and highlighted the limitations of some textbooks in their role to foster intercultural understanding. However, insufficient attention and discussion has been paid to identifying cultural elements in textbooks that represent a diversity of learners' cultural backgrounds. The present study sought to contribute to a better understanding of cultural representations in L2 textbooks written for learners in multicultural societies. L2 Mandarin textbooks written for Malaysian learners provide the opportunity to discover how the incorporation of cultural elements of culturally diverse learners has been addressed by textbook writers.

Theoretical Perspectives

This section describes the theoretical perspectives that inform the study, including cultural elements in FL/L2 teaching, cultural elements in TCSL, and types of culture in FL/L2 textbooks.

The term 'second language' (SL/L2) has been increasingly used for all kinds of non-native language learning in recent decades (Stern, 1983). The ordinal 'second' refers to 'the chronology of language learning' and can be "any language acquired (or to be acquired) later than the native language" (Stern, 1983, p. 12). In the present study, Mandarin is learned as an 'L2' by speakers whose first language is not Mandarin in the Malaysian context.

Cultural Elements in FL/L2 Teaching

Various frameworks have been put forward for examining cultural elements in FL/L2 teaching. Prominent among them are the Big 'C' and small 'c' conceptions of culture. Chastain (1988), Paige et al. (1999), Pulverness (1995), and Tomalin and Stempleski (1993), conceptualised culture as having two aspects, which are Culture with a Big 'C' (or Capital 'C') and culture with a small 'c' (as cited in Dehbozorgi et al., 2014). Big 'C' Culture refers to a set of facts and statistics relating to the arts, geography, history, education, customs, and festivals of a target language society (Lee, 2009). These are obvious to language learners. Small 'c' culture refers to the invisible and hidden aspects of a target language culture, that is, the mainstream sociocultural values, beliefs, and norms (Lee, 2009). The framework of Big 'C' and small 'c' cultures has been emphasised in the teaching of foreign language curriculum in Europe and North America (Chan et al., 2018).

Cultural Elements in TCSL

Scholars have highlighted the importance of teaching cultural awareness in TCSL (see for example Chen, 1992, 1997; Li, 2015; Liu, 2014; Liu, 2018; Zhang, 1990). Two kinds of cultural elements are said to comprise culture: knowledge-culture and communicative-culture (Zhang, 1990).

Zhang (1990) classifies culture into knowledge-culture and communicative-culture from the functional perspective of culture in FL teaching. Knowledge-culture (Kc) refers to cultural background knowledge, which does not affect the understanding and use of certain words or sentences when two people from different cultural backgrounds communicate with each other (Zhang, 1990). It contains knowledge of the literature and art, movies, music, beliefs and values, history, and geography of a particular country. Communicative-culture (Cc) refers to the cultural knowledge that directly influences the accuracy and effectiveness of communication between people from different cultural backgrounds (Zhang, 1990). Lack of the cultural knowledge associated with certain words, sentences, or non-verbal behaviour may result in misunderstandings (Zhang, 1990). For instance, in Chinese names, the surname precedes the given name. L2 Chinese learners may confuse the given names and surnames of Chinese people if they do not know the order of Chinese names.

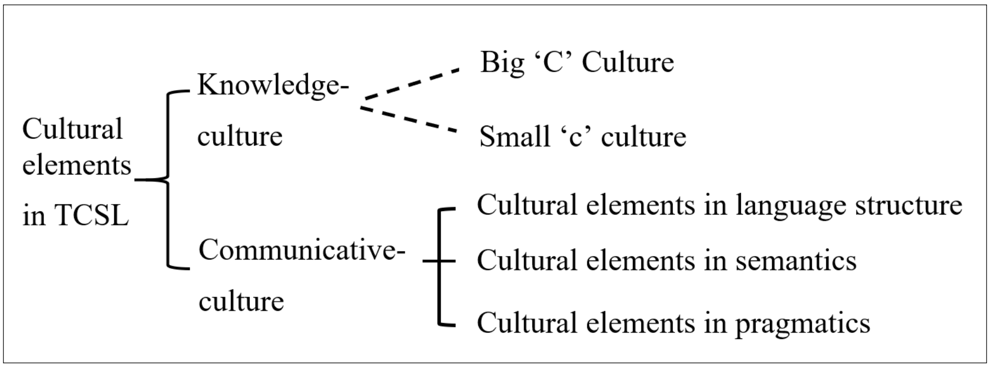
Furthermore, Chen (1992, 1997) divides communicative-culture into three aspects: cultural elements in language structure (yu gou语构), semantics (yu yi语义), and pragmatics (yu yong语用). Cultural elements in language structure (Cc-i) are embodied in the structure of vocabulary, phrases, and sentences, and reflects cultural characteristics through grammatical structure (Chen, 1992). For instance, Chinese people show respect for their ancestors by giving precedence to the surname. In the structure of Chinese appellation, the phrase 'Li xiao jie李小姐 (Miss Li)' (Chan et al., 2018) puts the surname 'Li 李' first, contrary to English and Malay language structure in which the title 'Miss' appears before the surname.

Cultural elements in semantics (Cc-ii) are contained in lexis and idioms (Chen, 1997). They carry the unique meanings, concepts, and thoughts of the Chinese people (Chan et al., 2018; Chen, 1992, 1997). For example, the word 'Xin fang 新房' means new house, but can also mean wedding house in the context of getting married in China (Chen, 1997). The cultural elements in semantics may cause misunderstandings in communication if learners do not learn the cultural meaning of the lexis and idioms (Chen, 1997).

Cultural elements in pragmatics (Cc-iii) refer to the cultural norms or conventions in language use in a social context and in interpersonal relationships (Chen, 1992, 1997). They are embodied in the speech acts of greeting, leave-taking, praise, apology, appreciation, honorifics, and taboos in communication (Chan et al., 2018; Chen, 1992, 1997). For instance, 'Shang nar?上哪儿？' or 'Nar qu? 哪儿去？' (Where are you going?) is normally used in greetings among people familiar with each other; 'Zai jian再见' (goodbye) can be replaced with 'Man zou慢走' (walk slowly) and 'Lu shang dang xin路上当心' (take care) in leave taking to show care to visitors.

The framework recommended by Chen (1992) and Zhang (1990) is concerned with cultural elements in communication. Knowledge-culture emphasises the cultural knowledge at the non-linguistic level such as the arts (Zhao, 1989, as cited in Zhang, 1990, p.17). Communicative-culture emphasises accuracy in the use of words or sentences when exchanging information. Both knowledge-culture and communicative-culture are important in communication.

Figure 1
Cultural Elements in Teaching Chinese as a Foreign/Second Language



Note. The dotted line in the figure means ‘subsumes’.

Based on the description above, it is clear that both Big ‘C’ and small ‘c’ elements are aspects of knowledge culture (Chan et al., 2018). Figure 1 shows these relationships.

Types of Culture in FL/L2 Textbooks
As Cortazzi and Jin (1999) explained, three basic types of cultural information can be included in textbooks. They are information related to the source culture, target culture, and international culture. Source culture (SC) refers to the learners’ own culture (Cortazzi & Jin, 1999). It is commonly regarded as equivalent to national culture (Byram, 1997; Spencer-Oatey, 2012). What is worth noting is that most nations nowadays contain different cultural groups within their boundaries, for instance, the United States, the U.K., Japan, and Malaysia. A source culture, therefore, may contain cultures of different cultural groups in different contexts. Target culture (TC) refers to the culture in which the target language is used as a first language (Cortazzi & Jin, 1999). International culture (IC) refers to cultures that are neither a source culture nor a target culture (Cortazzi & Jin, 1999). For example, there may be diverse cultures set in English-speaking and non-English-speaking countries, where English is not a first or second language but is used as an international language. Another example of international culture is when textbooks in Malaysia have a topic such as ‘travelling’ to Japan and Thailand for English lessons. With reference to L2 Mandarin learners in Malaysia, source culture involves information that is based on Malaysian society and culture, covering the cultures of the diverse ethnic groups in the country; target culture involves information that is based on Malaysian-Chinese culture and Chinese culture of people in China, and international culture involves information that is based on the cultures of other countries apart from Malaysian culture and the Chinese culture of China.

Based on Cortazzi and Jin’s (1999) work, Chao (2011) developed five categories for cultural information: source/local culture, target culture, international culture, intercultural interaction, and universality across culture. Intercultural interaction (ICI) refers to the comparison, reflection, or awareness of the similarities and differences between the source/local and the international or target culture through activities (Chao, 2011). These activities relate to problem-solving, case studies, and role play to help learners “develop knowledge, positive attitude, skills and awareness in intercultural communication” (Chao, 2011, p. 197). Universality across culture (UC) refers to general knowledge or content that is not specific to any particular country or culture. It is similar to the culturally neutral notion proposed by Tajeddin and Teimournezhad (2014). For instance, food can be considered a cultural theme; however, a food-related dialogue in a language textbook can be presented in such a way that it does not represent the cultural information of a source culture, target culture, or any other particular culture. Both Cortazzi and Jin’s (1999) and Chao’s (2011) frameworks on cultural information in textbooks provide a useful lens for examining cultural elements in textbooks with a focus on what is helpful for learners’ intercultural communication. The present study is informed by these frameworks on types of culture.

A textbook that reflects different cultures has different perceived benefits (Cortazzi & Jin, 1999). A textbook based on source culture can help learners talk to visitors and directly reinforce their own identity. A textbook based on target culture can help learners talk to visitors, be a visitor, develop knowledge and awareness of other cultures, as well as indirectly reinforce their own identity. Textbooks based on international culture can help learners talk to others, develop knowledge, awareness and skills of other cultures, and intercultural skills, as well as indirectly reinforce their own identity (Cortazzi & Jin, 1999). Thus, the importance of cultural information in textbooks cannot be denied since textbooks determine what benefits learners can obtain to a certain extent.

Purpose of the Study

The present study aimed to examine cultures as represented by cultural elements depicted in L2 Mandarin textbooks for Malaysian learners that were written by local authors in Malaysia. It sought to answer the following questions: (1) What cultural elements (knowledge-culture or communicative-culture) and types of culture (source, target, international cultures or intercultural interaction) are present in the selected textbooks? and (2) To what extent are these categories of cultural elements and types of culture represented in the textbooks? Content analysis using a framework combining the categories of cultural elements (Chen, 1992; Zhang, 1990), classification of Big 'C' and small 'c' (Lee, 2009; Paige et al., 1999; Pulverness, 1995; Tomalin & Stempleski, 1993), and types of culture (Chao, 2011; Cortazzi & Jin, 1999) was carried out to identify and describe the distribution of cultural elements in the textbook samples.

Methodology

Content Analysis

Content analysis as a research method "that uses a set of procedures to make valid inferences from text" (Weber, 1990, p. 117) is one of the most frequently used methods for analysing cultural content in FL textbooks (Sándorová, 2014). In the present study, content analysis was used to interpret and identify the cultural elements depicted in the textbooks. First, categories of cultural elements and types of culture were obtained from past related works and combined into an analytical framework. The categories of culture were then identified in the data, and frequencies and percentages were presented to illustrate the distribution of each category found in the data set. In the present study, qualitative content analysis, which is usually referred to as "latent level analysis because it concerns a second-level, interpretative analysis of the underlying deeper meaning of the data" (Dörnyei, 2007, p. 246), was used to identify the meanings of the text in the context of their use and link them to concepts of culture.

Sample

Four L2 Mandarin textbooks were purposively sampled for this study. The textbooks were all written by local authors and published in Malaysia. They are popular Mandarin textbooks currently in use at the university-level in Malaysia. Gateway to Mandarin Chinese 1 华语入门1 coded as TM1, was published by Xueer Publisher in 2015. Bahasa Mandarin (1) 华语1 coded as TM2, was published by Universiti Malaysia Kelantan Press in 2017. Mandarin for All (1) 华语1 coded as TM3, was published by Universiti Malaysia Pahang Publisher in 2018. Kick-start Mandarin Book 1一起来学华语1 coded as TM4, was published by UNIMAS Publisher in 2018. The textbooks were written for non-native speakers from different ethnic groups in Malaysia, namely the Malays, Indians, Iban, Kadazan, and various other minority ethnic groups. The target users are learners who do not know Mandarin and have no foundation knowledge of it. All the textbooks are currently being used in Malaysian public universities.

Defining the Categories

A categorisation scheme was created based on the works of several researchers. To examine the cultural elements in textbooks, the notions of knowledge-culture and communicative-culture (Chen, 1992; Zhang, 1990), as well as Big 'C' and small 'c' cultures (Lee, 2009; Paige et al., 1999; Pulverness, 1995; Tomalin & Stempleski, 1993) were used (see Table 1).

Table 1
Cultural Elements in Textbooks

Categories/Code	Sub-categories	Explanation	Themes or Examples
Knowledge-culture (Kc) (Zhang, 1990)	Big 'C' Culture (Lee, 2009; Paige et al., 1999; Pulverness, 1995; Tomalin & Stempleski, 1993)	A set of facts and statistics of a target speech society	Arts, literature, geography, history, education, religions, social customs, festivals, music, food, architecture, etc.
	Small 'c' culture (Lee, 2009; Paige et al., 1999; Pulverness, 1995; Tomalin & Stempleski, 1993)	The invisible and hidden aspects of a target speech society	Social values, beliefs, and norms
Communicative-culture (Cc) (Chen, 1992; Zhang, 1990)	Cultural elements in language structure (Cc-i)	Cultural elements that are embodied in the structure of vocabulary, phrases and sentences	Forms of names, dates, and times, greetings in Chinese
	Cultural elements in semantics (Cc-ii)	Cultural elements that are contained in lexis and idioms	Family kinship titles in Chinese
	Culture elements in pragmatics (Cc-iii)	Cultural norms or conventions in language used in the social context and interpersonal relationships	Linguistic taboos of a particular country

To explore the selection of different types of culture and intercultural aspects presented in the textbooks, Cortazzi and Jin's (1999) and Chao's (2011) frameworks of source culture, target culture, international culture, and intercultural interaction were used (see Table 2).

Table 2
Types of Culture in Textbooks

Categories/Code (Cortazzi & Jin, 1999; Chao, 2011)	Explanation
Source Culture (SC)	It refers to the information related to Malaysian society, covering the cultures of different ethnic groups.
Target Culture (TC)	It refers to Malaysian-Chinese culture and Chinese culture of China.
International Culture (IC)	It refers to cultures that are neither a source culture nor a target culture. In this study, it consists of all cultures except for Malaysian culture and the Chinese culture of China.
Intercultural Interaction (ICI)	It refers to comparisons, reflections, or awareness of the similarities and differences between the source/local and the international/target culture.

Detailed explanations and examples of the categories are described in the theoretical perspective section of this paper.

Coding Procedure

It is necessary to establish the units before starting coding (Cohen et al., 2007). A unit in content analysis may contain sentences, paragraphs, or the whole text, as well as people or themes (Sándorová, 2014). The units for data analysis in this study consisted of two parts: (1) each unique theme in each textbook was counted as one unit and (2) the introduction and appendix sections in each textbook were counted as one unit as a whole. Thus, the total units for the study were 38 units, including 35 themes and three sections of introductions and appendices in four textbooks.

As observed by Adaskou et al. (1990, p. 5), "Almost everything in a language course is capable of carrying a cultural load of some sort." Thus, guided by the research questions, the following elements were given attention in the analysis: illustrations, lexis, idioms and sentences in dialogues (or reading passages), descriptive text material (or notes), and contextual practice. Based on the categories (see Tables 1 and 2), the content related to cultural elements and types of culture in the four sampled textbooks was coded using the Atlas.ti software. Atlas.ti is suitable for qualitative and mixed methods data analysis because of its utility in facilitating analysis of images, text, video, and audio data, creating co-occurrence trees, as well as calculating the frequency. After finishing the coding, the frequencies of cultural elements and types of culture in the textbooks were then exported from Atlas.ti. Following the quantitative analysis, the details of cultural representation were then interpreted qualitatively highlighting specific examples from the texts.

Trustworthiness

The trustworthiness of a qualitative research study is important in evaluating its worth. Lincoln and Guba (1985) proposed four criteria: credibility, transferability, dependability, and conformability. Trustworthiness is warranted based on the accurate reporting on the stability of the data, how the data are collected, the reliability of the analysis, and how the findings represent the range of realities (see Elo et al., 2014). These principles are adhered to in the study.

Credibility

The sample has been adequately described. In the present study, the sample comprises a collection of four textbooks. The books were selected based on explicit criteria: that they were L2 Mandarin textbooks written for Malaysian learners by Malaysian writers and were currently being used for elementary learners in Malaysian public universities. The learners in the universities comprise students of various ethnicities.

Dependability

The sample, being print materials, are stable over time, and available for scrutiny by other researchers. The titles and publishers of the books are provided to enable future researchers to reanalyse the data.

Conformability

The categorisation matrix was developed after the careful examination of published works on cultural elements in L2 Mandarin textbooks to ensure they are appropriate for a detailed identification of cultural representation in textbooks. The coding process was conducted by the first researcher. Ten percent of the data were re-coded by two postgraduate students after training. One of the coders was a Malaysian who was familiar with Malaysian cultural elements. The percentage of agreement between coders was found to be more than 90% among the three coders. For each category of culture identified, examples were drawn from the textbooks to illustrate their meanings in context.

Transferability

Transferability parallels external validity (Bryman, 2008), which shows that the findings have applicability in other similar contexts or settings (Lincoln & Guba, 1985; Yin, 2014). While most qualitative studies are not generalisable in the statistical sense, the findings may serve as a useful reference for similar future studies on other samples of L2 Mandarin textbooks for Malaysian learners, and possibly textbooks for multicultural learners in other contexts.

Results

The results related to the two research questions are presented in this section.

Cultural Elements and Types of Culture in Textbooks

The two aspects of culture were examined in the four L2 Mandarin textbooks that were coded TM1, TM2, TM3, and TM4. The four textbooks were all designed for one semester of an L2 Mandarin course. There was a total of 38 units in the four textbooks, which contained different themes such as self-introductions, family, occupations, and daily expressions. From the analysis of all 38 units across the four textbooks (see appendix for complete proofs), the results are presented in Table 3.

For cultural elements, knowledge-culture, which does not directly influence the accuracy of communication between people from different backgrounds, is explicitly presented in all four textbooks (see Table 3). It is identified in almost all of the units of TM1, TM2, and TM4, but is relatively lower in TM3. Additionally, Big 'C' culture is dominant in all four textbooks while small 'c' culture is limited. Likewise, communicative-culture is also reflected in all four textbooks. Cultural elements in language structure (Cc-i), semantics (Cc-ii), and pragmatics (Cc-iii) are all embodied in TM1, TM2, TM3, and TM4 to different extents.

Table 3

Percentage of Cultural Elements and Types of Culture in Sampled Textbooks

Sample	Total units	Cultural elements					Type of culture			
		Kc (%)		Cc (%)			SC (%)	TC (%)	IC (%)	ICI (%)
		Big 'C'	Small 'c'	Cc-i	Cc-ii	Cc-iii				
TM1	7	100 (7/7)	0 (0/7)	71.43 (5/7)	14.29 (1/7)	28.57 (2/7)	100 (7/7)	100 (7/7)	0 (0/7)	0 (0/7)
TM2	11	100 (11/11)	9.01 (1/11)	45.45 (5/11)	18.18 (2/11)	27.27 (3/11)	90.91 (10/11)	81.82 (9/11)	9.01 (1/11)	0 (0/11)
TM3	10	80 (8/10)	0 (0/10)	50 (5/10)	20 (2/10)	40 (4/10)	70 (7/10)	100 (10/10)	0 (0/10)	0 (0/10)
TM4	10	100 (10/10)	0 (0/10)	20 (2/10)	0 (0/10)	40 (4/10)	40 (4/10)	100 (10/10)	10 (1/10)	0 (0/10)
Total	38	94.74 (36/38)	2.63 (1/38)	44.74 (17/38)	13.16 (5/38)	34.21 (13/38)	76.32 (29/38)	94.74 (36/38)	5.26 (2/38)	0 (0/38)

Note. Percentage (%) is for the total units in the sampled textbooks.

In terms of types of culture, target culture is the most dominant compared to source culture, international culture, and intercultural interaction, as it is present in almost all 38 units across all four textbooks (see Table 3). It is followed by source culture, which accounts for 18.42% less than target culture. International culture is only integrated into one unit each in TM2 and TM4, and intercultural interaction cannot be found in any of the textbooks and is thus the most under-represented.

Tables 4 to 7 show the analysis of categories of culture and specific examples portraying them in individual textbooks. In TM1, both source and target cultures are represented by elements of knowledge-culture (see Table 4). For source culture, it seems that only Malay culture is depicted, including the use of characters from the Malay ethnic group and portrayal of the structure of a Malay family in illustrations and dialogues. Source culture is represented by depicting the map of Malaysia, making reference to local institutions or places ('Ma liu jia马六甲 Melaka', 'Ji long po吉隆坡 Kuala Lumpur'), as well as the Muslim practice of praying on Friday. There are also elements representing the target culture, for instance, making references to Chinese movies and music (see Table 4). Further, knowledge-culture in TM1 consists mainly of the Big 'C' culture, while small 'c' cultural elements can hardly be found.

Table 4

Cultural Elements and Types of Culture in TM1

Type of culture	Knowledge-culture (Kc)		Communicative-culture (Cc)
	Big 'C'	Small 'c'	
SC	Using Malay characters in dialogues (U1-6). Organization or structure of a Malay family (U2/6). Material conditions: phone numbers in Malaysia context (U3). Geography: map of Malaysia (U7*), names of local places and institutions (U3). Religion: Muslim praying activity (U5).	\	\
TC	Using Chinese characters in dialogues (U1-6). Jackie Chan's Movie (U6). Music: children's songs in Chinese, the 'Happy Birthday' song in a Chinese context (U7*).	\	Cc-i: Forms of Chinese names (U1), Chinese appellation (U1/4), dates and times (U3/5) and addresses (U3). Cc-ii: Symbolic plant words in Chinese and 12 symbolic animals in the Chinese zodiac (U7*). Cc-iii: Norms of greetings in formal situations and informal situations (U1), Chinese number gestures (U2).
IC	\	\	\
ICI	\	\	\

Note. U=unit, for instance, U1 refers to unit 1. U7* refers to the sections of introduction and appendix in TM1.

Elements of communicative-culture that are embedded in language structure, semantics, and pragmatics are used mainly to represent the target culture in TM1, for instance, forms of expressions of dates and times and addresses, as well as norms of greetings in formal or informal situations. In the expression of dates and times, for example, the order in Chinese is 'nian年 (year)-yue月 (month)-hao号 (date)', such as '2016年12月20号 (2016, December 20)'. This is different from time expression in English or Malay where the year would appear last.

For TM2 (see Table 5), few examples of international culture and intercultural interaction can be found. Knowledge cultural elements are present in both source and target cultures. Similar to TM1, the Big 'C' elements are more prominent than the small 'c' cultural elements. For instance, a famous Chinese song entitled 'Shi shang zhi you mama hao世上只有妈妈好 (Mum is the best in the world)', representing the value of a mother's love for her children is a common song in both Malaysia and China. In contrast, communicative cultural elements are found only in the target culture. At the level of pragmatics, for example, people are depicted as wishing 'Gong xi ni ya恭喜你呀! (Congratulations to you!)' to others when good things happen, which is the cultural norm of congratulations in Chinese.

Table 5
Cultural Elements and Types of Culture in TM2

Type of culture	Knowledge-culture (Kc)		Communicative-culture (Cc)
	Big 'C'	Small 'c'	
SC	Using Malay characters in illustrations (U7) and the text (U1-10). Organization or structure of a family (no ethnic group specified) in Malaysia (U4). Material conditions: phone numbers in Malaysian contexts (U3). Geography: names of local places (U4/11*) and institutions (U9).	\	\
TC	Using Chinese characters in dialogues (U1-10). Music of Chinese songs (U2/4/8/9/10).	Social value of maternal love (U4)	Cc-i: Forms of Chinese name (U2), appellation (U1/5), dates and times (U6/8). Cc-ii: Family kinship titles in Chinese (U4), and 12 symbolic animals in the Chinese zodiac (U5). Cc-iii: Norms of greeting in formal situations (U1), responses to someone's thanks and apology (U1), congratulation (U6); honorifics and self-abasing forms (U2).
IC	Music of English songs (U6).	\	\
ICI	\	\	\

Note. U11* refers to the section of appendix in TM2.

Similar to TM2, TM3 pays more attention to source and target cultures than international culture and intercultural interaction (see Table 6). Knowledge-culture is mainly Big 'C' elements and found in source culture, related to using Malaysian local characters in illustrations, describing the structure of the Malay family, showing the Malaysian map, portraying different Malaysian ethnic groups and showing the phone number in the Malaysian format. Also, information relating to Malaysian local fruits (eg. 'Liu lian榴莲durian', 'Hong mao dan红毛丹rambutan') can be found in TM3. Small 'c' cultural elements are also limited in TM3. For communicative-culture, both source and target cultures are found to contain cultural elements in language structure, semantics and pragmatics. For example at the semantics level, the culture-loaded word 'Ken zhi min 垦殖民settler (especially a Felda settler)' is found in the text, which relates to Malaysia's history (source culture).

Table 6
Cultural Elements and Types of Culture in TM3

Type of culture	Knowledge-culture (Kc)		Communicative-culture (Cc)
	Big 'C'	Small 'c'	
SC	Using Malays characters in illustrations (U2). Religion: Muslim praying activity (U7/8). Organization or structure of the Malay family (U2). Geography: map of Malaysia (U1), races/ethnic groups in Malaysia (U5), reference to local places (U1/2/7/10) and institutions (U2/3). Food: local fruits (U10). Material conditions: phone numbers in Malaysia context (U1).	\	Cc-ii: culture-loaded words (U3)
TC	Using Chinese characters in dialogues (U1-10). Music of Chinese song (U5).	\	Cc-i: Forms of Chinese name (U1/5), appellation (U1/5), date and time (U6/8/10). Cc-ii: Family kinship titles in Chinese (U2). Cc-iii: Norms of greeting in formal situation (U1/4/7), response to someone's thanks (U1) and praise (U4), bargaining in buying things (U10).
IC	\	\	\
ICI	\	\	\

Finally, similar to TM2 and TM3, source and target cultures are dominant in TM4 (see Table 7). Cultural elements in knowledge-culture are all Big 'C' culture, which represent source, target, and international cultures, including using the Malaysian phone number format, introducing the structure of a Malay family, and showing the Malaysian map. For instance, a Malay family in Malaysia can have more than two children, whereas Chinese couples may only have up to two children under China's population policy. Other than that, Malay, Indian, Chinese, and Western food or drinks are mentioned in the text. Elements of communicative-culture are only found in target culture, covering cultural elements of language structure and pragmatics. Taking cultural elements in pragmatics as an example, TM4 portrays characters using the honorific ('Nin您you') to show respect to someone and 'Bu ke qi/Bu xie不客气/不谢 (you are welcome)' to respond to someone's thanks in the Chinese language.

Table 7
Cultural Elements and Types of Culture in TM4

Type of culture	Knowledge-culture (Kc)		Communicative-culture (Cc)
	Big 'C'	Small 'c'	
SC	Material conditions: phone numbers in the Malaysian context (U1). Organization or structure of a Malay family (U6). Geography: map of Malaysia (U1), races/ethnic groups in Malaysia (U6/10*), local places (U6), and institutions (U5). Food: Malay and Indian food/drinks (U7).	\	\
TC	Chinese opera in illustrations (U1-10*). Introduction of Chinese characters (U1-10*). Reference to Chinese food and tea (U7).	\	Cc-i: Forms of appellation (U7), dates and times (U4/5). Cc-iii: Norms of greetings in formal situations (U1), response to someone's thanks (U1), making phone calls (U8), using honorifics (U1/8/9), and bargaining in buying things (U9).
IC	Reference to Western food (U7).	\	\
ICI	\	\	\

Note. U10* refers to the sections of introduction and appendix in TM4.

Discussion

The results from the qualitative content analysis are discussed in the following sections.

Knowledge and Communicative Cultural Elements in Textbooks

The results showed that elements of knowledge-culture and communicative-culture were present to different extents in the four sampled textbooks. This is consistent with Zhang's (1990) recommendation that both knowledge-culture and communicative-culture should be included in textbooks at the elementary level of language teaching. It is possible that the local writers of the newly published textbooks tried to stress the importance of the cultural elements of both knowledge-culture and communicative-culture.

Appropriate Representation of Big 'C' and Small 'c' Culture

The results also revealed that there were different attitudes towards Big 'C' and small 'c' culture in the sampled textbooks. Among the elements of knowledge-culture, the small 'c' culture in the selected textbooks was remarkably limited while the majority of the knowledge-culture were related to Big 'C' source-culture learning. Nearly all the selected textbooks failed to reflect sociocultural values, norms, and beliefs. This seems a clear indication that the textbook writers did not regard small 'c' culture as an important aspect for elementary-level language learners. In contrast, the attempt at integrating Big 'C' source-culture learning in the textbooks realised predominantly through illustrations and dialogues was more fruitful. The inclusion of Big 'C' source-culture elements can help learners develop their interest in learning and practising Mandarin. However, it is also necessary to pay more attention to the Big 'C' in target-culture, since it may indirectly affect communication between L2 learners and native Mandarin speakers. Knowledge cultural elements, including Big 'C' and small 'c' elements, should not be ignored because they are beneficial to the development of communication competence from a long-term perspective (Zhang, 1990). Hence, writers should consider including an appropriate balance of Big 'C' and small 'c' elements in textbooks for L2 Mandarin elementary-level learners.

Explicit Representation of Communicative Cultural Elements

Communicative cultural elements as proposed by Chen (1997), which included cultural elements in language structure, semantics, and pragmatics, were presented in the sampled textbooks. The present study revealed that communicative cultural elements were explicitly depicted in the textbooks, particularly in TM1 and TM2, where explanatory notes on cultural elements are provided. In comparison, there was not as much relevant communicative-culture material in TM3 and TM4. Explicit information or instruction on these elements is essential in textbooks (Byram, 1997; Lee, 2009). Knowledge of communicative-culture helps learners to avoid making assumptions or stereotypes based on their own cultural background or pre-existing cultural knowledge (Chan et al., 2018). Through this, learners can benefit from a better understanding of the target language and use it more appropriately (Zhang, 1990). Therefore, it is important that L2 Mandarin textbook writers place more emphasis on providing explicit explanations on communicative cultural elements in the textbooks.

Source, Target, International Cultures, and Intercultural Interaction in Textbooks

With regard to the types of culture, both target and source cultures were represented more prominently than international culture and intercultural interaction in all four textbooks. This confirms the tendency towards localisation rather than internationalisation in L2 Mandarin textbooks as asserted by previous researchers (Lam & Hoe, 2013; Tan, 2007). However, there were differences in the manner in which target culture and source culture were implemented in the textbooks examined in the present study.

Emphasis on Teaching Culture

The present study discovered that the textbooks emphasised the target culture in L2 Mandarin learning rather than focusing only on linguistic skills. The target culture was found to be represented in all four textbooks. Elementary-level textbooks are often the first point of contact for learners with regard to the target culture (Bewley, 2018). Knowledge of the target culture not only benefits learners in multicultural communication (Ye & Wu, 2010), it may also arouse the learners' interest in visiting China and developing their knowledge and awareness of Chinese culture. The results of this study suggest that writers of the textbooks regard culture as an important aspect of L2 Mandarin education in Malaysia, consistent with the notion that teaching language entails teaching culture (Chan et al., 2018; Lam & Hoe, 2013).

Consideration of Learners' Diverse Cultural Backgrounds

The textbooks provided information on source culture such as Malaysian characters, places, food, maps, family structure, and religious beliefs. These cultural elements are different from what were reported in Chan et al.'s (2018) study. In the textbooks examined, source culture was oriented towards Malay culture. The Malay ethnic group was clearly represented in the textbooks whereas the Indian, Iban, Kadazan, and other ethnic groups were nearly absent. Only TM3 and TM4 had a slight mention of Indian food and the names of different ethnic groups. In contrast, Malay characters, food, family structure, and religious beliefs were dominantly represented. Lam and Hoe (2013) stated that although the target users of L2 Mandarin textbooks consisted of learners of various ethnicities, the Malay culture was put at the forefront possibly because the target users shared a common national language, Malay. However, considering the multicultural and multi-ethnic context in Malaysia, where different ethnic groups speak their own mother tongues and practice their unique cultures, regarding Malay culture as the learners' source culture would be unsuitable for the purpose of bridging learners' source culture, target culture, international culture, and intercultural interaction in FL teaching and learning. Other benefits of accurately representing learners' source culture include the reinforcement of learners' own identity (Cortazzi & Jin, 1999) and promoting cultural exchange and understanding amongst different ethnic groups (Ye & Wu, 2010).

Representation of International Culture and Intercultural Interaction

Unlike the studies of Ye and Wu (2010) and Chan et al. (2018), this study showed that there were only a few references to international culture and intercultural interaction found in the textbooks. This finding is similar to that reported in Lam and Hoe's (2013) study. Only two of the textbooks depicted international culture. For instance, TM2 made reference to English songs and TM4 had information about Western food. The result suggests that internationalisation was not a major concern in the textbooks (Lam & Hoe, 2013). Low internationalisation in textbooks may be considered a shortcoming, as learning about other cultures aside from target and source cultures is beneficial for L2 learners. Textbooks with representations of international culture can act as a cultural and linguistic resource for learners to practice speaking with others and develop their intercultural skills (Cortazzi & Jin, 1999). Textbooks that highlight intercultural interaction can help the learners compare or reflect on the differences and similarities between the source culture and the target or international cultures (Chao, 2011). Thus, textbooks that integrate international culture and intercultural interaction can help learners acquire intercultural communication skills in an increasingly globalised world.

The findings of this study complement Lam and Hoe (2013), Tan (2007), Yang (2019), and Ye and Wu (2010) by combining the analysis of cultural elements with the types of culture represented in textbooks. The study also extends Chan et al.'s (2018) work investigating cultural and intercultural elements in a similar context by addressing the multicultural dimension of the source culture in Malaysia's culturally diverse context. The findings of this study are limited to four selected L2 Mandarin textbooks and are not representative all of those written for Malaysian learners. Future studies can expand the sample size to analyse cultural representation in textbooks.

Conclusion

Two aspects of culture were analysed in this study, namely, cultural elements and types of culture. The results of the present study showed that the selected textbooks contained the cultural elements of both knowledge-culture and communicative-culture to different extents. The knowledge cultural elements reflected the Big 'C' more than the small 'c' culture. The larger part of the Big 'C' cultural elements had to do with source culture. The source culture, in turn, was mainly depicted as Malay culture with weaker representation of the other ethnic groups. Finally, communicative cultural elements were present, but lacked explicit explanatory notes that could help learners understand the implicit meanings of these elements. Overall, the results of this study revealed that the target culture was the most dominant in the textbooks, followed by the source culture, while little attention was paid to international culture and intercultural interaction.

The findings of this study suggest that there is a need to appropriately increase the content for small 'c' culture in L2 Mandarin textbooks for elementary learners. It is also necessary to integrate more content related to Big 'C' target-culture. Increasing the quantity and quality of cultural elements representing the cultures of the diverse ethnic groups in Malaysia is also worth considering. The textbooks need to provide more information on explaining the cultural elements of communicative-culture. It is recommended that an appropriate balance of source, target, and international cultures and intercultural interaction be given serious thought and due consideration by textbook writers. It is hoped that the findings of the study will be useful in raising awareness for the need to integrate cultural knowledge into teaching materials that are balanced and meaningful to L2 learners, particularly among authors, editors, publishers of textbooks, and Mandarin teachers. Finally, future research should look into how a diversity of cultural elements can be more effectively integrated into L2 Mandarin textbooks with the aim of helping learners develop cultural competence, and more importantly, gain global cultural consciousness.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for profit sectors.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

Adaskou, K., Britten, D., & Fahsi, B. (1990). Design decisions on the cultural content of acourse for Morocco. *ELT Journal, 44*(1), 3-10.

Ajideh, P., & Panahi, M. (2016). An analysis of culture-related content in English textbooks for Iranian students entitled 'Prospect' and 'Vision' Series. *International Journal of Language and Linguistics, 3*(6), 87-93.

Baker, W. (2011). From cultural awareness to intercultural awareness: culture in ELT. *ELT Journal, 66*(1), 62–70. http://dx.doi.org/10.1093/elt/ccr017

Bewley, F. (2018). *Culture and language ideology in Chinese foreign language textbooks: A thematic analysis* [Unpublished Master's dissertation]. The University of British Columbia.

Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). Addison Wesley Longman.

Bryman, A. (2008). *Social research methods* (3rd ed.). Oxford University Press.

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.

Chan, S. F., Dewitt, D., & Chin, H. L. (2018). The analysis of cultural and intercultural elements in Mandarin as foreign language textbooks from selected Malaysian public higher education institutions. *Jurnal Kurikulum & Pengajaran Asia Pasifik, 3*, 66-90.

Chao, T. (2011). The hidden curriculum of cultural content in internationally published ELT textbooks: A closer look at new American Inside Out. *The Journal of Asia TEFL, 8*(2), 189-210.

Chastain, K. (1988). *Developing second-language skills: Theory and practice*. Harcourt Brace Publishers.

Chen, G. L. (1992). 语言教学中的文化导入 [Cultural introduction in language teaching]. *Language Teaching and Linguistic Studies, 3*, 19-30.

Chen, G. L. (1997). 关于对外汉语课中的文化教学问题 [Cultural teaching in teaching Chinese as a second language]. *Applied Linguistics, 1*, 1-4. http://dx.doi.org.10.16499/j.cnki.10035397.1997.01.005

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. Routledge.

Cortazzi, M., & Jin, L. (1999). Cultural mirrors: Materials and methods in the EFL classroom. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 196-219). Cambridge University Press.

Dehbozorgi, M., Amalsaleh, E., & Kafipour, R. (2014). Exploring cultural content of three prominent EFL textbooks in Iran: A case study of American English Files, Top Notch and Four Corners. *Acta Didactica Napocensia, 7*(1), 69-81.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford University Press.

Elo, S., Kariainen, M., Kanste, O., Polkki, T., Ultriainen, K., & Kyngas, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open, 4*, 1–10. https://dx.doi.org/10.1177/2158244014522633

Hanban. (2014). *International curriculum for Chinese language education* (Revised ed.). Beijing Language and Culture University Press.

Hoe, F. T. (2014a). A study of Malaysian national universities Mandarin course teaching materials. *Overseas Chinese Education, 70*, 101-107.

Hoe, F. T. (2014b). 马来西亚国立大学华语教学及教师能力标准研究 [A study of teachers' competency standards and teaching Mandarin as a second language in Malaysian public universities] [Unpublished doctoral dissertation]. Universiti Tunku Abdul Rahman, Malaysia.

Hoe, F. T. (2016). 马来西亚国立大学华语课程中的文化教学 [The teaching of culture in Mandarin courses in Malaysia public universities]. Paper presented at *International Conference on Teaching Mandarin as A Second Language in Kongzi Institute*, UM, Malaysia.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford University Press.

Kumaravadivelu, B. (2008). *Cultural globalization and language education*. Yale University Press.

Lam, K. C., & Hoe, F. T. (2013). Mandarin classroom materials for Malay-speaking students: A case in UNIMAS. *Overseas Chinese Education, 3*, 284-289.

Lee, K. Y. (2009). Treating culture: What 11 high school EFL conversation textbooks in South Korea do. *English Teaching: Practice and Critique, 8*(1), 76-96.

Li, Q. (2015). 汉语教材的"国别化"问题探讨 [On country specialization of Chinese language textbooks]. *World Chinese Teaching, 4*, 526-540.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

Liu, C. Y. (2014). On the discourse perspectives in Chinese textbooks for learners in a specific country. *Journal of International Chinese Teaching and Research, 4*, 73-79.

Liu, X. J. (2018). The choice of cultural topics in textbooks of teaching Chinese as a foreign language and Chinese culture going global. *Journal of Guangxi University for Nationalities, 40*, 200-204.

Paige, R. M., Jorstad, H., Siaya, L., Klein, F., & Colby, J. (1999). Culture learning in language education: A review of the literature. In R. Paige, D. Lange & Y. Yershova (Eds.), *Culture as the core: Integrating culture into the language curriculum* (pp. 47-114). The Centre for Advanced Research on Language Acquisition, University of Minnesota.

Pan, B. S. (2011). Communication and development of Mandarin in Malaysia: A survey on the TCSL (Mandarin classes) in national universities. *Journal of Chinese language Teaching, 7*, 242-248.

Pulverness, A. (1995). Cultural studies, British studies and EFL. *Modern English Teacher, 4*(2), 7-11.

Sándorová, Z. (2014). Content analysis as a research method in investigating the cultural components in foreign language textbooks. *Journal of Language and Cultural Education, 2*(1), 95-128.

Spencer-Oatey, H. (2012). What is culture? A compilation of quotations. *Global PAD Open House*. http://www.warwick.ac.uk/globalpadintercultural

Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford University Press.

Tajeddin, Z., & Teimournezhad, S. (2014). Exploring the hidden agenda in the representation of culture in international and localized ELT textbooks. *The Language Learning Journal, 43*(2), 180-193. http:// dx.doi.org /10.1080/09571736.2013.869942

Tan, L. N. (2007). 马来西亚对外汉语教材的走向 [Trend of Chinese language materials in Malaysia]. *Proceedings of International Conference for Overseas Chinese language Textbook in Australian National University*. https:// www.researchgate.net/publication/315382429

Tomalin, B., & Stempleski, S. (1993). *Cultural awareness*. Oxford University Press.

Vernier, S., Barbuzza, S., Giusti, S. D., & Moral, G. D. (2008). The five language skills in the EFL classroom. *Nueva Revista de Lenguas Extranjeras, 10*, 263–291.

Weber, R. P. (1990). *Content analysis*. In Basic content analysis (2nd ed., pp. 117-124). Sage Publications.

Yang, M. (2019). 马来西亚本土华语教材词汇编写研究：以《华语入门》为例 [Study on vocabulary compilation of local Mandarin textbook in Malaysia: Taking Introduction to Chinese as an example]. *Today's Mass-media, 4*, 142-145.

Ye, T. T. (2011). 马来西亚高校汉语作为二语教学发展研究 [A study on the development of teaching Chinese as second language in Malaysian universities] [Unpublished doctoral dissertation]. Minzu University of China.

Ye, T. T., & Wu, Y. H. (2010). A review of TCSL teaching materials in Malaysia. *Journal of Yunnan Normal University, 4*, 63-67.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage Publications.

Zhang, Z. Y. (1990). 试议交际文化和知识文化 [Communication culture and knowledge culture]. *Language Teaching and Linguistic Studies, 1*, 15-32.

Zhao, X. Z. (1989). 文化差异与文化导入论略 [Cultural differences and cultural guidance]. *Language Teaching and Linguistic Studies, 1*, 76-83.

Appendix

Coding Table of Cultural Elements and Types of Culture in Four Sampled Textbooks

Sample	Unit	Cultural Elements					Type of Culture			
		Kc		Cc			SC	TC	IC	ICI
		Big ‘C’	Small ‘c’	Cc-i	Cc-ii	Cc-iii				
TM1	1	+	-	+	-	+	+	+	-	-
	2	+	-	+	-	+	+	+	-	-
	3	+	-	+	-	-	+	+	-	-
	4	+	-	+	-	-	+	+	-	-
	5	+	-	+	-	-	+	+	-	-
	6	+	-	-	-	-	+	+	-	-
	7*	+	-	-	+	-	+	+	-	-
	Total	***7/7***	***0/7***	***5/7***	***1/7***	***2/7***	***7/7***	***7/7***	***0/7***	***0/7***
	F (%)	***100%***	***0%***	***71.43%***	***14.29%***	***28.57%***	***100%***	***100%***	***0%***	***0%***
TM2	1	+	-	+	-	+	-	+	-	-
	2	+	-	+	-	+	+	+	-	-
	3	+	-	-	-	-	+	+	-	-
	4	+	+	-	+	-	+	+	-	-
	5	+	-	+	+	-	+	+	-	-
	6	+	-	+	-	+	+	+	+	-
	7	+	-	-	-	-	+	+	-	-
	8	+	-	+	-	-	+	+	-	-
	9	+	-	-	-	-	+	+	-	-
	10	+	-	-	-	-	+	+	-	-
	11*	+	-	-	-	-	+	-	-	-
	Total	***11/11***	***1/11***	***5/11***	***2/11***	***3/11***	***10/11***	***9/11***	***1/11***	***0/11***
	F (%)	***100%***	***9.01%***	***45.45%***	***18.18%***	***27.27%***	***90.91%***	***81.82%***	***9.01%***	***0%***
TM3	1	+	-	+	-	+	+	+	-	-
	2	+	-	-	+	-	+	+	-	-
	3	+	-	-	+	-	+	+	-	-
	4	-	-	-	-	+	-	+	-	-
	5	+	-	+	-	-	+	+	-	-
	6	-	-	+	-	-	-	+	-	-
	7	+	-	-	-	+	+	+	-	-
	8	+	-	+	-	-	+	+	-	-
	9	+	-	-	-	-	-	+	-	-
	10	+	-	+	-	+	+	+	-	-
	Total	***8/10***	***0/10***	***5/10***	***2/10***	***4/10***	***7/10***	***10/10***	***0/10***	***0/10***
	F (%)	***80%***	***0%***	***50%***	***20%***	***40%***	***70%***	***100%***	***0%***	***0%***

Sample	Unit	Cultural Elements					Type of Culture			
		Kc		Cc			SC	TC	IC	ICI
		Big 'C'	Small 'c'	Cc-i	Cc-ii	Cc-iii				
TM4	1	+	-	-	-	+	-	+	-	-
	2	+	-	-	-	-	-	+	-	-
	3	+	-	-	-	-	-	+	-	-
	4	+	-	+	-	-	+	+	-	-
	5	+	-	+	-	-	-	+	-	-
	6	+	-	-	-	-	+	+	-	-
	7	+	-	-	-	+	+	+	+	-
	8	+	-	-	-	+	+	+	-	-
	9	+	-	-	-	+	-	+	-	-
	10*	+	-	-	-	-	-	+	-	-
	Total	***10/10***	***0/10***	***2/10***	***0/10***	***4/10***	***4/10***	***10/10***	***1/10***	***0/10***
	F (%)	***100%***	***0%***	***20%***	***0%***	***40%***	***40%***	***100%***	***10%***	***0%***
TM1-4	***Total***	***36/38***	***1/38***	***17/38***	***5/38***	***13/38***	***29/38***	***36/38***	***2/38***	***0/38***
	F (%)	***94.74%***	***2.63%***	***44.74%***	***13.16%***	***34.21%***	***76.32%***	***94.74%***	***5.26%***	***0%***

Note. Kc refers to knowledge-culture, Cc refers to communicative-culture. Cc-i refers to cultural elements in language structure, Cc-ii refers to cultural elements in semantics, Cc-iii refers to cultural elements in pragmatics. SC refers to source culture, TC refers to target culture, IC refers to international culture, ICI refers to intercultural interaction. The asterisk (*) refers to the sections of introduction and appendix in textbooks.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Nazim, M., & Hazaea, A.N. (2020). Enhancing Language Assessment Skills through Need-based Training of Faculty Members in EFL Tertiary Context. *Journal of Language and Education, 6*(4), 138-152. https://doi.org/10.17323/jle.2020.10124

Enhancing Language Assessment Skills through Need-based Training for Faculty Members in EFL Tertiary Contexts

Mohd Nazim, Abduljalil Nasr Hazaea
Najran University

Correspondence concerning this article should be addressed to Abduljalil Nasr Hazaea, Department of English Language, Najran University, Saudi Arabia. E-mail: agaleel@gmail.com

In Saudi higher education, assessment has shifted to incorporate intended learning outcomes rather than merely textbook content. Subsequently, faculty members unwillingly participate in high-stakes competitive and harmonized assessment in English as a Foreign Language (EFL) courses during the preparatory year (PY). These challenges emphasize the importance of need-based training for faculty members. Accordingly, this context-specific study scrutinized faculty members' needs as well as the impact of a training program on engaging the participants and on fostering language assessment skills (LASs) among them. In so doing, an action research design used pre- and post-questionnaires and included a training portfolio to collect data from 31 faculty members. The study first identified those needs as instructional skills, design skills, and educational measurement skills. In the context of professional development, the researchers designed a training program based on those reported needs. During training, the participants expressed their satisfaction with the language assessment. After the training, the participants greatly improved their LASs. The paired tests indicated that the faculty members increased their instructional skills, design skills, and skills of educational measurement. Further research is recommended for enhancing LASs among EFL students.

Keywords: language assessment skills, faculty members, preparatory year, high-stakes tests, professional training

Introduction

In Saudi higher education, assessment has shifted to be based on intended learning outcomes rather than only on textbook content. This is specifically for high-stakes tests which determine the career of first year students. In a Saudi university preparatory year (PY), the context of the present study, English as a Foreign Language (EFL) assessment has become harmonized, responsibility is shared, and it is based on learning outcomes. These shifts have led to several reasons for fostering language assessment among faculty members. First, it has become a challenge for many faculty members to design an assessment matrix and a question paper based on learning outcomes. To cope with this challenge, faculty members need new language assessment skills (LASs) after employing content-based assessment for so long. Second, faculty members have different and overlapping roles in harmonized assessment. Their roles can be that of classroom teacher, test writer, course coordinator, and/or exam committee member. These roles make them disengaged and demotivated to participate in language assessment. Third, faculty members joined the English department with various teaching backgrounds, qualifications, and experience levels. Subsequently, the authors experienced poor representations of learning outcome-based assessment in the submitted exam proposals of many test writers. For these reasons, the researchers are interested in conducting a training program for faculty members teaching courses in a Saudi university PY program.

Several studies investigated teachers' language assessment literacy (LAL) in terms of a needs analysis such as Öz and Atay (2017) in Turkey, Lam (2015) in Hong Kong, and Lan and Fan (2019) in China. This is because LAL remains a neglected area in many programs that prepare language teachers (Fulcher, 2012). In fact, LAL involves principles, knowledge, and skills in language assessment. Theoretically speaking, Taylor (2013) reported that classroom language teachers need the highest level of skills. Subsequently, this study puts a focus on LASs for two reasons. First, faculty members usually need practical solutions when encountering challenges at their workplace. Second, skills coincide with the concept of 'training'. Under the umbrella of LAL, therefore, LASs are a potential area that no study has concentrated on or addressed in the form of a need-based training program.

At the same time, workplace training is essential for promoting LASs among faculty members to meet 21st century literacies for foreign language assessment in Saudi higher education. Faculty members need to understand certain issues about assessment such as its purpose and outcomes in addition to working with assessment data (Boyd & Donnarumma, 2018). When faculty members are trained in assessment skills, they become more confident with high-stakes tests. It is assumed that the participants in this study will overcome assessment challenges and become more confident aligning learning outcomes and textbook contents after training them on LASs.

Therefore, this context-specific study investigates the effectiveness of need-based training on LASs among faculty members working in a Saudi university PY program. Employing an action research design, this study addresses these three research questions:

1. What is the average level of LAS knowledge among faculty members in a PY EFL context?
2. To what extent are the participants engaged in the training treatment?
3. How effective is professional training on LAS development?

Literature Review

LAL, an emerging area in language assessment, is defined as the knowledge, skills, and principles in language assessment (Fulcher, 2012; Inbar-Lourie, 2008; Pill & Harding, 2013; Taylor, 2013). As a result, several models have been developed to investigate LAL (Brindley, 2001; Fulcher, 2012; Inbar-Lourie, 2008; Pill & Harding, 2013; Taylor, 2013), and how it can enhance the knowledge, skills, and principles in language assessment among various stakeholders (Davies, 2008; Language Assessment Literacy Symposium, 2016). Taylor (2013) summarized a conference on LAL and called for further research in this emerging and challenging field. The conference reported diverse stakeholders including "test writers, classroom teachers, university administrators and professional language testers" (p.410) who need various levels of LAL. As for classroom language teachers, Taylor further showed that they need the highest level of language pedagogy and technical skills. However, they need the lowest level of principles, concepts, and scores and decision making. In other words, language teachers need practical and teaching skills to achieve their assessment duties in a team-work context. On the contrary, the deanship and the program coordinator determine the assessment approach, plan, and matrix. Subsequently, language teachers need the lowest theoretical orientation about language assessment.

Several studies investigated perceptions, principles, and knowledge of assessment among language teachers. Lan and Fan (2019) identified the training needs for assessment among Chinese EFL teachers. Boyd and Donnarumma (2018) concentrated on the principles and perceptions of language teachers in London towards language assessment. Although they conducted a need-based training assessment program, they did not put a focus on LASs. Öz and Atay (2017) examined the perceptions and knowledge of assessment among Turkish EFL instructors in a PY program. The same authors found an imbalance between teachers' perceptions and classroom reflections. LASs still a great deal of potential for further research. Djoub (2017) advocated teacher training into assessment literacy through conscious professional development courses; something that includes multi-dimensional awareness about the importance, benefits, strategies, and techniques of language assessment. Galichkina (2016) investigated assessment awareness among EFL teachers in Russia through a project-based learning activity. This activity triggered a deep awareness of and confidence about assessment among the participants. The same study recommended integrating this activity in an ESL course. Subsequently, the present study incorporates this activity into its training program.

Gebril and Boraie (2016) designed a training program for EFL teachers in Egypt. They recommended that a similar program be conducted for EFL teachers. Similarly, Lam (2015) explored the training landscape in language assessment among EFL teachers. The study found inadequate language assessment in Hong Kong and recommended LAL. Hatipoğlu (2015) investigated the issue of language assessment among language teachers in Turkey. The study found a great effect on the assessment culture and assessment experience on pre-service teachers. These findings indicate that changes in assessment practices need well-designed training to cope with new assessment changes. Vogt and Tsagari (2014) identified the training needs in language assessment in seven European countries. They found that EFL teachers need training in language assessment. Instead, teachers compensated for insufficient training by using existing materials for language assessment. Walters (2010) applied 'standards reverse engineering' for the sake of critical awareness among ESL teachers and recommended applying it to classroom-based assessment to examine its effect on test construction and use in the classroom. Inbar-Lourie (2008) investigated language assessment courses in language programs to find gaps in teachers' qualifications. So far, it seems that no study has concentrated on LASs and addressed them through professional development need-based training programs in higher education EFL contexts.

Previous research also called for fostering LASs in the Saudi EFL context (Hakim, 2015; Hazaea, 2019; Hazaea & Tayeb, 2018; Mansory, 2016; Umer, Zakaria, & Alshara, 2018). Hazaea (2019) examined the need for professional development in ELT and recommended further research on LAL. Hakim (2015) suggested the use of training workshops to provide EFL facilitators with a clear and deep consideration of effective LAL. Similarly, Mansory (2016) called for long-term collaborative training and support that would greatly help in raising assessment literacy among teachers. Hazaea and Tayeb (2018) recommended training for EFL teachers to develop LAL. Umer et al. (2018) found a mismatch between Saudi EFL teachers' assessment tasks and course learning outcomes. Thus, need-based training is critical for promoting LASs among EFL faculty members to meet the 21st century literacies required for foreign language assessment.

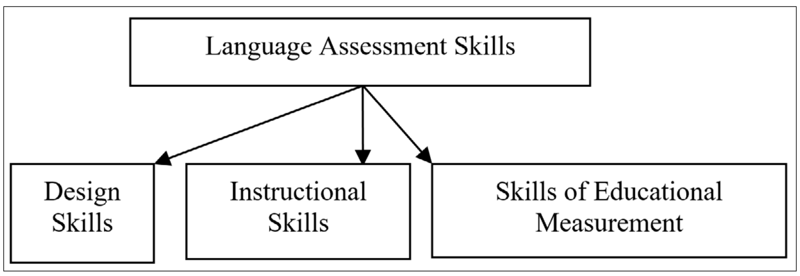
Previous research also showed that several questionnaires have been used for LAL. Fulcher (2012) developed a questionnaire to find out the training needs regarding assessment among language teachers. Herrera, Macías, and Fernando (2015) developed a questionnaire for EFL teachers on assessment literacy. Crusan, Plakans, and Gebril (2016) developed a survey to examine the perspectives and background of writing teachers on assessment. Djoub (2017) developed a questionnaire for measuring the effectiveness of LAL on practices of assessment and Giraldo (2018) developed a questionnaire for LASs among teachers. The present study adopted the content of this recent questionnaire as it best fits the context of PY programs.

Methodology

Approach

In language assessment, unique stakeholders need different levels of LASs. Language teachers need three sub-skills in language assessment. Giraldo (2018) proposed a list of LASs for language teachers. These skills are categorized as instructional skills, design skills, skills of educational measurement, and technological skills. The last two categories can be combined in one category where technological software is used for educational measurement. This list of skills coincides with Stabler-Havener's (2018) groupings of LASs: skills in item-writing, skills in test construction, and skills in analysis. Accordingly, the present study classifies list of assessment skills as instructional skills, design skills, and skills of educational measurement (Figure 1).

Figure 1
Language Assessment Skills (LASs)



Design skills refer to skills for writing test syllabuses that are based on learning outcomes (Taylor, 2009). These design skills also include true and false questions, multiple-choice questions, and so on (Giraldo, 2018). Design skills also involve the sub-skills of using rubrics and the skills for providing security for soft and hardcopies of tests. Writing exam proposals, which includes presenting, sequencing, and assisting in the achievement of learning objectives, is an integral part of design skills. Accordingly, faculty members should be able to design assessments that are "reliable, authentic, fair, ethical, practical, and interactive" (Fulcher 2012, p.116).

Instructional skills refer to the process of collecting formal and informal data for students' language development. Formal means of data include portfolios and tests (Rea-Dickin, 2001). Informal data includes classroom observation. It also involves the skills of giving feedback and motivating students based on their results (Hill & McNamara, 2012). Instructional skills also subsume skills of reporting and interpreting results to various stakeholders such as students and parents (Giraldo, 2018). It is assumed that a firm foundation in instructional skills sharpens LASs among faculty members.

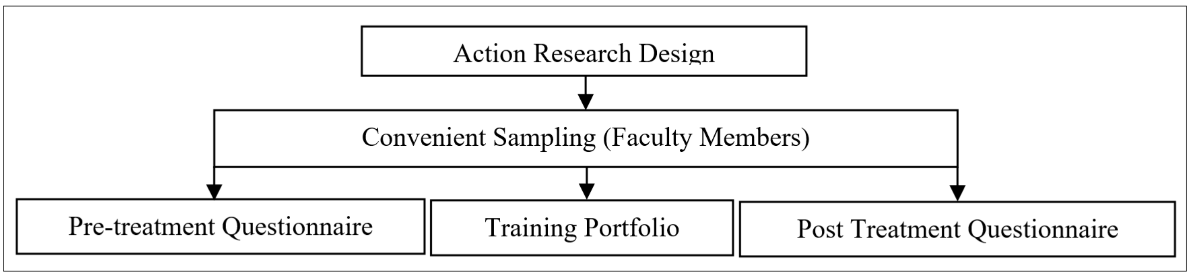
Skills of educational measurement include the ability to use statistical software such as Excel and SPSS. In so doing, faculty members will be able to calculate descriptive statistics such as means, curves, standard deviations, reliability, and correlations. Skills of educational measurement go beyond analysis into the interpretation of data showing students' strengths and weaknesses (Giraldo, 2018). This view is supported by Davies (2008) and Inbar-Lourie (2012).

Several training approaches guide language assessment (Malone, 2017). Recently, Boyd and Donnarumma (2018) suggested a trainee-centered approach for LAL. In general, this approach deals with teachers not as individual novices, but as collective experts. This approach builds teachers' engagement and confidence regarding assessment. Another feature of this approach is its flexibility for the issues of time and resources. It is a form of inductive learning that is suitable for workplace training. The trainer is a 'regulator' and the participants play the role of problem solvers. The current study uses this approach for its need-based training program.

Research Design

This study employed action research (Figure 2). First, a needs analysis questionnaire was administered to a convenient group of male faculty members because of the gender segregated system in Saudi universities. During the treatment, three workshops were conducted to address their needs in the form of a training program. After each workshop, a workshop evaluation form was distributed among the participants. At the end of the program, a post-treatment questionnaire was used.

Figure 2
Research Design



Setting and Participants

Thirty-five male faculty members deliver six English courses across approximately 30 sections in the Department of English Language Skills in the PY program of Saudi University. These faculty members joined the department with different degrees (Ph.D. [n=8], Master's [n=24] and BA [n=3], different teaching experiences (1-15 years), and various statuses of English language ability (EFL, ESL/NS). In terms of age, 45.2% of the participants are between 30 and 39 years of age; 35.4% are over the age of forty; and about 20% of the participants are under the age of thirty (Table 1).

Table 1
Background Information of Participants

Age	Percent	Teaching Experience	Percent	Qualification	Percent
25-29	19.4%	1-5 years	25.8%	BA	9.7%
30-39	45.2%	6-10 years	29%	Master	71%
+40	35.4%	11+ years	45.2%	Ph.D.	19.3%

For teaching, 45.2% of the participants have more than ten years of experience. As their terminal degree, 71% of the participants hold a master's degree in English. The researchers informed the participants about the nature of the study and received their consent to use the data for research purposes with anonymity.

Faculty members have different overlapping roles and shared responsibility in harmonized assessment as classroom teachers, test writers, course coordinators, and/or exam committee members. All faculty members are considered classroom teachers. For testing, however, the department initially invited proposals from selected faculty members (test writers). They are provided with an assessment plan for which they must design questions based on intended learning outcomes. They then send the proposals to the course coordinators. The coordinators review the proposals and prepare one suggested exam draft, which is then sent to the exam committee. The exam committee reviews the items based on the characteristics of a good test and prepares valid and reliable tests. Necessary adjustments are made after multiple reviews. The exam paper is proofread again before it is printed.

The exam is administrated according to the schedule of the central examination committee. Exam papers are collected and distributed for evaluation, which is conducted by groups of 2-3 members. An exam paper goes through the various stages of evaluation, marking, rechecking, and teacher filtering. Marks are inserted into Excel spreadsheets for data analysis including question-wise analysis, learning outcomes, figures, means and standard deviations. Therefore, the target faculty members have overlapping roles in terms of designing and interpreting high-stakes assessment.

Instruments

Two instruments were used in this study: a questionnaire and a training portfolio. The pre-treatment questionnaire consisted of three parts: personal information, format and delivery method, and content of the questionnaire. Personal information included their name (optional), age, terminal degree, and teaching experience. Format and delivery method involve faculty members' preferences in terms of delivery format and trainer . The content of the questionnaire emphasized three main areas of LASs: design skills (12 items), instructional skills (12 items), and skills of educational measurement (8 items). A 5-point Likert scale was used; (from strongly disagree to strongly agree). The pre-treatment questionnaire was then distributed among the faculty members to identify their levels of LASs.

The data was analyzed using SPSS 20 software in which several analyses were employed: descriptive analysis, independent t-test, and paired t-test. Cronbach's Alpha was 0.974, which indicates that the questionnaire was reliable for internal consistency. After the needs analysis, the study addressed the reported needs in the form of a training program. The post-treatment questionnaire followed the content of the questionnaire used in the needs analysis. Cronbach's Alpha was 0.986.

The training portfolio was another instrument used in this study. It should be a systematic and organized tool based on participants' reflections and assessments (Khan & Begum, 2012). A training portfolio consists of observation sheets, training materials, and workshop evaluation form (Binet, Gavin, Carroll, & Arcaya, 2019). On the workshop evaluation form, the researchers asked the participants to point out their level of satisfaction and engagement with the workshops. A 1 to 5 scale was used where 1 showed the least level of satisfaction, and 5 the highest level of satisfaction. The evaluation forms were analyzed by calculating the degree of engagement under each scale. The observations sheets were analyzed using thematic analysis.

Procedure

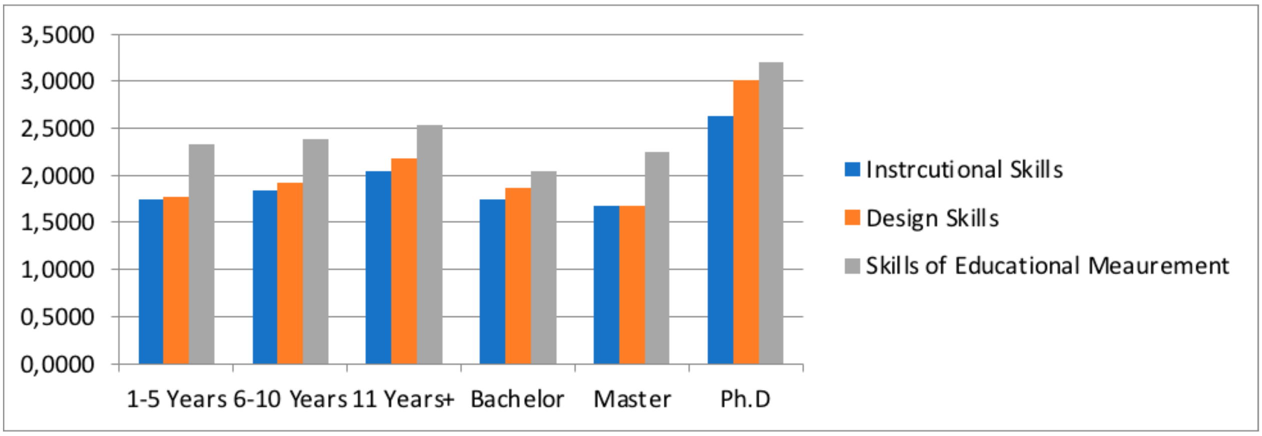
Needs Analysis

Thirty-one faculty members responded to the questionnaire. The two authors, Ph.D. holders, and two master's holders did not respond to the questionnaire. To collect objective data, the two authors avoided responding to the items on the questionnaire because their responses might affect the results of the study. The researchers analyzed the responses of the participants on the format of activities for LASs. As for the activities, the participants preferred workshops, group discussions, and seminars: 45%, 35%, and 20%; respectively. While 64.5% preferred conducting workshops at the workplace, 35.5% preferred online activities. As for preferred speakers, 32.5% of the participants wanted their colleagues to share their experiences. While 45% of the respondents preferred an expert from outside the institution, 22.5% preferred an expert from inside the institution.

The needs analysis shows low skill levels for educational measurement, design skills, and instructional skills (2.41, 1.97, 1.86); respectively. Then, the data analysis was conducted to see the relationships between LASs and teaching experience as well as qualifications (Figure 3).

Figure 3

LASs vs. Teaching Experience and Qualification (Needs Analysis)



It was found that teaching experience had no effect on instructional skills, design skills, and skills of educational measurement. All values were below the average except for skills of educational measurement where the participants have more than ten years of experience. Their value reached the average (m=2.53). This result can be attributed to assessment based on learning outcomes being a new experience for the participants who used to use content-based assessment. On the contrary, qualifications had a role in shaping the levels of LASs. Compared with other qualifications, Ph.D. holders (n=6) were found to be somewhat in need on skills of educational measurement (m=3.2), design skills (m=3), and instructional skills (m=2.6). Both holders of bachelor's degrees (n=3) and master's degrees (n=22) were still highly in need of training on LASs.

Accordingly, the researchers designed a training program. As most of the participants were master's degree holders, the researchers designed the training program to suit their need level and requested that all of the participants join the training program. Similarly, the researchers used workshops and group discussion activities. In addition, the activities were conducted at the workplace and online tutorials were shared among the participants. Although 45% of the participants preferred external experts, the researchers conducted the training program with an internal expert from the e-learning unit. It was beyond the researchers' limit to invite and support external experts to train the participants on LASs.

Training Program

In the context of professional development, the researchers as 'regulators' designed and conducted a training program consisting of six one-hour training sessions and two-and-a-half-hour online tutorials. The researchers reviewed the training materials between workshops and the reported needs to best enhance LASs among the participants. In addition, the researchers are qualified enough to facilitate this job since the first author is the head of the examination committee and the second author's area of expertise are language assessment (Hazaea & Tayeb, 2018) and professional development (Hazaea, 2019). The participants, through the workshops, were trained (Ørngreen & Levinsen, 2017) on LAS areas as revealed by the needs analysis report. Each training session started with a theoretical orientation and then a group discussion. While two of the workshops were conducted by the two authors, one workshop was conducted by an expert from the e-learning unit.

To address the need for design skills, the first author conducted a training workshop entitled "Testing the test". It aimed to investigate question papers in terms of process and product. It involved the planning and preparation of exam proposals. It covered assessment methods and test construction. Using the revise engineering method (Walters, 2010), the participants were divided test-wise into four groups: Listening and Speaking, Reading, Writing, and Grammar. First, each group was given the submitted proposals for the previous final exam. They were then asked to evaluate the exam proposals and to find the strengths and weaknesses, and give suggestions for improving the exam proposals. After that, each group was given the final version of the same exam papers. They evaluated them in the same way they evaluated the exam proposals. Finally, each group compared the exam proposal with the final version of the exam papers.

To address the need for instructional skills, the second author conducted a training workshop entitled "Enhancing instructional skills for language assessment". The workshop first contextualized language assessment as high-stakes assessment. It then explained instructional skills for English courses. Because of time constraints and to cope with the participants' preferences, the second author shared *YouTube* links with the participants on three topics: How to use 1) Blackboard, 2) Google forms and 3) students' portfolios for language assessment. During group discussions, the participants were divided course-wise into four groups. They were given copies of the first midterm papers, course learning outcomes, covered syllabi, and a question paper review template. They were asked to review the exam papers, align learning outcomes, content, and exam papers, and to think about the strengths, weaknesses and suggestions for improvement. As an individual long-term task, the participants were asked to collect their students' scripts (formal data) and the participants' observations (informal data), and analyze them for students' language development and feedback. The participants were also asked to introduce students' portfolios for assessing writing and reading classes, and to use Blackboard for continuous assessment of their students.

To address the need for educational measurement skills, a training workshop entitled "Excel program as an educational measurement tool" was conducted in the computer lab by an expert from the e-learning unit. It aimed to apply Excel software to language assessment. To highlight the importance of Excel, the expert trained the participants on the types of data (texts, numbers and equations, formulas) and on how to create them in the program. They were also trained on how to protect the data inside the worksheet as well as on how to protect the Excel file with a password. The participants were also trained on how to generate means, standard deviations, and use of figures to report students' results, learning outcomes, and question-wise analysis.

The training program was intensively conducted during the second midterm week, a period where faculty members can be involved in an assessment environment. During this week, faculty members had a chance to participate in the workshops. After each workshop, a workshop evaluation form was distributed among the participants. Later, a post-questionnaire was distributed among the participants to evaluate the effectiveness of the treatment.

Results

Engagement in the Treatment

The analysis of the training portfolio shows that the participants were satisfied with the training program. They became more engaged and confident about test design and interpretation.

The analysis of the observation sheets shows that participants asked many questions during the theoretical part, especially during the sessions on instructional skills and design skills. These participants were curious about some of the assessment terms. When they were asked about 'language assessment literacy', it was observed that there were no responses, which indicates that the term seemed to be new to them. They indicated that although they practiced some instructional skills such as reporting slow learners and good performance, they depended on direct methods for assessment data. They rarely used informal assessment data tools. In terms of design skills, some participants were surprised at the processes and the hidden work needed to assemble an exam paper before its final production. Furthermore, some participants asked for the training materials to be shared after the workshops.

During the group discussions in the workshops, the researchers observed that the participants were debating, which indicates overlapping roles such as classroom teachers, test writers, course coordinators, and members of the exam committee. The researchers also recorded the following observations during the training sessions or later when they were mentioned by the participants.

> Next time, I am going to align learning outcomes, assessment plan in my exam proposal.
> It is a fruitful chance for developing test skills especially alignment of learning outcomes with questions.
> Today, I got new understanding to testing … I hope to participate in similar workshops.
> It is a perfect chance for triggering the essence of test interpretation.
> I learnt from the discussion to improve the current situation to widen skills about testing and framing a balanced fair test to assess learners.
> The exam questions should be based on learning outcomes.
> I have to keep in mind Excel measurement when designing a test proposal.

The most challenging part for the participants was aligning learning outcomes, content, and the assessment plan. For test designing, the participants showed that they learnt how to construct good test structures. The participants also realized that they had different roles and shared responsibility in language assessment. They also realized that students' results are reported in several ways for different purposes. Students' results are reported to the competitive programs based on the highest scores where program of medicine, for example, selects the highest level of students. However, students' results are reported to other programs as 'pass' and 'fail'

The researchers also observed that the allocated time was not sufficient to address the needs for instructional skills and design skills. During the discussion sessions, the researchers had to stop the group work and collect data while the participants still needed more time to discuss the issues of the workshops. To overcome that challenge, the researchers shared the training materials with the participants. They also shared online tutorials for them to watch during their free time. During the workshop on measurement skills, the participants faced some technical problems with the Excel software including opening shared worksheets, and adding and calculating mathematic formulas.

The analysis of the workshop evaluation forms shows that the participants were satisfied and engaged with the new experience (Table 2).

Table 2
Satisfaction and Engagement with the Workshops

No.	Level of satisfaction and engagement	5	4	3	2	1
1.	I am satisfied with the topics, the tasks, and the structure of the workshop.	58%	32%	10%	0	0
2.	I have interacted and participated in the workshop.	45%	39%	16%	0	0
3.	The workshop is appropriate for me.	87%	13%	0	0	0
4.	I have developed new understandings of testing.	81%	19%	0	0	0
5.	I have improved in terms of testing practices as a result of the workshop.	58%	23%	13%	6%	0
6.	The skills practiced during the workshop can be applied in language testing.	94%	6%	0	0	0
7.	The skills practiced during the workshop can be transferred to my evaluation and marking situations.	87%	13%	0	0	0
8.	The information can be transferred to interpreting and reporting test results.	81%	19%	0	0	0

Note. Adapted from "Assessment literacy for teachers: A pilot study investigating the challenges, benefits and impact of assessment literacy training", by E. Boyd and D. Donnarumma, 2018, pp.123-125, *Teacher involvement inhigh-stakes language testing* (pp. 105-126). Springer International Publishing (http://dx.doi.org/10.1007/978-3-319-77177-9_7). Copyright 2018 by Springer International Publishing

This table shows that the participants built their confidence in testing. Most of the participants reported that the information gained during the workshop could be applied in language testing. They also showed that the workshops were appropriate for them, and the skills practiced during the workshops could be transferred to their evaluation and marking situations.

Enhancing LASs

The participants improved their LASs. The independent t-tests of the pre-treatment and post-treatment (see appendix) showed that the participants increased their instructional skills, design skills, and skills of educational measurements to 4.25, 4.21, and 4.04; respectively. Similarly, the paired t-test showed significant differences between the pre-treatment and the post-treatment at a 95% confidence interval of the differences (Table 3).

Table 3
Paired Differences of the T-Test

Pre-treatment – Posttreatment of LASs	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Dev.	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper			
Instructional Skills	-2.38440	.31262	.09025	-2.58303	-2.18577	-26.421	31	.000
Design Skills	-2.26345	.41716	.12042	-2.52850	-1.99840	-18.796	31	.000
Skills of Educational Measurement	-1.62901	.42059	.14870	-1.98064	-1.27739	-10.955	31	.000
LASs	-2.15020	.48236	.08527	-2.32411	-1.97629	-25.216	31	.000

Table 3 shows the differences in the means and the statistical levels of significance of the pre-treatment and post-treatment of the LASs. Instructional skills and design skills scored very high means, 2.38 and 2.26, with a standard deviation of .312 and .417 respectively. Skills of educational measurement scored a high level, 1.62, with a standard deviation of .420. All in all, the paired test revealed the appearance of very high significances at the level of each LAS and the total LASs (.000).

The independent analysis of the post-treatment questionnaire The results of the treatment in the appendix shows that the participants improved instructional skills.

> I can improve instruction based on assessment results and feedback. (57%)
> I can provide feedback on students' assessment performance. (54%)

On the one hand, these two items indicated the highest level of improvement. On the other hand, the analysis showed that the participants need further training on using technology to assess students as in the following item *I can incorporate technologies in assessing students (35%)'*. They also need to be trained to use alternative means such as the students' portfolios of language assessment as was indicated in the analysis of the following item. *I can utilize alternative means for assessment, for example, portfolios. (41%)'*.

> The independent analysis also showed the participants improved skills for designing tests (See appendix).
> *I can write selected-response items such as multiple-choice, true-false, and matching. (54%)*
> *I can clearly define the language construct(s) a test will give information about. (53%)*

These two items indicated that the participants highly improved in terms of item writing and language constructs. On the contrary, the participants were still in need of workshops for raters and for designing rubrics for alternative means of assessment as can be seen in the following least developed items.

> I can design workshops for raters, whenever necessary. (22%)
> I can design rubrics for alternative assessments such as portfolios and peer assessment (39%)

The participants also improved their skills in educational measurement (See appendix). As evidenced by the following two items, the participants can infer students' points of strengths and weaknesses based on the analysis of questions and learning outcomes. They also benefitted from the shared tutorial links about the use of Excel in language assessment.

I can infer students' strengths and weaknesses based on data. (45%)
I can use internet tutorials for particular language assessment needs. (43%)

However, the participants were still in need of critical analysis on external tests as evidenced by the following item analysis. *I can criticize external tests and their qualities based on their psychometric characteristics (22%).* This item reported the least developed measurement skill. Unexpectedly, the participants reported that they needed to use Excel for language assessment as it was evidenced by the analysis of the following item.

I can use software such as Excel for language assessment. (24%)

This result can be attributed to the nature of work in a PY program where the exam committee interprets and reports assessment data. The role of faculty members lies in inserting the data into the Excel sheets prepared by the exam committee. It seems that the participants need to know more about how Excel sheets are prepared and designed.

Discussion

This study aimed to scrutinize the impact of need-based training on LASs among faculty members in the PY program of Saudi University. First, it identified levels of LASs through a needs analysis. Then it designed a training program based on those reported needs; something to engage and motivate the participants regarding language assessment. Finally, it aimed to enhance LASs among the participants. This section revisits these three issues.

The nature of language assessment was a challenging endeavor for a PY program. Shifting the assessment to be based on learning outcomes had great effect on shaping the LASs among the participants. The needs analysis report showed that the participants lacked training on instructional skills, design skills, and skills of educational measurement. This need coincided with previous studies in the Saudi context. For example, Mansory (2016) found that EFL teachers at the Language Institute of Saudi University might be excluded from language assessment due to their need for developing LASs in technical areas, which necessitates technical preparation and training that can usually be expensive and time consuming. The study recommended long-term collaborative training and support that would greatly help in raising the assessment literacy among teachers. Similarly, Hakim (2015) reported poor practicing assessment techniques among EFL teachers at an English language institute in a Saudi university. The needs analysis was also in agreement with Hazaea and Tayeb (2018) who found that washback had the least effect on content assessment among the four factors of language teaching.

The present needs report is also supported by studies from European countries. On examining the training needs among language teachers in Europe, Vogt and Tsagari (2014) showed that teachers need training in language assessment. Because of inadequate training, teachers either learn about assessment at the workplace or assess their students with ready-made tests. Similarly, Boyd and Donnarumma (2018) found that EFL teachers need to be aware of the importance and impact of testing. They also need training on test design.

In Turkey, Öz and Atay (2017) suggested ongoing professional development sessions for EFL instructors to raise their awareness and to improve their performance in their own assessment. The same study found an imbalance between assessment literacy and classroom reflection. Hatipoğlu (2015) explored the role of assessment experience on EFL teachers' needs relevant to LAL.

In other countries, Lan and Fan (2019) found that EFL teachers in China were underdeveloped, especially in technical skills and language pedagogy. In Hong Kong, Lam (2015) concluded that potential research could investigate how in-service language teachers developed their LAL via workshops especially during the first few years of teaching. Sultana (2019) found that the insufficient academic background of assessment among Bangladeshi EFL teachers hindered their assessment skills.

Moving to another issue, the data analysis revealed that the participants engaged in the treatment. The training program moved the stagnant waters of language assessment for the participants. They participated actively in the workshops. These findings coincided with those reported by Boyd and Donnarumma (2018) who found that tutors were confident about testing after training sessions on principles of language assessment. In an assessment training program for EFL teachers in Egypt, the participants reported high levels of agreement regarding the program outcomes. The participants also stated that the workshops were very useful (Gebril & Boraie, 2016). On the contrary, some disagreements and negotiations emerged during the group discussions due to the overlapping roles of the participants. These findings coincide with Malone (2017), who reported certain challenges including disagreements between language testers and teachers about the topics to be covered in an exam paper.

After the treatment, the data analysis shows that the training program was effective in terms of enhancing LASs (2.15). As the paired t-test showed, instructional skills (2.38) and design skills (2.26) scored a very high level of significance. Skills on educational measurements (1.62) also scored a high level of significance. These findings coincided with the existing research. Walters (2010) found that “prior formal training in test specification writing (or lack of it) affected [Standard Reverse Engineering] processes and outputs” (p.337). Umer et al. (2018) found an overlap between course learning outcomes and question papers in an English language program at Saudi University, which can lead to surface-level learning among students. The findings also coincided with Xie and Lei (2019) who proposed instructional assessment strategies with the L2 writing instruction processes. Chen and Zhang (2017) also found that feedback from teachers played a vital role in EFL writing.

Yet, the participants still need further training on online assessment, assessing students’ portfolios, designing assessment workshops and rubrics, criticizing external tests, and programming Excel sheets for language assessment. These results suggest that training without practice is not effective. Some quizzes are sent online through the Blackboard, but online assessment where faculty members use Blackboard, for example, to generate questions, to give immediate feedback to students is not practiced in the PY program. That is why the participants feel that they need further training on online assessment. Similarly, it seems that the participants never use student portfolios in their assessment.

Guided by the LAL approach, the present study focused only on language teachers. On the contrary, students are still an important, yet neglected, stakeholder in language assessment. As teaching has moved into student-centered learning, student-centered assessment (Baranovskaya & Shaforostova, 2017) is still under-investigated among LAL theorists. For better language learning, recent research has shown the importance of involving students in their ‘alternative’ language assessment. For example, students can participate in their continuous assessment using student portfolios (Llarenas, 2019), peer assessment (Stognieva, 2015), self-assessment (Baranovskaya & Shaforostova, 2017), Olympiads (Bolshakova, 2015) and roundtable discussions (Rodomanchenko, 2017) to name a few.

Conclusion

This context-specific study reports the levels of LASs among EFL faculty members and the effectiveness of professional training on LASs for designing and interpreting tests. It first identified three levels of LASs, and then addressed those needs in the context of professional development. The needs analysis report showed that the participants need training in the three overlapping areas of instructional skills, design skills, and educational measurement skills. During treatment, the study found that the participants reported their confidence and satisfaction with language assessment. The analysis of the post-treatment questionnaire showed that the participants improved their instructional skills, design skills, and skills of educational measurements.

These improvements could have positive effect on the participants’ assessment practices. These findings imply constructive, valid, and fair assessments for PY students. These students are the indirect beneficiary of this experiment. Students’ voices also need to be heard in language assessment. In a student-centered environment, students can be involved in pre-test preparation. For example, once students produce their own quizzes, classroom teachers can better understand the way students think of the exam and can address their needs in language assessment accordingly. In the context of high-stakes tests, students always ask about the nature of the exam. Such questions can be answered in the context of learning outcomes. In other words, when students ask about the nature of the test, faculty members need to highlight the intended learning outcomes of the courses. However, further research is recommended on enhancing LASs among students.

If we were to repeat the training program, we would make several adjustments. First, we would include other stakeholders such as EFL schoolteachers, students, and the coordinators of the concerned programs that stipulate PY high scores before admission. Second, we would invite experts from outside the university to share in training faculty members. We would sort out the need for assessment skills for reading, vocabulary, writing, speaking, and listening. The allocated time for the training program was not enough to accommodate the busy schedules of faculty members. An intensive program on design and instructional skills could be better implemented during a semester break. A remedial training program could also be implemented to re-address the skills that still need further training and discussion. Other researchers could also employ a similar training program to their unique teaching contexts.

Acknowledgements

We would like to thank the Ministry of Education in Saudi Arabia and the Deanship of Scientific Research at Najran University for their financial funding and technical support for this project in the eighth research phase grant code No. NU/ SHED/16/109. Our gratitude extends to the editorial board and the anonymous reviewers of the manuscript for their constructive comments.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Baranovskaya, T., & Shaforostova, V. (2017). Assessment and evaluation techniques. *Journal of Language and Education, 3*(2), 30-38. https://doi.org/10.17323/2411-7390-2017-3-2-30-38

Binet, A., Gavin, V., Carroll, L., & Arcaya, M. (2019). Designing and facilitating collaborative research design and data analysis workshops: Lessons learned in the healthy neighborhoods study. *International journal of environmental research and public health, 16*(3), 324-339. http://dx.doi.org/10.3390/ijerph16030324

Bolshakova, E. (2015). Olympiad in the English language as a form of alternative language assessment. *Journal of Language and Education, 1*(2), 6-12. https://doi.org/10.17323/2411-7390-2015-1-2-6-12

Boyd, E., & Donnarumma, D. (2018). Assessment literacy for teachers: A pilot study investigating the challenges, benefits and impact of assessment literacy training. In D. Xerri & P. V. Briffa (Eds.), *Teacher involvement inhigh-stakes language testing* (pp. 105-126). Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-77177-9_7

Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with uncertainty: Essays in honour of Alan Davies* (pp. 137-143). Cambridge University Press.

Chen, D. & Zhang, L. (2017). Formative assessment of academic English writing for Chinese EFL learners. *TESOL International Journal* , *12(2)*, 47-64

Crusan, D., Plakans, L., & Gebril, A. (2016). Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices. *Assessing Writing, 28*, 43-56. http://dx.doi.org/10.1016/j.asw.2016.03.001

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing,* 25(3), 327-347. http://dx.doi.org/10.1177/0265532208090156

Djoub, Z. (2017). Assessment literacy: Beyond teacher practice. In R. Al-Mahrooqi, C. Coombe, F. Al-Maamari & V. Thakur (Eds.), *Revisiting EFL assessment: Second language learning and teaching* (pp. 9-27). Springer. http://dx.doi.org/10.1007/978-3-319-32601-6_2

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132. http://dx.doi.org/10.1080/15434303.2011.642041

Galichkina, E. (2016). Developing teacher-trainees' assessment awareness in the EFL classroom through project-based learning activity. *Journal of Language and Education, 2(3),* 61-70. http://dx.doi.org/10.17323/2411-7390-2016-2-3-61-70

Gebril, A., & Boraie, D. (2016). Assessment literacy training for English language educators in Egypt. In V. Aryadoust & J. Fox (Eds.), *Trends in language assessment research and practice: The view from the Middle East and the Pacific Rim* (pp. 416-437). Combridge scholars publishing.

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile Issues in Teachers' Professional Development, 20*(1), 179-195. http://dx.doi.org/10.15446/profile.v20n1.62089

Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education & Literacy Studies, 3*(4), 42-48. http://dx.doi.org/10.7575/aiac.ijels.v.3n.4p.42

Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal, 4*(2), 111-128.

Hazaea, A. (2019). The needs on professional development of English language faculty members at Saudi University. *International Journal of Educational Researchers, 10*(1), 1-14

Hazaea, A. N., & Tayeb, Y. A. (2018). Washback effect of LOBELA on EFL teaching at preparatory year of Najran University. *International Journal of Educational Investigations, 5*(3),1-14

Herrera M., L., Macías, V., & Fernando, D. (2015). A call for language assessment literacy in the education and development of teachers of English as a foreign language. *Colombian Applied Linguistics Journal, 17*(2), 302-312. https://doi.org/10.14483/udistrital.jour.calj.2015.2.a09

Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. Language testing, *29*(3), 395-420. http://dx.doi.org/10.1177/0265532211428317

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402. http://dx.doi.org/10.1177/0265532208090158

Inbar-Lourie, O. (2012). Language assessment literacy. In C. Chapelle (Ed.), The encyclopedia of applied linguistics (pp. 1-9). John Wiley & Sons. http://dx.doi.org/10.1002/9781405198431.wbeal0605

Khan, B., & Begum, S. (2012). Portfolio: A professional development and learning tool for teachers. *International Journal of Social Science and Education, 2*(2), 363-377.

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197. http://dx.doi.org/10.1177/0265532214554321

Lan, C., & Fan, S. (2019). Developing classroom-based language assessment literacy for in-service EFL teachers: The gaps. *Studies in Educational Evaluation, 61*, 112-122 https://doi.org/110.1016/j.stueduc.2019.1003.1003.

Language Assessment Literacy Symposium. (2016). Enhancing language assessment literacy: Sharing, broadening, innovating. *Language Assessment Literacy.* URL http://wp.lancs.ac.uk/ltrg/files/2015/12/Language-Assessment-Literacy-Symposium-2016-full-programme.pdf

Llarenas, A.T. (2019). Vignettes of experiences in portfolio assessment. *Asian EFL Journal, 21*(2), 196-223

Malone, M. E. (2017). Training in language assessment. In E. Shohamy, L. G. & S. May (Eds.), *Language Testing and Assessment, Encyclopedia of Language and Education* (vol. 3, pp. 225-239). Springer. http://dx.doi.org/10.1007/978-0-387-30424-3_178

Mansory, M. (2016). *EFL teachers' beliefs and attitudes towards English language assessment in a Saudi university's English language institute* [Unpublished doctoral dissertation]. University of Exeter. URL http://hdl.handle.net/10871/25765

Ørngreen, R., & Levinsen, K. (2017). Workshops as a research methodology. *Electronic Journal of E-learning, 15*(1), 70-81.

Öz, S., & Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal, 6*(1), 25-44.

Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing, 30*(3), 381-402. http://dx.doi.org/10.1177/0265532213480337

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462. http://dx.doi.org/10.1177/026553220101800407

Rodomanchenko, A. (2017). Roundtable discussion in language teaching: Assessing subject knowledge and language skills. *Journal of Language and Education, 3*(4), 44-51. https://doi.org/10.17323/2411-7390-2017-3-4-44-51

Stabler-Havener, M. L. (2018). Defining, conceptualizing, problematizing, and assessing language teacher assessment literacy. *Working Papers in Applied Linguistics and TESOL, 18*(1), 1-22.

Stognieva, O. (2015). Implementingp peer assessment in a Russian university ESP classroom. *Journal of Language and Education, 1*(4), 63-73. https://doi.org/10.17323/2411-7390-2015-1-4-63-73

Sultana, N. (2019). Language assessment literacy: an uncharted area for the English language teachers in Bangladesh. *Language Testing in Asia, 9*(1), 1-14. http://dx.doi.org/10.1186/s40468-019-0077-8

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21-36. http://dx.doi.org/10.1017/S0267190509090035

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403-412. http://dx.doi.org/10.1177/0265532213480338

Umer, M., Zakaria, M. H., & Alshara, M. A. (2018). Investigating Saudi University EFL teachers' assessment literacy: Theory and practice. *International Journal of English Linguistics*, 8(3), 345-356 http://dx.doi.org/10.5539/ijel.v8n3p345

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374-402.http://dx.doi.org/10.1080/15434303.2014.960046

Walters, F. S. (2010). Cultivating assessment literacy: Standards evaluation through language-test specification reverse engineering. *Language Assessment Quarterly, 7*(4), 317-342. http://dx.doi.org/10.1080/15434303.2010.516042

Xie, Q., & Lei, Y. (2019) formative assessment in primary English writing classes: A case study from Hong Kong. *Asian EFL Journal, 23*(5), 55-95

Appendix

Language Assessment Skills

Statement	Pre-treatment		Post Treatment		Improv.
	Means	Std. Dev	Means	Std. Dev	Percent
Instructional Skills	1.868283	0.939779	4.252683	0.797118	47%
I can improve instruction based on assessment results and feedback.	1.5161	.81121	4.3548	.75491	57%
I can provide feedback on students' assessment performance.	1.5484	.76762	4.2581	.77321	54%
I can provide motivating assessment experiences, giving encouraging feedback.	1.5806	.76482	4.2581	.81518	54%
I can collect formal data (e.g. through tests) on students' language development.	1.6774	.79108	4.3226	.70176	53%
I can plan, implement, monitor, record, and report student language development.	1.8710	.84624	4.3548	.75491	50%
I can align learning outcomes, instruction, and assessment.	1.8710	.92166	4.3226	.74776	49%
I can collect informal data (e.g. while observing in class) on students' language development.	1.9677	.94812	4.2903	.82436	46%
I can use language assessment methods appropriately.	1.8387	1.00322	4.1613	.82044	46%
I can use multiple methods of assessment to make decisions based on substantive information.	2.0323	1.01600	4.2581	.92979	45%
I can communicate test results to a variety of audiences: students, department, and deanship.	2.0000	1.09545	4.1290	.84624	43%
I can utilize alternative means for assessment, for example, portfolios.	2.0968	1.19317	4.1290	.76341	41%
I can incorporate technologies for assessing students.	2.4194	1.11876	4.1935	.83344	35%
Design Skills	1.970664	0.990725	4.212367	0.820873	45%
I can write selected-response items such as multiple-choice, true-false, and matching.	1.7742	.95602	4.4516	.72290	54%
I can clearly define the language construct(s) a test will give information about.	1.6774	.79108	4.3226	.74776	53%
I can clearly identify and state the purpose for language assessment.	1.7097	.78288	4.2258	.84497	50%
I can design assessments that are valid not only in terms of course contents but also course tasks.	1.7419	.81518	4.1935	.79244	49%
I can improve test items after item analysis, focusing on items that are either too difficult, too easy, or unclear.	1.8065	1.01388	4.2581	.89322	49%
I can design assessments that are reliable, authentic, fair, ethical, practical, and interactive.	1.8710	1.02443	4.2258	.84497	47%
I can design constructed-response items (for speaking and writing), along with rubrics for assessments.	1.8710	.95715	4.1613	.82044	46%
I can construct test specification to design parallel forms of tests.	1.9032	1.01176	4.1613	.68784	45%
I can write test syllabuses to inform test users of test formats, where applicable.	1.9677	1.07963	4.2258	.84497	45%
I can provide security to ensure that unwanted access to tests is deterred.	1.9032	1.07563	4.0968	.94357	44%
I can design rubrics for alternative assessments such as portfolios and peer assessments.	2.1935	1.07763	4.1613	.77875	39%
I can design workshops for raters, whenever necessary.	2.9677	1.30343	4.0645	.92864	22%
Skills of Educational Measurement	2.415338	1.155141	4.04435	0.949653	33%
I can infer students' strengths and weaknesses based on data.	2.0000	.89443	4.2258	.88354	45%
I can use internet tutorials for language assessment needs.	2.1613	1.15749	4.2903	.90161	43%
I can interpret data related to test design such as item difficulty and item discrimination.	2.2903	1.21638	4.1290	.88476	37%
I can calculate reliability and validity indices by using appropriate methods such as Excel spreadsheets.	2.2581	1.18231	4.0000	.89443	35%
I can calculate descriptive statistics such as means and curves, reliability, and correlations.	2.4839	1.15097	3.8710	1.14723	28%
I can investigate facility and discrimination indices statistically.	2.5484	1.26065	3.9677	.87498	28%
I can use software such as Excel for language assessment.	2.6452	1.22606	3.8387	1.09839	24%
I can criticize external tests and their qualities based on their psychometric characteristics.	2.9355	1.15284	4.0323	.91228	22%

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Nurie, Y. (2020). Graduate Students' Perceived Needs and Preferences for Supervisor Written Feedback for Thesis Writing. *Journal of Language and Education, 6*(4), 153-170. https://doi.org/10.17323/jle.2020.10340

Graduate Students' Perceived Needs and Preferences for Supervisor Written Feedback for Thesis Writing

Yenus Nurie

Bahir Dar University

Correspondence concerning this article should be addressed to Yenus Nurie, Faculty of Humanities, Bahir Dar University, P.O. Box 79, Ethiopia. Email: yenusn@gmail.com

A plethora of previous research has explored students' preferences for written feedback from teachers to respond to students' writing in the classroom. However, little or no research has investigated graduate students' needs and preferences regarding written feedback provided by their supervisors in response to thesis writing. This study examined the feedback needs and preferences of EFL graduate students to the three nominated themes defining supervisor written feedback to thesis writing in this study: content, genre, and linguistic feedback to thesis writing. Data was collected from 32 master's students from the TEFL and Media and Communications at Bahir Dar University, Ethiopia using a questionnaire and an unstructured interview. The participants' responses were tabulated and analyzed using descriptive statistics. The results indicated that both groups commonly preferred feedback on content the most. When sub-categories of this feedback were examined further, it was found that TEFL students favored gaps in theoretical understanding, but Media and Communications students preferred coverage and gaps in the literature. They also showed discrepancies regarding their preferences for part-genres in thesis writing and the various features of linguistic accuracy. The results of this study suggest that feedback on theses should be realized in regard to the needs and preferences of graduate students. Finally, implications for further research that could shed light on the resonant understanding of feedback on thesis writing are presented.

Keywords: thesis, feedback preference, content feedback, genre, linguistic feedback

Introduction

Written feedback on a student thesis is a critical aspect of supervision in higher education. However, it could also prove to be a stumbling block to the success of thesis/dissertation if supervisors do not consider their students' perceived needs and preferences for various forms of written feedback. Supervision in a postgraduate program is of paramount importance in higher education where supervising students on the research projects they conduct as a requirement for their graduation is one of the responsibilities of professors all over the world (Yinager, 2019). Feedback is interactive and involves continuous two-way communication that encourages teaching and learning among educators and students (Manjet, 2016). Indeed, the role of feedback in facilitating student learning has been advocated by many educational researchers (Aitchison, Catterall, Ross, & Burgin, 2012; Evans, Hartshorn, McCollum, & Wolfersberger, 2010; Pokorny & Pickford, 2010; Stracke & Kumar, 2010).

The mounting evidence on the role of feedback has also attracted considerable attention in higher education where effective and high quality feedback is found to be a key element of quality teaching in higher education (Ali, Watson, & Dhingra, 2016; Basturkmen, East, & Bitchener, 2014; Ghandi, & Maghsoudi, 2014; Hoomanfard, Jafarigohar, Jalilifar, & Hosseini, 2018; Leng, 2014; Wang & Jiang, 2015). Another interesting recent development in the field of writing is that while the value of feedback in teaching and learning is consistently reiterated in educational literature, a considerable amount of research has also provided important assertions about the critical role of feedback in helping graduate students determine the parts of writing that are correct and the parts that need correction for improved writing as well as alerting students of their strengths and weaknesses in academic writing (Manjet, 2016). Thesis writing is a daunting experience for all graduate students who conduct research for the first time. Consequentially, this insight has significant pedagogical implications for supervisors of thesis writing, suggesting that instruction should be usually given either on written or oral feedback or both by supervisors when students are writing a thesis (Bitchener, Basturkmen, & East, 2010).

Given the multifarious functions of feedback on thesis writing discussed in the preceding paragraphs, the development of students' higher-order and critical thinking skills is an important one. Feedback helps in stimulating critical thinking among students to ensure the development of constructive ideas for writing (Manjet, 2016). As supervisors provide feedback, they teach their students a gradual release of responsibility that includes instilling higher-order transformative skills such as reflecting critically or interpreting meanings logically by assessing a thorough evaluation of their own work independently. In such a supportive environment, the students can begin to consider the real-life implications of the insights provided by their supervisors by inculcating them in their thinking. This will help them find solutions for the problems they encounter during the thesis writing process. Feedback is indispensable in teaching and learning and its role is documented in educational literature (Price, Handley, Millar & O'Donovan, 2010; Yu & Lee, 2013). In addition, Rowe and Wood (2008) posit that in order to achieve quality teaching and supervising, effective and quality feedback should be provided. Taken together, written feedback by the graduate students' supervisors and lecturers is a fundamental source of input for academic writing such as thesis writing (Bitchener et al., 2010).

At this juncture, it is essential to make clear distinctions between the two concepts related to supervisor feedback on thesis writing: effective and quality feedback on thesis writing. Quality and effective education has recently been an increasingly important concern in universities that policymakers, teachers, and other stakeholders are struggling to achieve. Although they seem to be similar, it is essential to make clear distinctions between the terms 'effective' and 'quality' as they are used in the context of this study. Education experts differ widely on meaning of the terms 'effective' and 'quality' feedback in regard to thesis writing in higher education, but this study defines 'quality' as any typical feedback used to attain the desired objectives of supervision on thesis writing. Similarly, the term 'effective feedback' refers to various means of supervision that are conducted with skills and strategies used for the ultimate goal of the supervision. Using the terms interchangeably, some researchers posit that in order to achieve quality teaching and supervising, effective and quality feedback should be provided (Rowe & Wood, 2009).

One possible reason for the effectiveness of feedback can be ascribed to addressing the needs and preferences of students for supervisor feedback to thesis writing. It is also recognized that understanding students' perceived needs of feedback can be helpful for evaluating the appropriateness and relevance of current supervisory practices in thesis writing. A considerable amount of literature has discussed various types of feedback, such as direct and indirect feedback, peer and teacher feedback, oral and written feedback, and grammar and content feedback. Although a large body of research has explored these types of feedback, little is still known about graduate students' perceived needs regarding the three aspects of feedback in response to thesis writing: Content knowledge: its accuracy, completeness and relevance, Genre knowledge: the functions of different parts of a thesis, and Linguistic accuracy and appropriateness. Further, in spite of differences across various disciplines, the two major challenges that have hindered graduate students when writing their theses are a limited understanding of the characteristics of the thesis genre and its component parts (part-genres like the introduction and discussion sections/chapters) and uncertainty about the expectations and requirements of their discipline-specific communities of practice (Bitchner et.al., 2010).

Having established the appropriateness of the three themes defining supervisor written feedback for thesis writing nominated in this study, it is then important to understand the ways of delivering feedback on student thesis writing. Instruction should be usually given either in written or oral forms. This is because students will benefit twice when they receive both oral and written feedback. Oral feedback, such as teacher-student conferencing, is deemed necessary as it sheds some light on students' misunderstandings and difficulties while the supervisor discusses the written comments provided. The importance of written feedback is also acknowledged in thesis writing contexts since written feedback helps the student improve their academic writing. A study by Bitchener, Young, and Cameron (2005) found that combining written and oral feedback made for significant improvements in student writing over time, but it was oral feedback that had the added potential of allowing constructive comments that could meet the students' individual needs. Consequentially, a large body of research results also revealed that there is a growing trend that believes such blended forms of supervisory feedback should be given simultaneously. Supervisors almost always provide written feedback on their students' thesis drafts (Tee, Kumar, & Abdullah, 2013), and the written feedback provided is usually supported with oral feedback through face-to-face meetings (Bitchener & Basturkmen, 2006; Bitchener, Basturkmen, East, & Meyer, 2011; Bitchener et al., 2005).

A remarkable variant that has been emphasized in the literature regarding supervisor feedback in response to student thesis writing is students' preference variations towards the value of feedback to thesis writing. In connection with written feedback, for instance, various characteristics of feedback preferences have been reiterated in the literature (Bitchener et al., 2010). Previous research has also asserted that students value and appreciate written feedback on their grammar (Bitchner et al., 2011; Manjet, 2016). On the other hand, research has also reported various features of content or subject-area feedback as graduate students' most preferred feedback (Chokwe, 2015; Lee, 2008). Based on the interview results obtained from a similar English as a foreign language context, three factors (genre knowledge, content knowledge, and appropriation) were found to affect Iranian TEFL graduate students' perceived needs on the type of written feedback (Hoomanfard, Jafarigohar, Jalilifar, & Hosseini, 2018).

Irrespective of these variations, one of the most notable assertions is that feedback for thesis writing can only be viewed as an important process if supervisors carefully consider the salient features of each feedback type in light of the literature when providing feedback. Concerning this, Hyland (2009) posits that the most helpful feedback is that which helps students understand the expectations of their disciplinary community. It "conveys implicit messages" about the values and beliefs of the discourse community, the nature of disciplinary knowledge, and student identities in the community (Hyland, p. 132). Bitchner et al. (2011) also suggested coherence and cohesion in argument creation as an example of graduate students' writing difficulties in thesis writing. Moreover, students valued feedback when they perceived that the feedback providers believed in their potential, cared about the improvement of their skills, and tried to be helpful (Can & Walker, 2011).

Literature Review

Theoretical Framework

This study is informed by the premise of the socially mediated learning theory as it emphasizes the essence of social interaction and mediation. Vygotysky (1978) postulates that the potential for the cognitive development of a learner is limited to the Zone of Proximal Development (ZPD). According to Vygotysky (1978), this 'Zone' According to Vygotysky (1978), this 'Zone' is an important area that the student is cognitively prepared to examine things carefully with the help of a more experienced other to fully develop. Bitchener, et.al (2011) maintained that feedback helps students by enhancing the supervisor-supervisee relationship to help it take on a more egalitarian «peer-to-peer» relationship, with the former being the more capable peer whose role is to scaffold and negotiate the ZPD to ultimately develop as an independent thesis writer.

As the more experienced other, the supervisor can give provide scaffolding and other strategies to enhance learner curiosity until the learner develops complex skills. Further, the theory underscores the relevance of the holistic approach toward teaching (e.g. grammar should be taught not as a discrete element but it must be taught holistically to help the learner construct meaning). The fact that the theory underscores the importance of an integrated approach toward teaching helps the socially mediated learning theory of learning fit the purpose of this research, which is specific to thesis writing. Thus, supervisors should provide one form of feedback not just as a discrete element, but it must be provided holistically using various supervisor feedback mechanisms to thesis writing that help students elucidate their intellectual knowledge and research skills.

Taken together, this research was guided by the socially mediated learning theory as it underlines the essence of social interaction and mediation that aims to mediate how students think, suggesting that supervisory written feedback could act as remedial instruction to develop students' problem solving skill in thesis writing. This skill can be developed under the guidance of the supervisor, and hence the theory fits the purpose of this research as it sheds light on significance of a supervisor as someone who can support the learner with the development of complex skills through modeling, collaborative learning, and independent problem solving techniques to foster the intellectual knowledge and skills of learners through scaffolding and mediation.

Supervisors' Written Feedback Types

In the literature, various types of feedback are highlighted. For instance, on the basis of the amount of feedback provided for student writing, written corrective feedback can be comprehensive (unfocused), which refers to teacher's feedback in response to all of the errors committed by the student, a common time-consuming practice used by teachers (Lee, 2004, 2008), or selective (focused), where the teacher provides written feedback selectively on a number of linguistic features, such as sentence style, active/passive, verb tense, etc. Other forms include direct feedback, referring to the provision of the correct form or structure by the teacher (Ferris, 2003), or indirect feedback, referring to the indication of an error by the teacher through circling or underlining an error, indicating the number of errors in the margin, or placing a code such as SP for spelling or WW for wrong word (Ferris & Roberts, 2001). However, as the purpose of this study is to explore the perceived written feedback needs of postgraduate students, this study focuses on three distinct feedback types provided by supervisors in response to student thesis writing; these are feedback on content, feedback on genre, and feedback on linguistic accuracy and appropriateness. A detailed account of these categories will be elaborated on in the methodology section.

Despite the fact that feedback constitutes a major form of instruction for higher-degree research students, the general focus of advisors has been reported as the struggle to articulate implicit knowledge (Paré, 2011). Micro-levels errors include ambiguity in references, the misuse or over use of conjunctions and repetition, the misuse of lexical items, etc., and written feedback on a 'micro-level' refers to the types of feedback given to treat learner errors with regard to learners' use of grammatical accuracy and appropriate word order patterns, as well as acceptable grammatical systems and forms (e.g., tense, agreement, pluralization), patterns, and rules (Brown, 2007). With regard to this, having analyzed the considerable discrepancies in the techniques of teacher error correction in his stud. Similarly, Leng (2014) conducted a study analyzing the written feedback on ESL students' written assignments to shed light on how feedback acts as a type of written speech between the lecturer and student. The results from the study indicated that the written feedback provided to students was helpful and useful in their essay revision. The study concludes that the feedback was effective for the students because they were able to attend to the revision of their second draft well.

Graduate Students' Written Feedback Preferences

The growing amount of research on supervision practices has revealed graduate students' written feedback preferences for supervisor written feedback in response to thesis writing. One of the other functions of feedback that has been proposed in the literature is part-genre, which is feedback offered to comment on students' theses regarding the different purposes, structures and organization, and parts of a thesis (Bitchner, et al., 2010). Manjet (2016) indicated that graduate students preferred the content of their academic work to be emphasized in the feedback and not merely focus on language and formatting issues.

In a similar vein, Hooman Saeli (2019) employed semi-structured interviews to obtain data from 14 teachers and 15 students to investigate teachers' practices and students' preferences regarding grammar-centered written corrective feedback (WCF) in an Iranian EFL context. The study revealed that the teachers' practices were driven by the students' preferences in that they mostly provided teacher-generated grammar feedback simply because they believed that their students preferred this type of correction. Esra (2020) also examined mentoring practices and mentor growth areas as perceived by three student-teachers in a 12-week practicum period in a Turkish English Language Teaching (ELT) context using interviews and reported that mentors rarely employed practices related to feedback, pedagogical knowledge, and system requirements.

Regarding postgraduate students' preferences about supervisors' written feedback, Hoomanfard, et al. (2018) used a mixed-method design to examine Ph.D. and MA students' needs for supervisor written feedback on their theses/dissertations. The results indicated that there were similarities (argument, logical order, transition, clarity, and references decisions) and differences (inclusion of information, formatting, grammar, conclusion, introduction, and consistency) between the priorities given by MA and Ph.D. students. With a particular emphasis on MA students, the results showed that the MA students' expressed priorities that were not similar to those of the supervisors except in three areas (argument, formatting, and grammar). The survey study of Can and Walker (2014) also revealed the Ph.D. students' feedback needs in order of priorities as arguments, the conclusion, clarity and understandability of the statements, inclusion or exclusion of information, the introduction, consistency in the overall paper, logical order of ideas, transition, paragraphs or sections, grammar, formatting, and references. The study further reported that Ph.D. students' ratings were lower for more mechanical aspects such as grammar, formatting, and referencing.

Zhan (2016) also investigated a teacher's and her students' preferences regarding written teacher feedback in a college English as a foreign language (EFL) writing class in China. The study used essays, questionnaires for 62 students, and interviews from the teacher and her six students to identify the types of feedback given by the teacher, the perceptions and preferences of the students, and the perceptions of the teacher. The findings of the study indicated that while the teacher's written feedback covered content, organization, vocabulary, grammar, and mechanics, the students reported that they benefited from and preferred feedback on organization the most, which was not expected by their teachers. Consequently, the study suggested that foreign language writing teachers communicate more with their students about their feedback practices, and be aware of students' perceptions and preferences so that their writing instruction could be more effective.

Significance of the Study

This study is based on the notion that understanding a great deal about student perceived needs and preferences is helpful because of the potential of this factor to boost the quality of written feedback in thesis writing. The results of the present study can hopefully familiarize EFL supervisors with the importance of considering the needs of their students during the thesis writing process. The findings of the study could also lay the foundations for new graduate school guidelines that could inform graduate supervisors about potential misunderstandings between supervisors and students, and help them understand how to make supervision in thesis writing effective.

This may also help to design better supportive mechanisms for supervisors in terms of scaffolding their students, which would ultimately produce more effective supervisory practices. Unlike other similar EFL studies that were conducted to investigate feedback preferences in classrooms (Amrhein & Nassaji, 2010; Ghandi & Maghsoudi, 2014), the present study examined the feedback needs of graduate students to thesis writing, which is an under-researched topic in a totally untouched context of supervision in universities in Ethiopia. All things considered, the results of this study could be significant in informing supervisors to critically analyze their students' writing preferences and adjusting how to balance these preferences in written feedback, thereby helping their students become more successful at thesis writing. It may also be used to provide insights into understanding the perceived needs of Ethiopian graduate students compared with the perceived preferences of graduate students in other EFL contexts.

Rationale for the Study

Two major interrelated points could primarily justify the need for further studies on feedback preferences. In Ethiopian universities, students who pursue master's degrees are required to write a thesis under the supervision of one or two advisors as a partial fulfillment for their master's degree program. Thus, it can be a challenge for these students without the support of the supervisors assigned to them. Moreover, despite the fact that research courses that serve as a vehicle to build students' research skills are offered in both undergraduate and graduate schools, student researchers have little or no capacity and understanding of the basic research components and skills necessary to undertake a study (Yenus, 2018). Local research has also indicated that many graduate students had no adequate writing skill when they started writing their theses (Zeleke, 2017).

The findings of numerous studies (Aitchison et al., 2012; Cotterall, 2011; Paré et al, 2011; Wang & Li, 2011) have showed considerable interest in postgraduate research writing, as little or no research has been conducted to specifically examine the perceptions of what graduate students think are the most important aspects of supervisor written feedback on thesis writing. In a similar vein, several studies have focused on teachers' written feedback to respond to their students' writing in the classroom rather than on student thesis writing (Katayama, 2007; Wang, 2010), indicating a dearth of research into writing practices at the postgraduate level. Therefore, this study was designed to respond to this need. In light of all this, the following research questions were formulated.

1. What are the perceived preferences of graduate students for their supervisors' written feedback?
2. Is there a significant difference in the mean preference scores between TEFL and Media and Communication Students?

Methodology

Research Design

This study adopted a multi-method approach so as to triangulate the data obtained from the students' survey questionnaire and the interview responses. The students' views regarding the needs and preferences of supervisory feedback were elicited initially from the responses of the survey questionnaire adapted from the exploratory, descriptive study of Bitchner et al. (2010) that investigated what supervisors said they focused on when giving feedback, which was followed by more in-depth questioning about the responses through the use of semi-structured interviews. Having established the data collection instruments used for this study, it is then important to understand the major feedback areas used to analyze the students' written feedback preferences.

The first theme that was derived regarding the types of supervisory written feedback to student thesis writing was feedback on content, which refers to comments on gaps in the literature, gaps in theoretical understanding and coverage, the potential irrelevance of the thesis, the wider significance of the work, and gaps in the justification or explanation of arguments. Part-genre feedback includes the structure of what goes where and examples of what is expected, purpose of part-genre, part-genre skills required for each genre and expected word count from a thesis constitutes the second major theme of the study. The third theme was a written feedback that contains punctuation, spelling and capitalization, systematic errors, writing style, how to develop ideas with supportive and counter evidence, writing coherence and cohesion, the quality of writing, grammar (verb tenses, subject/verb agreement, article use...etc.), and the appropriateness of vocabulary choice and academic register was also developed to help understand written feedback regarding the students' preferences for linguistic accuracy and appropriateness (Bitchener, et al., 2010).

Population and Participants

The target population of this study was graduate students who were pursuing a master's degree in various colleges at Bahir Dar University. The participants in the study were 20 English and 12 Media students who were currently enrolled at Bahir Dar University to earn their master's degree in Teaching English as a Foreign Language and Media and Communications respectively. Among the participants who were registered for the two disciplines, 68.8% were males and 31.2% were females. A total of 32 students participated in filling out the survey questionnaire in the faculty of humanities, where purposive sampling was used. Specifically, 14 (70%) of the participants in TEFL were males, and six (30%) were females, whereas eight (66.7%) of the participants in Media and Communication were males and four (33.3%) were females.

Research Procedures

The students were informed about the purpose of the questionnaire before they began filling out the form. Then, the researcher explained the instructions of the questionnaire that required them to rate each item based on their preferences regarding the various features of written feedback on thesis writing on a Likert scale ranging from "strongly agree" (5) to "strongly disagree (1)." After the participants were informed about the purpose of the questionnaire, the researcher distributed the questionnaire to the participants and asked that they be completed in 30 minutes. Finally, the researcher collected the questionnaires in person after making sure that the students completely filled them out. The results of the survey questionnaire were then used to analyze the participants' perceptual preferences for written feedback on thesis writing.

Data Collection Instruments

Data were obtained from a standardized questionnaire that was adapted by the researcher from Bitchner et al., 2010). The questionnaire consists of 17 items questions that were developed from an extensive literature review of published studies on various features of supervisory feedback for thesis writing. The Cronbach Alpha reliability of the needs for written feedback section of the questionnaire was .79. To ensure the consistency of the standard questionnaire in the context of the present study, a pilot study was carried out to test and improve the questionnaire used for this study. It was also reviewed by three experienced researchers to determine the content and face validity of the questionnaire. The researcher used the feedback to improve the questionnaire by removing unnecessary or repetitive statements. The survey instrument used for this study was also tested for reliability. The questionnaire was administered and pilot tested by other groups of students who were not part of the main study.

The pilot study helped the researcher revise unclear questions and make necessary changes for the administration of the questionnaire items before they were launched in the main study.

This standardized questionnaire was also field tested in the study of Yenus (2018) that examined 55 doctoral students' perceptions and preferences about supervisor written feedback to respond to student dissertations at Bahir Dar University, Ethiopia. In the study, to check the internal consistency of the reliability of the items, Cronbach's alpha coefficient was computed through SPSS version 20 and found to be 0.80. Meanwhile, since the level of the participants was still different from the study of Yenus (2018), the present questionnaire was again pilot tested and underwent a subsequent reliability test, which was found to be 0.82 Cronbach's alpha. The second instrument was an unstructured interview. Five TEFL students and three Media and Communications students were selected randomly for the interview. The purpose of the interview was to provide additional insights into the questionnaire results of this study and investigate their perceived needs and preferences regarding supervisor written feedback. The interviews were conducted after the responses obtained from students' questionnaire had been analyzed. The students' interview responses were recorded, categorized, and thematically analyzed.

Data Analysis

This study aimed to understand graduate students' feedback preferences for written feedback. Thus, the descriptive statistics results of the questionnaire were used to analyze the general feedback preferences of graduate students. An independent samples t-test was used to examine whether there was a significant difference between TEFL and Media and Communication graduate students regarding their feedback preferences. To conduct the independent samples t-test, the grand mean scores of both groups were compared to evaluate whether there was a statistically significant difference in their perceptions regarding written feedback. To be specific, the survey questionnaire aimed at examining feedback preferences of university students who enrolled in two disciplines regarding the various features of thesis writing as the content, part-genre, and linguistic accuracy and appropriateness. The students' responses obtained from the survey questionnaire were collected and imported into SPSS and calculated. The participants' responses to each item on the questionnaire were then summarized, tabulated, and analyzed using descriptive statistics such as mean and standard deviation and inferential statistics from an independent samples t-test. The interview data obtained from both the TEFL and Media and Communication students' explanatory responses was transcribed, coded, categorized into themes, and analyzed qualitatively.

Results

The results obtained from the questionnaire and interview data were organized according to the order in which the two research questions were structured: (1) the types of written feedback graduate students emphasized and (2) the comparison results of feedback preferences between TEFL and Media and Communication students.

The Students' Perceived Needs and Preferences for Various Types of Written Feedback

Table 1

Descriptive Statistics for Graduate Students' Preferences for the Specific Features of Content Feedback

No.	Content Feedback	Group	Mean	SD	MD	F	t	p-value
1.	Gaps in theoretical understanding and coverage	TEFL	4.45	.671				
		Media	3.50	.527	.955	.322	5.185	.000
2.	Gaps in the literature	TEFL	3.55	1.011				
		Media	4.30	.949	-.755	.445	-1.721	.055
3.	Irrelevance of thesis	TEFL	2.64	1.177				
		Media	3.10	1.101	.464	1.429	-1.214	.301
4.	Wider significance of the work	TEFL	2.18	1.435				
		Media	3.80	1.135	-1.618	1.778	-3.452	.004
5.	Gaps in the justification or explanation of arguments	TEFL	3.09	1.109				
		Media	3.40	1.174	-.309	.018	.794	.478

*SD: Standard Deviation *MD: Mean Difference

Table 1 displays the means, standard deviation, mean difference, and p-value results of graduate students' preferences for the specific features of content feedback on thesis writing. With a mean of 4.45 and standard deviation of .671, commenting on gaps in theoretical understanding and coverage was the most preferred type of supervisor feedback for the TEFL participants, whereas the most popular form of feedback for Media and Communications students (with a mean of 4.30 and standard deviation of .949) was Gaps in the literature. While both groups agreed that feedback on content was the most important feedback for their thesis writing, there is a glaring discrepancy between TEFL and Media and Communications students' views regarding some of the sub-categories for content feedback.

Based on the independent samples t-test results, we can conclude that TEFL participants showed significant differences on their perceptions regarding two items out of the five sub-categories of feedback on content. "Gaps in theoretical understanding and coverage" had a mean difference of .955 was significant at 0.000 for TEFL participants, whereas "Wider significance the work" had a mean difference of -1.618 with a significance level of $p < 0.05$. for Media and Communication participants. However, no statistically significant difference was found among the participants by discipline regarding their perceived needs toward the remaining three sub-categories.

The interview results of TEFL students' priorities were close to those expressed by Media and Communication students in comments related to content or subject areas. The thematic analysis of the students' interviews indicated that all of the students prioritized feedback on content. Both groups particularly acknowledged the indispensable necessity of comments on gaps in the literature and theoretical understanding. The interviews were particularly enlightening in regard to the types of gaps that both groups of students shared as the most important feedback. A TEFL student stated:

> As a graduate student, I must have a clear understanding of the topic and the related literature and the conceptual and theoretical framework so that my research would be scientific. Thus, I expect my advisor to read my thesis thoroughly and comment on how the writing pieces are linked with the review of the related literature as I am not familiar with how to accurately review an article or books.

The verbal account of the following interviewee result from Media and Communications signifies his strong preference for feedback on content.

> I would prefer my supervisor to comment on the accuracy, completeness, and relevance of content written in my thesis. As to me, commenting on content is the most important feedback. I am unaware of the difference between conceptual and theoretical frameworks, and it is usually difficult for me how to review the literature and how to relate my thesis to the review of related literature.

Table 2

Descriptive Statistics for Graduate Students' Preferences for the Specific Features of Part-Genre

No.	Genre Feedback	Group	Mean	SD	MD	F	t	p-value
1.	Content specific to the various part-genres of the thesis	TEFL	3.64	1.217				
		Media	3.50	1.650	.136	2.349	.847	.795
2.	Structure of what goes where and examples of what is expected	TEFL	3.23	1.110				
		Media	3.30	1.160	-.073	.027	.326	.867
3.	Purpose of part-genre	TEFL	2.45	1.299				
		Media	4.20	.789	-1.745	4.293	-3.248	.000
4.	Expected word count from a thesis	TEFL	2.64	1.560				
		Media	2.70	1.567	-.267	.882	-.030	.916
5.	Gaps in the justification or explanation of arguments	TEFL	3.09	1.109				
		Media	3.40	1.174	-.309	.018	.794	.478

*SD: Standard Deviation *MD: Mean Difference

Table 2 displays the means, standard deviation, mean difference, and p-value results of graduate students' preferences for the specific features of part-genre feedback on thesis writing. With respect to the specific components of feedback on part-genres, TEFL respondents agreed with a mean of (3.64) and standard deviation of 1.217 that content specific to the various part-genres of the thesis was the most useful component of the part-genres on thesis writing. With a mean of (4.20) and standard deviation of .789, the most preferred part-genre-related feedback for Media and Communications students was supervisor comments regarding the purpose of part-genre. Based on the data, both groups of participants were overwhelmingly positive about the role of all of the sub-categories of feedback on part-genre except for one item. The students' responses from the two disciplines on part-genre revealed some significant discrepancies in their concerns regarding the importance of various structures and components of the thesis. Students in Media and Communication differed significantly in their perception regarding the importance of the part-genre sub-category "Purpose of part-genre" with a mean difference of -1.745 and a significance level of $p < 0.05$.

The interview responses also revealed consistent results with the questionnaire data regarding the students' shared values for the sub-category of feedback on genre-content specific to the various part-genres of the thesis. Both TEFL and Media and Communications students stressed the value of comments on content specific to the various part-genres of the thesis. The following two samples were taken from the students who participated in the interview. A student from TEFL noted:

> Generally, as I am not an experienced researcher, I think that I have inadequate research skills necessary for thesis writing. That is why I chose feedback on content specific to the various part-genres of the thesis. I am not sure if I am writing research to an acceptable standard. I hope my advisor's feedback on this may help me decipher the quality of academic writing in research.

Likewise, another student from Media and Communications expressed:

> I have trouble in academic writing. Particularly, I have a lot of problems structuring and organizing my thesis. I don't have any idea of what to write in various components of a thesis. Therefore, it is important for me to receive such comments on my thesis.

Table 3

Descriptive Statistics for Graduate Students' Preferences for the Specific Features of Linguistic Feedback

No.	Linguistic Feedback	Group	Mean	SD	MD	F	t	p-value
1.	Punctuation, spelling, and capitalization	TEFL	2.00	1.113				
		Media	2.90	1.449	-.900	.028	-2.240	.063
2.	Systematic errors	TEFL	2.18	1.402				
		Media	3.60	1.174	-1.418	1.773	-3.106	.009
3.	Writing style	TEFL	2.50	1.225				
		Media	3.80	1.229	-1.300	.734	-3.149	.009
4.	How to develop ideas with supportive and counter evidence	TEFL	1.95	.999				
		Media	3.70	1.160	-1.745	.144	-4.526	.000
5.	Writing coherence and cohesion and quality of writing	TEFL	3.23	1.378				
		Media	3.90	.876	-.673	4.683	-1.388	.168
6.	Grammar (tenses, subject/verb agreement, article use, etc.)	TEFL	2.09	1.151				
		Media	2.80	1.135	-.709	.468	-2.104	.115
7.	Appropriateness of vocabulary choice and register	TEFL	3.05	1.174				
		Media	4.60	.516	-1.555	4.422	-3.704	.000

*SD: Standard Deviation *MD: Mean Difference

Table 3 displays the means, standard deviation, mean difference, and the p-value results of graduate students' preferences for the specific features of linguistic feedback on thesis writing. The data showed that graduate students from TEFL gave little value to all linguistic aspects of feedback except for Writing coherence and cohesion and quality of writing with a mean of 3.23 and standard deviation of 1.378. Media and Communications students favored Appropriateness of vocabulary choice and register with a mean score 4.60 and standard deviation of .516. The students from TEFL also expressed feedback on Appropriateness of vocabulary choice and register with a mean of 3.05 and standard deviation of 1.174 as their second priority. Writing coherence and cohesion and quality of writing received a mean of 3.90 and standard deviation of .876 from Media and Communications students. Seven specific feedback types were listed under feedback on the linguistic accuracy section. The two groups didn't show any statistical difference on three of the questionnaire items.

On the other hand, there were statistically significant differences between the adjusted mean scores of the remaining four items with regard to the sub-categories of linguistic accuracy and appropriateness at significant levels. This implies that students' responses differed according to their discipline in these four examples of linguistic feedback: 'Systematic errors', 'writing style', 'how to develop ideas with supportive and counter evidence', and 'appropriateness of vocabulary choice and register' with a mean difference of -1.418, -1.300, -1.745, and -1.555 respectively at $p<0.05$. Surprisingly all four of these sub-categories of linguistic feedback was attributed to Media and Communication students.

The interview results of the TEFL graduate students replicated the students' lower preferences for linguistic accuracy. The low level of priority given for this type of feedback might have stemmed from the students' prior experiences as all of the interviewees were experienced EFL secondary school English teachers for many years as one of the student noted:

> I don't see the relevance of receiving feedback on linguistic accuracy and appropriateness at this level because I don't think that comments on structural use of the language help me achieve the big picture of thesis writing. I also don't think that it adds something to my thesis writing.

Another TEFL student stated:

> As an English teacher, I have studied English in a university. I have also been teaching English for several years, and hence I believe that I gained adequate knowledge on the correct use of the English language. Therefore, I don't see the point of receiving grammar and language use feedback as this is not so difficult to manage for graduate students like me.

Conversely, feedback on linguistic accuracy was a big concern that mattered most for Media and Communications students. Apart from commenting on punctuation, spelling, capitalization, and grammar in their thesis writing, the participants from Media and Communications expressed that they needed the other sub-categories of linguistic accuracy, acknowledging that they had trouble producing a coherent and quality academic thesis as illustrated by one respondent:

> I want my advisor to comment on how to develop ideas, present arguments, and produce effective scientific writing. As the purpose of writing research is to communicate with the scientific community through my thesis, I think that my thesis should be well organized and structured, which is difficult to achieve without the proper linguistic feedback.

Results of Feedback Preferences between TEFL and Media and Communication Students

Table 4

Independent Sample t-test Results for Graduate Students' Preferences for General Written Feedback

No.	Feedback Category	Group	No	Mean	SD	MD	F	t	p-value
1.	Content	TEFL	20	3.18	.614.				
		Media	12	3.62	.676	-.438	.016	-1.583	.080
2.	Part-Genre	TEFL	20	2.99	.811				
		Media	12	3.43	.541	-.446	.268	-.894	.113
3.	Linguistic	TEFL	20	2.43	.784				
		Media	12	3.61	.547	-1.186	2.413	5.189	.000
	Total	TEFL	20	2.87	1.017				
		Media	12	3.55	.967	.68	.098	-3.888	.002

Notes: N = 32 * p<.05 (significant at less than .05 level)

Table 4 displays a summary of the independent samples t-test results for graduate students' preferences for the general written feedback on thesis writing. Regarding general feedback preferences, the overall responses obtained from the two groups showed that both groups were similarly concerned with feedback on content the most. This value underscores the point that students really appreciated the importance of content or subject-area feedback. Specifically, TEFL students placed great importance on comments regarding the content and subject area with a mean average value of 3.18 and standard deviation .614, whereas Media and Communications students showed their preferences for comments related to content in their thesis writing with a mean average value of 3.62 and standard deviation .676. Recalling from Table 3, it was evident that the quantitative data revealed statistically significant differences among the two groups of students' perceptions toward various sub-categories of linguistic feedback, which resulted in an overall statistical difference between them regarding the linguistic feedback and appropriateness.

Overall, Media and Communications students gave high ratings for feedback on content and linguistic features of supervisor comments, which indicates that the students would like their advisors to provide these comments in detail. From the researcher's shared experience, the researcher had hypothesized that students from both groups would prefer feedback on linguistic accuracy most. In contrast, with the mean difference of 1.186, participants from TEFL tended to have little or no interest in feedback on the linguistic aspects of their thesis work, indicating that TEFL students did not want or expect their advisors to comment on this.

The independent samples t-test was used to compare the two groups' scores on the same variable feedback and its sub-categories. The mean scores of TEFL and Media and Communication students were compared to evaluate whether there was a difference in their perception regarding written feedback. Inferring from the data, although both groups were similarly concerned with feedback on content the most, it was also evident that with a mean difference of .68, a statistically significant difference was observed between the two groups with regard to their overall perceived needs of feedback at $p<0.005$. This result, could be ascribed to the linguistic accuracy and appropriateness in favor of Media and Communication students.

Table 5

Descriptive Statistics of Each Aspect of the Students' Perceived Needs of Feedback Instrument

	Mean	SD	Skewedness	Kurtosis	Min	Max	N
Feedback on Content	3.32	.656	-.585	1.298	1	5	32
Feedback on Part-genre	3.13	.757	-730	.749	1	5	32
Linguistic Feedback	2.80	.903	-.050	-.081	1	5	32

*SD: Standard Deviation *Min: Minimum *Max: Maximum *N: Total number

As indicated in the table above, in all four of the feedback needs, the minimum and maximum values were one and five, respectively. The variables' normal distribution of the data was supported by the results. Since both the degrees of skewedness and kurtosis were less than the absolute value of one, the variables were normally distributed.

Table 6

The Questionnaire Items and Final Three-Factor Structure of Students' Perceived Needs

		Factor		
		1	2	3
	Factor 1: Feedback on Content needs			
1.	Gaps in theoretical understanding and coverage	.820		
2.	Gaps in the literature	.785		
3.	Irrelevance of the thesis	.729		
4.	Wider significance of the work	.652		
5.	Gaps in the justification or explanation of arguments	.740		
	Factor 2: Feedback on Part-genre needs			
6.	Content specific to the various part-genres of the thesis		.754	
7.	Structure of what goes where and examples of what is expected		.680	

		Factor		
		1	2	3
8.	Purpose of part-genre		.842	
9.	Expected word count for a thesis		.714	
	Factor 3: Linguistic Feedback needs			
10.	Punctuation, spelling, and capitalization			772
11.	Systematic errors			.626
12.	Writing style			.761
13.	How to develop ideas with supportive and counter evidence			.758
14.	Writing coherence and cohesion and quality of writing			.663
15.	Grammar (Tenses, subject/verb agreement, article use, etc.)			.687
16.	Appropriateness of vocabulary choice and register			.793

Table 7

Cronbach's Alpha for Each Aspect of the Students' Questionnaire

	Cronbach's Alpha	Number of items
Feedback on Content needs	.782	5
Feedback on Part-genre needs	.790	4
Linguistic Feedback needs	.769	7
Total	.776	16

Discussion

A growing body of research is investigating students' preferences for supervisor written feedback for thesis writing as examining student needs and preferences may help researchers and pedagogic experts fully understand the benefits of various types of supervisor written feedback to improve student thesis writing. This study sought to examine how EFL graduate students perceive the three major written feedback types provided by supervisors for thesis writing. Although a small-sized survey, overall, the results of this study demonstrate that the students held a very positive view towards the three major feedback types on thesis writing and preferred feedback on content the most. This conceptualization, as well as the more positive tendency towards written feedback reported in this study, can be considered vital, and signals the importance of considering the needs of students when providing supervision on thesis writing, since it might be a good idea for teachers to offer more feedback on the content and structure of their students' work for advanced EFL learners (Chen, Nassaji, & Liu, 2016).

The results also revealed that both TEFL and Media and Communication students really appreciated the importance of content or subject-area feedback. It was also indicated that both groups of students still expressed a favorable preference for part-genre feedback on their thesis writing. This is a reminiscent of Manjet's (2016) conviction that graduate students preferred feedback to be on the content of their academic work and not merely focus on language and formatting issues. Given that student researchers have little or no capacity and understanding of basic research components and the skills necessary to undertake a study (Yenus, 2018), the results of this study that showed students' highest preference for feedback on content that includes key research skills such as gaps in the literature, gaps in theoretical understanding and coverage, irrelevance of the thesis, wider significance the work, gaps in the justification or explanation of arguments is not surprising. Similar EFL student perceptions towards content feedback were reported in the study by Bitchener et al. (2011), in which the students who participated in the study thought content feedback was the most useful for their thesis writing. In other similar studies, it was reported that although teachers' practices were driven by the students' preferences, these teachers mostly provided teacher-generated grammar feedback simply because they believed that their students preferred this type of correction (Hooman Saeli, 2019).

The present study also aimed to understand whether the students' perceptions significantly differed across the two disciplines: TEFL and Media and Communications. The results of this study revealed that there were more similarities than variations among the two groups of students' feedback preferences. The second sub-theme was along the lines of the specific components of each piece of written feedback. It was found that gaps in theoretical understanding and coverage and gaps in the literature was the most popular form of feedback for both groups. However, when the students' needs and preferences for the sub-categories of content feedback were further examined, the students showed significant differences. The sub-category "gaps in theoretical understanding and coverage" could be ascribed to TEFL participants, and "wider significance the work" could be ascribed to Media and Communication participants. While the graduate students' perceived needs might have been affected by their prior experiences in the form of academic writing experience, journal feedback, and exposure to a similar genre (Hoomanfard, et. al., 2018), taking this variation into account can be seen as an important reminder for supervisors since evaluating feedback quality needs careful consideration on the extent to which staff and students agree on the purpose of the feedback (Price et al., 2010).

Another interesting result was the students' wide range of preferences for feedback on part-genre, although both groups of respondents alluded to their difficulties writing and organizing their theses. These discrepancies might be because the academic writing practices the two groups of students brought from their prior academic learning backgrounds might have affected their expectations in their current thesis writing (Manjet, 2015), indicating the necessity for considering the particular culture of academic writing trends in each discipline when providing feedback on student thesis writing. The students' responses about their gaps in content specific to the various part-genres of their theses were particularly enlightening as supervisors are expected to guide the student researchers throughout their study and provide the time, expertise, and support to foster the candidate's research skills and attitude as well as to ensure the production of research of an acceptable standard (Heath, 2002).

A substantial amount of research on students' preferences for supervisor written feedback for thesis writing has been documented. The students' lowest preferences for linguistic accuracy reported in this study were in contrast to previous research which asserted that students valued and appreciated written feedback about grammar (Bitchner et al., 2011; Manjet, 2016). Particularly, the students' real concern on the two items pertinent to organization in academic writing: appropriateness of vocabulary choice and register, and writing coherence and cohesion and quality of writing could provide corroborative evidence for the results of previous research in that graduate students face challenges in their academic writing practices in the context of expressing ideas, linking ideas, sequencing their assignment, and ensuring clarity in their writing (Manjet, 2016). Although why these students reported the lowest priority for receiving comments pertinent to feedback on linguistic accuracy requires further study, it can be assumed that their prior knowledge might have affected their preferences as students' linguistic ability and linguistic self-confidence could affect the their perceptions regarding the need for feedback on their thesis/dissertation (Hoomanfard, et. al., 2018).

The convergent preference between TEFL and Media and Communications students in this study where students in both disciplines acknowledged that they needed more comments on content areas was very strong. However, the researcher took a very cautious approach to this finding as this realization may be due to the influence of the constant use of the same correction methods that lead students to think they are the best methods (Lee, 2004). On the other hand, this result was consistent with the results of the study of Yenus (2018) that compared the focus of feedback between TEFL and Media and Communications supervisors by analyzing the samples of supervisors' written feedback in the final version of student theses. Concurrently, the views of the respondents of the present study were, therefore, in tandem with his assertion that supervisors in both disciplines focused feedback on content the most, implying the congruence between graduate students' feedback preferences and supervisors' practices in responding to student thesis writing. Therefore, the results of this study that showed that various features of content or subject-area feedback as graduate students' most preferred feedback is also in line with previous research (Bitchner et al., 2011; Chokwe, 2015; Lee, 2008; Manjet, 2016).

Conclusion

The quantitative analysis of the questionnaire results and the interview responses of the participants revealed the following conclusions. They showed that feedback on content was the graduate students' most preferred feedback, indicating that this type of feedback was the most important type of feedback that the participants were concerned with. The results of this study also showed that both groups of students seemed to equate the importance of the provision of feedback on content. Concerning the second research question, congruence and discrepancies between students from TEFL and Media and Communication students emerged. While the students from both groups valued feedback on content the most, the students' preferences were also incongruent with supervisor feedback, which could be ascribed to linguistic accuracy and appropriateness in favor of Media and Communication students. Further, the participants also showed discrepancies regarding their preferences for part-genres on thesis writing and on the various features of linguistic accuracy. The study also showed that students across the two disciplines differed in their feedback preferences regarding the specific components of each category in thesis writing.

Based on the major findings, it is suggested that efforts be made for the better realization of students' need and preferences in thesis writing. Advisors also need to approach their supervision more thoughtfully by linking theory and practice with regard to feedback for thesis writing. All things considered, instructors should critically analyze their students' writing preferences and adjust how to balance these preferences for written feedback thereby helping their students with successful thesis writing.

It is hoped that this research provides a significant contribution to scientific knowledge pertaining to supervision in the higher education context. The participants in this study that came from two disciplines, namely TEFL and Media and Communications were found to be divided in their opinions towards some aspects of feedback for thesis writing. Therefore, the reasons behind their preferences were not examined as this was not the purpose of this study, indicating the need for future studies that examine this issue including other data gathering tools, such as focus-group discussions and text analysis of student theses. Another limitation was that the sample size in the TEFL group was more largely represented than the students who participated from Media and Communications; hence this may have implications for future research examining students' preferences with a fairer representation of students from the two disciplines.

References

Aitchison, C., Catterall, J, Ross, P., & Burgin, S. (2012) 'Tough love and tears': Learning doctoral writing in the sciences, *Higher Education Research & Development, 31*(4), 435-447. http://dx.doi.org/10.1080/07294360.2011.559195.

Ali, P. A., Watson, R., & Dhingra, K. (2016). Postgraduate research students' and their supervisors' attitudes towards supervision. *International Journal of Doctoral Studies, 11*, 227-241. https://doi.org/10.28945/3541

Amrhein, H., & Nassaji, H. (2010). Written corrective feedback: What do students and teachers prefer and why? *Canadian Journal of Applied Linguistics*,*13*, 95-127.

Badenhorst, C., Moloney, C., Rosales, J., and Ru, L.(2015). Beyond deficit: Graduate student research-writing pedagogies, *Teaching in Higher Education, 20*(1), 1-11. https://doi.org/10.1080/13562517.2014.945160.

Basturkmen, H., East, M., & Bitchener, J. (2014). Supervisors' on-script feedback comments on drafts of dissertations: socializing students into the academic discourse community, *Teaching in Higher Education, 19*(4), 432-445. https://doi.org/10.1080/13562517.2012.752728

Bitchener, J., Basturkmen, H., & East, M. (2010). The focus of supervisor written feedback to thesis/dissertation students. *International Journal of English Studies, 10*(2), 79-97. https://doi.org/10.6018/ijes/2010/2/11920.

Bitchener, J., Basturkmen, H., East, M., & Meyer, H. (2011). *Best practices in supervisor feedback to thesis students.* Research report. Ako Aotearoa .

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy.* Pearson Education.

Can, G., & Walker, A. (2014). Social science doctoral students' needs and preferences for written feedback. *Higher Education, 68*(2), 303-318. https://doi.org/10.1007/s10734-014-9713-5.

Can, G., & Walker, A. (2011) A Model for doctoral students' perceptions and attitudes toward written feedback for academic writing. *Research in Higher Education, 52*, 508–536. https://doi 10.1007/s11162-010-9204-1

Caterall, J., Ross, P., Aitchison, C., & Burgin, S. (2011). Pedagogical approaches that facilitate writing in postgraduate research candidature in science and technology. *Journal of University Teaching and Learning Practice, 8*(2), Article 7.

Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: a case study of university students from Mainland China. *Asian Pacific Journal of Second and Foreign Language*

Education, 1(5), 1-17. https://doi10.1186/s40862016-0010-y.
Chokwe, J. M. (2015). Students' and tutors' perceptions of feedback on academic essays in an open and distance learning context. *Open Praxis, 7*(1), 39-56. https://doi.org/10.5944/openpraxis.7.1.154
Cotterall, S. (2011). Doctoral students writing: where's the pedagogy? *Teaching in Higher Education, 16*(4), 413-425. https://doi.org/10.1080/13562517.2011.560381.
Esra, O. (2020)."I would like her to…:" ELT student-teachers' reflections on mentoring practices. *PASAA: A Journal of Language teaching and learning, 59*(1), 48-76.
Evans, N., Hartshorn, J., McCollum, R., & Wolfersberger, M. (2010). Contextualizing corrective feedback in second language writing pedagogy. *Language Teaching Research, 14* (4), 445–463. https://doi.org/10.1177/1362168810375367
Ferris, D. (2003). *Response to student writing: Implications for second language students*. Lawrence Erlbaum.
Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, *10(3)*, 161-184. https://doi.org/10.1016/S10603743(01)00039-X
Ghandi, M., & Maghsoudi, M. (2014). The effect of direct and indirect corrective feedback on Iranian EFL learners' spelling errors. *English Language Teaching, 7*, 53-61. https://doi.org/10.5539/elt.v7n8p53.
Ghazal, L., Gul, R, Hanzala, M., Jessop, T., & Tharani, A. (2014) Graduate students' perceptions of written feedback at a private university in Pakistan. *International Journal of Higher Education, 3*(2), 13-27. https://doi.org/10.5430/ijhe.v3n2p13
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487
Heath, T. (2002). A quantitative analysis of PhD students' views of supervision. *Higher Education Research Development, 21*, 41-61. https://doi.org/10.1080/07294360220124648
Hoomanfard, M. H., Jafarigohar, M., Jalilifar, A., & Hosseini Masum, S. M. (2018). A comparative study of graduate students' self-perceived needs for written feedback and the supervisors' perceptions. *Journal of Research in Applied Linguistics, 9*(2), 3-26. https://doi.org/10.1016/j.esp.2005.12.003
Hyland, K. (2009). *Academic discourse*. Continuum.
Katayama, A. (2007). Japanese EFL students' preferences toward correction of classroom oral errors. *Asian EFL Journal, 9*(4), 289-305. http://dx.doi.org/10.17507/tpls.0703.03
Kumar, V., & Stracke, E. (2007). An analysis of written feedback on a PhD thesis. *Teaching in Higher Education, 12*(4), 461–470. https://doi.org/10.1080/13562510701415433
Lavelle, E., & Bushrow, K. (2007). Writing approaches of graduate students. *Educational Psychology, 27*(6), 807–822.
Lee, A. & Murray, R. (2015) Supervising writing: helping postgraduate students develop as Researchers. *Innovations in Education and Teaching International, 52*(5), 558-570. https://doi.org/10.1080/14703297.2013.866329..
Lee, I. (2004). Error correction in L2 secondary writing classroom. *Journal of Second Language Writing, 13(4)*, 285-312. https://doi.org/10.1016/j. jslw:2004.08.001
Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing, 17(3),* 144-164. https://doi.org/10.1016/j.jslw.2007.12.001
Leng, K. (2014). An analysis of written feedback on ESL students' writing. *Procedia – Social and Behavioral Sciences, 123*, 389-397. https://doi.org/10.1016/j.sbspro.2014.01.1437.
Manjet, K. (2016). Graduate students' needs and preferences for written feedback on academic writing. *English Language Teaching, 9*(12), 79-88. https://doi.org/10.5539/elt.v9n12p79
Paré, A. (2011). *Speaking and writing: Supervisory feedback and the dissertation*. In L. McAlpine & C. Amundsen (Eds.), Doctoral education: Research-based strategies (pp. 59-74). Springer. https://doi.org/10.1007/978-94-007-0507-4_4
Pokorny, H. & Pickford, P. (2010). Complexity cues and relationships: Student perceptions of feedback. *Active& Learning in higher Education, 11(1),* 21-30. https://doi.org/10.1177/1469787409355872
Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: all that effort, but what is the effect? *Assessment and Evaluation in Higher Education, 35*(3), 277-289. http://dx.doi.org/10.1080/02602930903541007.
Rowe, A. (2009). The personal dimension in teaching: Why students value feedback. *International Journal of Educational Managemen,t 25(4)*, 343–360. https://doi.org/10.1108/09513541111136630 .
Saeli, H. (2019). Teachers' practices and students' preferences: Grammar-centered written corrective feedback in Iran. *Research in English Language Pedagogy, 7*(1), 46-70. https://doi.org/10.30486/relp.2019.663422
Stracke, E., & Kumar, V. (2010). Feedback and self-regulated learning: insights from supervisors' and PhD examiners' reports. *Reflective Practice, 11*(1), 19-32. http://dx.doi.org/10.1080/14623940903525140

Sutton, P. (2009).Towards dialogic feedback. *Critical and Reflective Practice in Education, 1,* 1-10. https://doi.org/10.1007/978-3-642-13166-0_27

Tee, P. L., Kumar, V., & Abdullah, M. H. (2013). Written feedback on ESL student writers' academic essays. *Pertanika Journal of Social Sciences and Humanities, 21*(2), 477-502.

Vygotsky, L. (1978). The interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.), *Mind and Society* (pp.79-91). Harvard University Press.

Wang, T., & Jiang, L. (2015). Studies on written corrective feedback: Theoretical perspectives, empirical evidence, and future directions. *English Language Teaching, 8*(1), 110-120. http://dx.doi.org/10.5539/elt.v8n1p110.

Wang, T., & Li, L. (2011). 'Tell me what to do' vs 'guide me through it': Feedback experiences of international doctoral students. *Active Learning in Higher Education, 12*(2), 101-112. https://doi.org/10.1177/1469787411402438

Wang, P. (2010). Dealing with English majors' written errors in Chinese universities. *Journal of Language Teaching and Research, 1*(3), 194-205. http://dx.doi.org/10.4304/jltr.1.3.

Weaver, M. (2006). Do students value feedback? Students' perception of tutors' written responses. *Assessment and Evaluation in Higher Education, 31*(3), 379-394. https://doi.org/10.1080/02602930500353061

Yenus, N. (2018a). Higher education supervision practices on student thesis writing: Language function and focus of written feedback. *International Journal of Teaching and Learning in Higher Education, 30*(3), 1-12.

Yenus, N. (2018b). Doctoral students' perceived needs and preferences for supervisors' written feedback. *PASAA, 56*(2),112-144.

Yin, K. (2011). *Qualitative research from start to finish*. Guilford Press.

Yinager, T. (2019). Views of summer TEFL postgraduate students and their supervisors on the practice of thesis supervision in the Ethiopian context. *PASAA, 58* (2), 131-165.

Yu, S., & Lee, I. (2013). Understanding supervisors' commentary practices in doctoral research proposal writing: A Hong Kong study. *Asia-Pacific Education Researcher, 22*(4), 473-483. https://doi.org/10.1007/s40299-012-0046-9.

Zeleke A. (2017). Ethiopian public university entrants' writing skills in English language: The case of Hawassa University entrants. *International Journal of Development Research, 7*(9), 15089-15092.

Zewdu, E. (2012). An *investigation into the research culture of Addis Ababa University: The case of a Teaching English as a Foreign Language (TEFL) PhD program* [Doctoral

Dissertation]. Addis Ababa University. URL www.bdu.edu.et>bdu.edu.et.fh>files

Zhan, L. (2016). Written teacher feedback: student perceptions, teacher perceptions, and actual teacher performance. *English Language Teaching, 9*(8), 73-84. https://doi.org/10.5539/elt.v9n8p73

Appendix A

Questionnaire for Graduate Students

The purpose of this questionnaire is to collect data on areas which you consider most important for you, and you want your advisor to look at. Therefore, your genuine response is useful for the completion of this study. The following are some postgraduate students' preferences towards supervisor feedback to thesis writing. Read each statement and then decide if you: (1) strongly disagree, (2) disagree,(3) neither agree nor disagree (undecided), (4) agree, or (5) strongly agree corresponding to the items provided.

Thank You

When responding to my thesis, I want my advisor to comment on the:	1	2	3	4	5
Content					
Gaps in theoretical understanding and coverage					
Gaps in the literature					
Irrelevance of thesis					
Wider significance the work					
Gaps in the justification or explanation of arguments					
Part-genre					
Structure of what goes where and examples of what is expected					
Purpose of part-genre					
Part-genre Skills required for each genre					
Expected word count from a thesis					
Linguistic accuracy and appropriateness					
Punctuation , spelling and capitalization					
Systematic errors					
Writing style					
How to develop ideas with supportive and counter evidence					
Writing coherence and cohesion and quality of writing					
Grammar (verb tenses, subject/verb agreement, article use…etc.)					
Appropriateness of vocabulary choice and academic register					

Appendix B

Interview Questions for Graduate Students Answer the following questions by illustrating your points. I am only interested in understanding graduate students feedback needs to thesis writing.

1. Do you prefer that your supervisor provide you all types of content feedback? Why?
2. Which specific aspects of content feedback are the most important to you when your supervisor provides feedback to your thesis writing? Why?
3. Do you prefer that your supervisor provide you all types of genre feedback? Why?
4. Which specific aspects of genre feedback are the most important to you when your supervisor provides feedback to your thesis writing? Why?
5. Do you prefer that your supervisor provide you all types of linguistic feedback? Why?
6. Which specific aspects of linguistic accuracy feedback are the most important to you when your supervisor provides feedback to your thesis writing? Why?
7. In general, which feedback type among (feedback on content, genre or linguistic feedback) to thesis writing, do you like, your supervisor to provide you most? Why?
8. What are your expectations with respect to your supervisor feedback to your thesis writing? Why?

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Thumvichit, A. (2020). ELT Master's Programmes in Thailand: Focused Areas and Research Trends. *Journal of Language and Education, 6*(4), 171-182. https://doi.org/10.17323/jle.2020.10532

ELT Master's Programmes in Thailand: Focused Areas and Research Trends

Athip Thumvichit
Ramkhamhaeng University

Correspondence concerning this article should be addressed to Athip Thumvichit, Ramkhamhaeng University, 2086 Ramkhamhaeng Rd, Hua Mak, Bang Kapi District, Bangkok 10240.
E-mail: athip.t@rumail.ru.ac.th

The increasing demand for competent users of English and qualified English teachers has accelerated the growth of graduate programmes in English language teaching (ELT). In Thailand, ELT master's programmes have been serving as a training ground for Thai English teachers for decades. This study explores the focused areas and research trends of Thai ELT master's programmes. The analyses involved ten ELT master's programmes offered by ten different universities and 201 master's theses submitted between the years of 2014 and 2018. Foundation and core courses were categorised into twelve content areas. The findings show that teaching and research methodology courses were the most common areas, indicating that the programmes were not only pedagogical but also research-oriented. In consistence with the international trends of ELT research, the *Instructional effects* research area still prevails. The number of studies on *Assessment* and *Curriculum/Programme* is relatively low compared with the number of courses in such areas. Also discussed are considerations for programme management, lecturers, and students. It is also recommended that all the courses offered be treated as a gateway to research opportunities in addition to teaching practice and professional practice improvement.

***Keywords*:** English language teaching, English teachers, master's programmes, master's theses, research trends

Introduction

English Language Teaching (ELT) master's programmes in Thailand were established to strengthen qualifications and career pathways of those entering or already employed in English language education. The programmes are considered a concentrated route for professional English teachers, especially ones aiming for advanced skills and knowledge in the field of pedagogy. The programmes are also considered a stepping stone to professional development and a gateway to a teaching career at the tertiary level. In this sense, ELT master's programmes are now serving as a training ground for Thai English teachers and even for those who are pursuing other English-related careers. As English becomes more important and pervasive, ELT master's programmes in Thailand are growing more diverse. Despite the increasing number, diversity, and benefits of ELT master's programmes, the programmes as a whole have not yet been comprehensively explored. With the absence of up-to-the-minute literature devoted to the progress of ELT at the graduate level, there is a paucity of understanding surrounding areas of focus and research trends in ELT master's programmes around the country. By gaining insight into the currently research trends and areas of focus, ELT scholars and programme management can be better informed about the present state of affairs in the ELT community (Lin & Chen, 2010). Addressing such issues also provides a better understanding of how to move the ELT community forward (Abdel Latif, 2018).

Considering the importance of ELT masters' programmes, as the first part of a larger effort[1] to explore various aspects of these programmes, this study is devoted to investigating areas of focus and research trends. Exploring these two areas will highlight the knowledge fields and research areas that have been emphasised and underemphasised in ELT master's programmes, leading to calls for changes to improve the programmes.

[1] The other part of this project explores reasons for attending ELT master's programmes and crucial factors influencing the choice of university and programme (Thumvichit, 2020).

From English Language Education Reforms to the Establishment of ELT Programmes

The Thai government promulgated the National Education Act (NEA) of 1999 to improve overall educational quality. One of the main areas of focus was tertiary education with an emphasis on internationalising it and making English as a medium of instruction standard practice. Such an attempt accelerated the growth of international programmes and the emergence of various English-related courses (Kaur et al., 2016). Proposed by the Ministry of Education (MoE) to replace the 2001 Basic Education Core Curriculum (BEC), the 2008 BEC was explicitly crafted to address the demands of globalisation, making English an essential language. It suggested that language learning should concentrate not only on communicative goals, but also culture, relationships with other disciplines, and relationships among communities. To align with the proposed ideas and the growing importance of English in the global arena, the stimulation of English language education came in the form of accredited international schools, English programmes (EPs) and bilingual programmes (BPs) at schools, and international programmes at higher education institutions.

Additionally, the MoE realised the importance of graduate programmes as a route to professional development and career advancement, promoting the establishment of university language institutes to fulfil the rising need for English language training for a wide variety of purposes ranging from general to professional communication to English teacher development. The 21st century has witnessed the considerable growth of the ELT graduate community in Thailand. ELT graduate programmes have now become diverse in focus and variously named. Some of them are called English Language Teaching (ELT), and others operate under the widely used name Teaching English to Speakers of Other Languages (TESOL). One programme that sets itself apart from others is named Applied Linguistics for English Language Teaching (ALELT). In response to the current status of English, there are also programmes named Teaching English as an International Language (TEIL) and Teaching English as a Global Language (TEGL), addressing the intercultural aspects of ELT. Unlike other major fields, the master's degree in ELT is offered as either a Master of Arts (MA) or Master of Education (MEd). Although various types of ELT graduate degrees are offered (e.g., certificates, master's degrees, doctoral degrees), the master's degree has tended to receive greater attention than others because, in general, it is the minimum requirement for a tertiary teaching position and the programme's duration is only two years on average.

Research into ELT Master's Programmes

Many studies have been conducted to develop a better understanding of ELT programmes. Hasrati and Tavakoli's (2015) work draws on interview and questionnaire data to explore how globalised contexts of higher education affected MA TESOL programmes in the UK. They found that globalisation has brought more international students to the programmes and brought about the need to modify the programmes' curricula and content. US survey studies reported that approximately 40% of the students enrolling in MA TESOL programmes in the US (England & Roberts, 1989) and as many as 90% in some US universities between the late 80s and 2001 (Brady & Gulikers, 2004) were L2 users of English. Since English-speaking countries are perceived to be a reliable destination for international students, topics related to international students' experiences in English-related graduate programmes have received tremendous attention from language scholars as they contribute immensely to the programmes' development. There have been numerous studies focusing on international students' experiences of ELT-related master's programmes in top higher education destination countries: the US (e.g., Baecher, 2012; Tseng, 2013), the UK (e.g., Copland & Garton 2012; Fordyce & Hennebry 2013; Schartner & Young, 2015), Australia (e.g., Hughes & Bruce, 2013; Phakiti & Li, 2011), and New Zealand (e.g., Li & Tin, 2013). While many prior studies focused on a single university or a particular group of universities, Copland et al. (2017) conducted a national-scale study focusing on ELT master's programmes in the UK. In that study, several aspects of ELT master's programmes were explored: the availability of the courses, modes and formats, target students, enrolment decisions, and students' expectations. Involving nearly all of the ELT master's programmes in the UK, the findings of this study can represent various aspects of ELT master's programmes in the UK.

Regarding content areas covered in ELT master's programmes, Stapleton and Shao's (2018) large-scale study reports on a survey of 241 MA TESOL programmes offered in 16 countries. In this study, courses offered in the programmes were examined for their coverage and prevalent knowledge fields. It revealed a wide range of knowledge fields not limited to the teaching and learning of English. That is, courses with an emphasis on society, culture, sociolinguistics, literature, and language arts were also covered in the programmes. It was also found that the most common knowledge field was *Teaching methods/issues*, followed by *Elements of linguistics* (e.g., lexis, phonetics, phonology, and semantics). Apart from the findings, that study provided insights into a certain community of ELT graduate programmes.

Despite the growth of the ELT graduate community, the rising demand for qualified English teachers and competent English users at professional levels, and the growing body of research on ELT master's programmes in the context of top education destination countries, there has been an absence of research committed to exploring ELT master's programmes in Thailand as a whole. The current study is, therefore, a timely response to this shortage and can trigger local researchers' interest in the Thai ELT community. Although this study does not directly reflect the current state of the global community of ELT graduate programmes that are growing rapidly along with the importance of English, it commits to moving this community forward by observing what the programmes have to offer to improve teaching practice, enhance research skills, and expand opportunities.

Previous Studies on Trends of Language Teaching Research

Over the past decade, a growing interest in ELT research among language scholars can be observed. Lin and Cheng (2010) investigated the prevalent research contexts and research areas of master's theses in Taiwan universities produced between 2003 and 2007. To identify the prevalent research areas, they adopted the list of content areas created by the TESOL International Association in 2009. Their findings suggested that the most preferred was *Language skills* followed by *Teaching methods*, and the majority of studies were conducted on the secondary education platform, which concurs with the programme objectives and aspects of the courses. Employing a similar list of research areas to categorise research papers, Kirmizi (2012) investigated the research trends of master's theses submitted in universities in Turkey between 2006 and 2011. Identically, *Language skills* and *Teaching methods* remained at the top of the list. Master's programmes, however, were not the only source of data for the identification of research trends. Tůma and Píšová (2013) drew on doctoral dissertations defended in the Czech Republic and abroad between 2006 and 2012. Instead of adopting the previously proposed categorisation framework, they employed thematic analysis to outline the research categories. The findings showed that *Management of teaching/learning processes* was the most prevalent category. However, this does not mean that the findings of Tůma and Píšová's (2013) study were entirely in line with those of aforementioned studies because *Management of teaching/learning processes* tended to incorporate a broader range of research themes than *Language skills* in the TESOL list, which concerns only different language skills. A more extensive study by Bani-Khaled's (2012) pointed to trends over four decades covering the period from 1974 to 2010. In that study, a corpus of 550 master's theses and doctoral dissertations was compiled. *Pedagogy/ Methodology* was ranked the most researched area of research with over 30% concerning teaching and teaching methods. It is interesting to learn that pedagogy and teaching methods were the most selected areas, even in non-teaching programmes.

Besides master's theses and doctoral dissertations, published articles also served as data in studies on ELT research trends. Wenfeng and Gao (2008), for example, investigated the research areas of research output on English language education in China. A corpus of 81 articles published between 2001 and 2006 in 24 international journals was compiled. The findings indicated that *English used in China* had more articles than others. This category covers a range of research areas including the development of English as a foreign language in China, varieties of English, and linguistic analyses of English used in China. Unlike master's theses and doctoral dissertations, internationally published articles tended to seek an understanding of how Chinese influenced the development and use of English in China, rather than teaching elements. Some studies were designed to investigate the research trends of a single journal. Stapleton and Shao (2017), for example, unveiled the research trends of the 359 articles published in the *Language Teaching Research* journal during the journal first's 20 years. *Instructional effects* was a more common theme than others. Although the journal invited a vast range of topics, teaching-related topics remained the most popular.

When the scope of research is restricted to the 'teaching' of a particular language, researchers can generate a list of research areas used specifically to categorise 'teaching' research. Recently, Abdel Latif (2018) proposed a list that emphasises the teaching elements of ELT research. All nine research areas presented in the list are related to the teaching and learning of English because the study attempted to identify the research trends of published research only in the field of ELT. To this end, unlike previous studies, the researcher did not add research areas that were not related to the teaching of English. In addition to identifying research areas, which was the main focus of the study, the researcher also identified the research methods and designs of the studies as well as the contexts in which the studies were conducted. As for the research methods and designs, each research article was labelled either ‘experimental’ or ‘non-experimental’, and either ‘quantitative’, ‘qualitative’, or ‘mixed’. The researcher proposed a set of research contexts covering a wide variety of educational stages ranging from the kindergarten level to the tertiary level to non-award programmes, leading to a firm understanding of the contexts in which the studies were conducted. Nevertheless, research trends undergo changes over time and thus the findings can represent only a snapshot of research trends. In this sense, the current study serves as a picture of the current state of ELT education at the graduate level in Thailand, investigating research trends to better understand research focuses and directions, and highlight the dearth of research on some knowledge fields.

Methodology

Identifying ELT Master’s Programmes in Thailand

First, I browsed through the list of higher education institutions provided by the Office of the Higher Education Commission (OHEC) (www.mua.go.th) to identify all higher education institutions and all curricula approved between the years of 2012 and 2018. Under the supervision of the OHEC, 156 higher education institutions were identified. I read through the list to obtain the names of English-related master’s programmes. Since this study focuses purely on English teaching programmes, those offering an English-related degree that was not entirely devoted to teaching (e.g., Applied Linguistics, English as an International Language, English for Specific Purposes) were excluded. In addition, the programmes that had ceased recruiting students were also excluded. Accordingly, I contacted the programme coordinators to ensure the latest status of the programmes before proceeding. I found that some programmes were temporary, if not permanently, closed due to the shortage of applicants and/or qualified lecturers. After several rounds of screening, ten programmes offered by ten different universities were identified. Six of them were offered under the MA degree, and the rest were offered under the MEd degree. The selection of the programmes, therefore, renders the results more representative of focused areas and research trends of ELT master’s programmes in Thailand. To maintain the ethical standards, the programme titles and the names of the universities are not disclosed. Throughout this paper, therefore, the codenames *ELTP1* to *ELTP10* are used to refer to the programmes.

Analysing Focused Areas

Courses Offered in ELT Master’s Programmes

To capture the focused areas, this study drew on taught courses offered in the programmes. In general, there were three types of courses: foundation courses, core courses, and elective courses. Foundation courses are courses that are required prior to or alongside core courses, all of which are compulsory for all students. Foundation courses generally aim to prepare students for graduate-level study and studies in the field of education. Foundation courses under the ELT master’s programme come in the form of English skills development or non-English-related educational principles. However, as indicated in curriculum documents, some foundation courses can be waived if students meet certain requirements, such as satisfactory results of a prescribed English test. Elective courses, on the other hand, were treated as choices. Therefore, all the elective courses were excluded from this analysis. In this regard, all foundation and core courses of each programme were collected for the analysis.

The programmes tended to offer two sets of courses: one for the thesis plan and one for the non-thesis plan. In some programmes, the two sets were slightly different as the non-thesis plan, also known as Plan B, required more course credits than the thesis plan, also known as Plan A. In *ELTP8,* for example, Plan B required six more core course credits than Plan A. Non-thesis students were offered some courses that were completely absent in Plan A. Thus, to align with the other focus of this study, namely research trends of master’s theses, only courses that were offered in Plan A were analysed.

Data Analysis

The foundation and core courses were categorised into different focused areas in ELT, employing thematic analysis, which helps identify a distinctive set of themes representing a particular dataset. To do so, course descriptions were read through to identify the content and scope of the courses, both of which were used for assisting categorisation of themes. The courses with the same or similar content areas were convened, after which distinct themes emerged representing different content areas. The findings are discussed for insight into the overall focused areas of ELT master's programmes in Thailand.

Analysing Research Trends

Identifying Theses in ELT Master's Programmes

This study draws on ELT master's theses made publicly accessible between 2014 and 2018 from ELT programmes. Master's theses and doctoral dissertations are normally not published in external academic journals but rather are available on library databases. Therefore, these databases were used to generate data for the analysis. After locating the library websites, I used built-in search engines to gain access to the full-text files of the theses. The built-in search engines were configured for searching master's theses submitted between 2014 and 2018. The rationale behind choosing these years was that they were the five most recent years and that they would reflect the most recent research trends. Since the libraries allow public access, the full-text files were downloaded and stored on a computer. In cases where the library's built-in search engine had no features that could help locate potential theses, librarians were contacted for support. However, some recently completed theses were not yet made publicly accessible, especially those approved in 2018. To this end, to avoid violating university copyright policies, only theses that were available on the library databases at the time of conducting this study were collected for the analysis. However, there were a few theses that were excluded because of the absence of full-text files. Visiting all the library databases yielded a corpus of 201 theses.

Data Analysis

This study employed Stapleton and Shao's (2017) list of research focuses as a starter, which comprised 13 areas: (1) *Instructional effects,* (2) *Teacher cognition,* (3) *Exploratory practice,* (4) *Teacher education,* (5) *Learner behaviors,* (6) *Learner cognition,* (7) *Theoretical essays, narratives, and descriptions*, (8) *Teacher practices*, (9) *SLA mechanism,* (10) *Instructional materials*, (11) *Assessment*, (12) *Curriculum and program development*, and (13) *Review, surveys, meta-studies*. The rationale behind this selection was that their framework covered a broad range of topics and clear-cut descriptors were provided, making adoption and adaptation plausible. However, the list was subject to modification. As coding proceeded, the extended list of research areas was revised until reaching the final scheme that could cover all the thesis titles. To handle the subjective nature of coding, apart from the researcher, another coder with considerable experience in teaching English and supervising ELT students was invited to ensure the reliability of the coding protocol. The external coder was given a modified list of eleven research categories with examples for each category. A short trial session was held to secure a comprehensive understanding of each category. Owing to the labour-intensiveness of coding, the external coder was assigned to code only 50% (100 theses) of the theses. The inter-coder agreement was measured with a chance-corrected agreement measure, Cohen's Kappa coefficient, κ, and the percentage to agreement. A κ value of 9.33 was generated with an agreement of 95%. Once the two coding systems were finalised, the theses were categorised based on the primary research aim. The main body of the papers were reviewed, in addition to the titles and abstracts, as that part of the paper provided information regarding the actual research objective of the theses. The frequency count of each research area was also generated to address the differences between research areas.

Results and Discussion

Areas of Focus

Analysing the themes of taught courses yielded twelve content areas (see Table 1). The number of compulsory courses varied from programme to programme ranging from five to 14 courses as some programmes tended to add more weight to elective courses giving students more freedom to choose courses that suited their goals and interests. While most of the areas presented here are language-related, if not English-related, there was an area that did not focus on language but was offered to prepare students for studying advanced content in the field of education. Offered in five programmes, *General teaching/education* courses covered a broad range of teaching practices and education, including pedagogical sciences, principles of education, educational management, teacher professional development, educational psychology, and educational philosophy. Students were required to take these courses prior to or at the same time as core courses. In addition to *General teaching/*

education, General language studies aimed to prepare students for language elements in the ELT programmes; however, this type of course did not specifically target any particular ELT elements but rather addressed the general foundation of language studies.

In some programmes, foundation courses could be offered as *English proficiency development,* which aimed to improve students' English language skills essential for graduate-level work such as reading scholarly articles, academic writing, and giving an academic presentation. Nevertheless, *English proficiency development* courses were not always compulsory. In *ELTP2*, for example, students who were L1 users of English or could demonstrate sufficient English language proficiency were able to waive such courses. *English proficiency development* was one of the most preferred areas, with twelve courses identified in six programmes. However, these courses were omitted in some programmes. *ELTP1, ELTP3, ELTP6,* and *ELTP9* did not offer any English skills courses but rather required all applicants to demonstrate a certain level of English proficiency before applying. Qualified results of English proficiency tests such as 79 for TOEFL IBT and 6.5 for IELTS were typically used as one of the admission requirements to ensure readiness for graduate-level work and meet the standard for English-related programmes. Apart from scores on universal standardised tests, those of some in-house English proficiency tests given by universities in the country are also accepted for some programmes these days.

Table 1
Foundation and core courses offered in ELT master's programmes

Areas	Number of Courses									
	ELTP1	**ELTP2**	**ELTP3**	**ELTP4**	**ELTP5**	**ELTP6**	**ELTP7**	**ELTP8**	**ELTP9**	**ELTP10**
Language teaching methodology	1	1	1	2	2	2	1	1	1	1
Research methodology	1	1	1	1	1	1	2	2	2	1
English proficiency development		1		1	2		4	3		1
General teaching/education			2		3		3	2		1
Language assessment/evaluation	1	1	1	1	1	1	1	1		
Language learning	1	1		1	2	1	1		1	
Applied linguistics				4				1	1	1
Language course/curriculum development		1	1		1	1				
Technology for teaching and learning					1			1		1
Cultural aspects		1			1			1		
Materials development	1						1			
General language studies	1		1							

As for content areas, *Language teaching methodology* was one of the most preferred areas with 13 courses offered, concurring with Stapleton and Shao's (2018) report on the most common knowledge fields of MA TESOL programmes in 16 countries. Offered in all the programmes, this content area covered a range of content: language teaching approaches, principles, and theories. Although some courses in this category were vaguely named, the course descriptions did indicate that, to some degree, they focused on at least one of the elements mentioned above. Another most preferred content area was *Research methodology*. This content area covered various dimensions of research including but not limited to principles, designs, procedures, instrumentation, data collection, and sampling. While most programmes tended to offer one research course that attempted to highlight various research elements, *ELTP2* offered a selection of research courses: quantitative, qualitative, and general methodology. That is, students chose only one of the research courses that they found relevant. Despite being offered as options, these courses were considered core courses, as noted by the curriculum document. Up to that point in my research, *Language teaching methodology* and *Research methodology* were not omitted from any of the programmes, as the former focuses on teaching practice and the latter focuses on research elements, making them key components of ELT master's programmes.

Language assessment/evaluation and *Language learning* were also highly preferred with eight courses identified in eight and seven programmes respectively. Assessment has been one of the essential constituents of education, playing a crucial role in the teaching and learning process, and ELT is no exception. *Language* *assessment/evaluation* deals with a wide variety of topics related to language testing, assessment, and evaluation. The availability of language assessment courses is in response to scholarly calls for the promotion of language assessment literacy amongst teachers, assessment tool developers, policymakers, and even students (e.g., Baker, 2016; Deygers & Malone, 2019). *ELTP9* and *ELTP10*, however, offered this content area as elective courses. While teaching has become the key focus of the programmes, learning has also received considerable attention from programme developers. Seven of the ten programmes offer courses focusing specifically on language learning elements. *Language learning* covers any content area related to the learning and acquisition of a second or foreign language. Language acquisition tends to be the most popular course in this content area as it addresses fundamental aspects of how one successfully acquires a (second or foreign) language. Another preferred content area is *Language curriculum development*, with four courses identified in four programmes. This content area moves beyond the teaching and learning aspects to the curriculum level of language education. It is undeniable that, apart from teaching and administrative duties, English teachers are often involved with course and curriculum development, and this aspect of English teaching renders this content area essential.

There was no other content areas that appeared in more than three programmes. The rest of the taught course credits were devoted to content areas in which the programmes specialised. For example, a content area that yielded the identity of *ELTP2* was *Cultural aspects*. In *ELTP2*, several cultural-related courses were offered as core and elective courses because it specialised in teaching English in global contexts, which involves topics such as English as a lingua franca (ELF), the influence of culture on ELT, and different varieties of English.

Despite being committed to the pedagogical aspects of English, some of the programmes reserved some credits for *Applied linguistics* with seven courses identified across the programmes. This content area allows students to explore the micro-level of language and ways to apply different types of linguistic elements (e.g., phonetics, syntax, semantics, pragmatics) for practical purposes. Nevertheless, some content areas deemed vital in ELT did not receive much attention from programme developers. Notwithstanding urgent calls for the incorporation of technology into ELT, technology courses have still not been made compulsory in most ELT master's programmes in Thailand. Only three compulsory technology-related courses were noted, all of which tended to focus on the roles and applications of information and communication technology (ICT) in ELT. *Materials development* is another content area that deserves more attention from programme developers. Like curriculum development, designing and selecting instructional materials are essential for English teachers at all levels. However, only two core courses on *Materials development* were found, both of which emphasised not only ELT materials design but also evaluation, analysis, and adaptation.

Research Trends

The research categories presented in this study are different from those of Stapleton and Shao (2017) because a new boundary for all the research categories was needed in response to the distinctiveness of the dataset. Thus, one main coding scheme was proposed for categorising the theses into different research areas and another for categorising theses that fell into the *Instructional effects* category into different target areas that researchers attempted to improve.

Table 2
Research areas of ELT master's theses

Research Area	Example	No. of Theses	%
Instructional effects	effects of instructional approaches on language skills, performance, and cognitive aspects	102	50.7
Learner cognition/practice	learners' beliefs, attitudes, self-reflection, processing, difficulties, awareness, and learning strategies	24	11.9
Linguistic analysis	discourse analysis, genre analysis, lexical analysis, and corpus linguistics	17	8.4
Teacher cognition/practice	teachers' beliefs, attitudes, self-perceived, senses, teaching strategies, and teaching styles	16	7.9
Culture/Identity	intercultural communication, World Englishes, learner identities, and gender	11	5.4

Research Area	Example	No. of Theses	%
English for professionals	professional communication and professionals' needs of English	8	3.9
Influential factors	factors affecting comprehension, causing willingness, and facilitating learning	8	3.9
Assessment	test development, test evaluation, and washback effects	7	3.4
Curriculum/Programme	programme evaluation, curriculum design, and lesson development	4	1.9
Instructional materials	materials development and textbook evaluation	3	1.4
Teacher education	pre- and in-service teacher practice, the teaching profession, and teacher professional development	1	0.4
Total		**201**	

Eleven research areas were identified, covering a wide variety of English-related topics. Table 2 shows that *Instructional effects* dominated ELT research conducted in master's programmes in Thailand with 50.7 % of the theses. *Instructional effects* covered studies that primarily focused on the influences of instructional approaches on students' skills, performance, and cognitive aspects. This finding concurs with previous thesis research trends in Taiwan (Lin & Cheng, 2010) and Turkey (Kirmizi, 2012) in that instructional approaches and language skills were the most researched in the field. The dominance of this research area also existed in the top international publication platforms such as the *Language Teaching Research* journal, in which almost half of the articles focused on the effects of instructional approaches (Stapleton & Shao, 2017). This figure aligns well with that of the taught courses under *Language teaching methodology*, which was the most common content area of the programmes.

While Stapleton and Shao's (2017) study focused on instructional approaches, which were further categorised into sub-categories, the current study emphasised target skills, performance, and cognitive aspects that researchers attempted to investigate. Thus, the theses under *Instructional effects* were further categorised into nine sub-categories (see Table 3). *Reading* was the most researched target area with 21.5% of the theses, followed by *Writing* and *Speaking*. Such findings align well with the ELT research trends in Egypt uncovered by Abdel Latif's (2018) study finding that, among the four skills of language, reading was the most researched area, followed by writing, speaking, and listening respectively. *Integrated skills* and *Cognitive aspects* also received some attention with 13 (12.7%) theses each. Studies on *Integrated skills* attempt to improve two or more language skills simultaneously (e.g., speaking and listening, reading and writing, communicative skills). On the other hand, *Cognitive aspects* cover studies that explore learner cognition towards instructional approaches (e.g., learner's motivation, attitudes, beliefs, satisfaction, willingness). However, *Vocabulary* and *Grammar* were hardly researched, compared with the four skills of language, with only nine and seven theses identified respectively. Despite being one of the four language skills, *Listening*, as an isolated skill, was least researched; however, listening and pronunciation were often treated as an integrated area along with speaking, turning them into *Integrated skills* rather than *Listening* or *Pronunciation*.

Table 3
Sub-areas of Instructional effects

Sub-areas	No. of Theses	%
Reading	22	21.5
Writing	18	17.6
Speaking	16	15.6
Integrated skills	13	12.7
Cognitive aspects	13	12.7
Vocabulary	9	8.8
Grammar	7	6.8
Listening	2	1.9
Pronunciation	2	1.9
Total	**102**	

Apart from *Instructional effects,* the other half of the theses covered a range of topics in the remaining research areas. The theses that fell into the rest of the research areas were not done in an experimental manner but rather aimed to explore, develop, or evaluate particular English-related elements. The second most researched area was *Learner cognition/practice* with 24 (11%) theses, followed by *Linguistic analysis* with 17 (8.4%) theses, and *Teacher cognition/practice* with 16 (7.9%) theses identified. *Learner cognition/practice* and *Teacher cognition/ practice* cover a broad scope of cognitive aspects and practices of learners and teachers (e.g., beliefs, attitudes, awareness, learning/teaching strategies). These two research areas differ from *Cognitive aspects* under *Instructional effects* in that they both were not treated as experimental but rather as exploratory research.

Unlike the aforementioned research areas, *Linguistic analysis* explores a particular linguistic feature in a specific text type such as research articles, master's theses, essays, songs, news articles, comics, or email transactions. This research area has to do with various text analytical approaches such as discourse analysis, genre analysis, lexical analysis, and corpus linguistics. Despite not being offered as a standalone compulsory course, these text analytical approaches were still highly researched. *Culture/Identity* was another topic that was researched in more than ten theses. This research area deals with cultural influences on English and identity issues related to ELT. Topics including, but not limited to, intercultural communication, World Englishes, learner identities, and gender fall into this category. The field of ELT is not limited to teaching and learning in the contexts of formal education. Some students also researched English in other professional settings. Studies on *English for professionals* deal with ELT and the use of English in professional contexts other than formal education. In this research area, a variety of careers researched by students were identified such as train conductors, customs personnel, Royal Army officers, health service professionals, and diving instructors. Some students chose *Influential factors* as their research topics. This covers any studies that aimed at exploring factors that influence teaching and learning.

Some research areas did not match the number of compulsory courses offered in the programmes. Only seven theses fell into *Assessment,* which, as its name suggests, covers any topic related to exams and tests, even though almost all of the programmes offered compulsory assessment courses. Such an imbalance can also be seen in the case of *Curriculum/Programme*, considering the number of curriculum courses offered, and in the case of *Teacher education*, considering the nature of the programmes. *Instructional materials* did not receive much attention with only three theses identified despite the vital role teaching materials play in language classrooms. English textbooks, for example, are considered an integral component of any English classroom lesson (Cortazzi & Jin, 1999; McGrath, 2016; Richards, 2006; Tomlinson, 2008), and materials evaluation is regarded as a global issue (McGrath, 2016; Tomlinson & Masuhara, 2017) leading to a growing body of international literature on textbook evaluation along with materials development.

Conclusion

Although the current study was conducted in a descriptive manner, the ten programmes reviewed in this study rendered results more representative of the areas of focus and research trends of ELT master's programmes in Thailand. Exploring each of the ten programmes revealed a huge variation in the courses deemed compulsory. *Language teaching methodology* as a standard content area indicated that the ELT master's programmes nationwide have emphasised instructional approaches and strategies. As Stapleton and Shao (2018) argued, Richards's (2008) claim that master's programmes generally focus on language acquisition rather than language teaching may not be entirely true, at least in the Thai context. The current study confirms that even at a national scale, pedagogical content still dominates master's research in Thailand. This trend is subject to change, but as of now, this content area has received considerable attention from international and Thai ELT master's programmes. Another prime focus of Thai ELT master's programmes is *Research methodology.* Although the research content area is not widely featured in international ELT master's programmes (Stapleton & Shao, 2018), it serves as a central focus in Thai ELT master's programmes along with the pedagogical content area. The programmes are well-suited for both classroom teachers and those who are seeking a teaching position at a higher education institution, which often requires research skills and experience. This also implies that the programmes have been adjusted to the growing demand for qualified English teachers at all stages of education and responded to the call for more research content. In this sense, ELT master's programmes in Thailand are aimed at preparing current and prospective English teachers for both classroom instruction and research so as to make them strong candidates for any English teaching position. Such findings, nevertheless, do not conform to degree-type programmes. Several higher education institutions in the US provide the benchmark for how to ascertain the difference between an MEd and MA degree. The Faculty of Education at the University of Kansas, for example, provides an MA degree that is scholarly-oriented, with an emphasis on sophisticated field knowledge (e.g., theories, procedures, research); on the contrary, their MEd degree is practice-focused, focusing on teaching practices (e.g., instructional approaches, material designs, curriculum development). While the MA degree prepares students for a teaching career in higher education institutions and further education, the MEd degree is suitable for licensed teachers who hope to improve their teaching practices or move into an administrative role. However, such rules may not apply to Thai ELT master's programmes. Upon browsing through the compulsory courses offered, it can be concluded that the MEd and MA degrees in teaching English are not treated differently. They did not seem to be one or the other but rather both were career- and research-oriented.

Dominating international ELT research for over two decades (Stapleton & Shao, 2017), *Instructional effects* research area is the most preferred topic in Thailand as well. The growth of this research area in ELT master's programmes may be influenced by students' experience as classroom practitioners and the fact that classroom teachers are expected to implement new teaching methods, even though there is an abundance of other topics to choose from. The imbalance between compulsory courses and research options for theses was noticeable. The number of studies focusing on assessment was relatively low, especially when considering the number of compulsory assessment courses offered in the programmes and the importance of assessment in the teaching and learning process. The areas of curriculum planning, instructional materials, and teacher education also suffered from a shortage of research.

In closing, this study casts light on the current situation with regard to ELT master's programmes in Thailand in terms of areas of focus and research trends. The findings presented here suggest issues that require the attention of programme management, lecturers, and students. Since ELT deals with various aspects of education, content and research areas should not be limited to classroom instruction or an investigation of the efficacy of instructional approaches. I would encourage programme management and lecturers to treat other content areas (e.g., assessment, curriculum development, materials development) as a gateway to new research opportunities. Students should be exposed to a wide variety of topics and encouraged to explore research areas relevant to their practices. International trends in ELT and ELT research should be reviewed from time to time to keep community members up to date and move the community forward. This study was conducted in response to the absence of literature on Thai ELT master's programmes and the first step towards the understanding of Thai ELT master's programmes as a teaching, learning, and research community.

Like any research, this study has its limitations that need to be addressed. The scope of this study was limited to English teaching programmes; however, there are several other language-related programmes (e.g., Applied Linguistics, Linguistics, English for as Specific Purposes, English as an International Language) that can be explored for a better understanding of the ELT graduate community as they can also reflect areas of focus, research trends, and other aspects of ELT. Future research should consider reviewing such programmes to obtain more conclusive findings. Although this study claimed that its findings indicated research trends from the last five years, some recently submitted master's theses were not collected because they were not yet available for public access, especially those submitted in late 2018. With all the theses submitted between the years of 2014 and 2018, the findings would have been more conclusive. The ELT master's programmes that were omitted in this study because of temporary closure should be considered by future researchers once they resume operation. Considering the growing importance of ELT master's programmes, future research should also consider exploring other aspects such as challenges faced by programme management, student expectations, and student satisfaction as they contribute immensely to the development of the programmes and the ELT community.

Acknowledgements

This work was financially supported by the Faculty of Education, Ramkhamhaeng University. My sincere gratitude goes to the external coder for participating in the inter-rater reliability evaluation. I deeply thank the anonymous reviewers and the editorial team for the careful reading and insightful suggestions.

References

Abdel Latif, M. M. (2018). English language teaching research in Egypt: Trends and challenges. *Journal of Multilingual and Multicultural Development*, *39*(9), 818–829. https://doi.org/10.1080/01434632.2018.1445259

Baecher, L. (2012). Feedback from the field: What novice preK-12 ESL teachers want to tell. *TESOL Quarterly*, *46*(3), 578–588. https://doi.org/10.1002/tesq.43

Baker, B. (2016). Language assessment literacy as professional competence: The case of Canadian admissions decision makers. *Canadian Journal of Applied Linguistics*, *19*(1), 63–83.

Bani-Khaled, T. A. (2012). Jordanian English language research: A meta-analysis. *International Journal of Humanities and Social Science*, *2*(14), 55–66.

Brady, B., & Gulikers, G. (2004). Enhancing the MA in TESOL practicum course for nonnative English-speaking student teachers. In L. D. Kamhi-Stein (Ed.), *Learning and teaching from experience: Perspectives on non-native English-speaking professionals* (pp. 206–229). The University of Michigan Press.

Copland, F., & Garton, S. (2012). Life after online learning. In L. England (Ed.), *Online language teacher education: TESOL perspectives* (pp. 64–77). Routledge.

Copland, F., Viana, V., Bowker, D., Moran, E., Papageorgiou, I., & Shapira, M. (2017). ELT master's courses in the UK: Students' expectations and experiences. British Council. https://englishagenda.britishcouncil.org/continuing-professional-development/teacher-educator-framework/taking-responsibility-own-professional-development/elt-masters-courses-uk-students-expectations-and-experiences

Cortazzi, M., & Jin, L. X. (1999). Cultural mirrors: Materials and methods in the EFL classroom. In E. Hinkel (Ed.), *Culture in second language teaching* (pp. 196–219). Cambridge University Press.

Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, *36*(3), 347–368. https://doi.org/10.1177/0265532219826390

England, L., & Roberts, C. (1989, March 7-11). A survey of foreign students in MA-TESOL programmes [Paper presentation]. The Annual Meeting of the Teachers of English to Speakers of Other Languages, San Antonio, TX. https://files.eric.ed.gov/fulltext/ED306764.pdf

Fordyce, K., & Hennebry, M. L. (2013, September 6). An investigation to the degree of alignment between student expectations and actual experiences on a one-year master's in TESOL at a UK university [Paper presentation]. The BAAL Conference, Edinburgh.

Hasrati, M., & Tavakoli, P. (2015). Globalisation and MA TESOL programs in the UK. *Higher Education*, *69*(4), 547-565. https://doi.org/10.1007/s10734-014-9790-5

Hughes, H. E., & Bruce, C. S. (2013). International students' experiences of informed learning: A pedagogical case study. *International Journal of Pedagogies and Learning*, *8*(2), 106–119. https://doi.org/10.5172/ijpl.2013.8.2.106

Kaur, A., Young, D., & Kirkpatrick, R. (2016). English education policy in Thailand: Why the poor results? In R. Kirkpatrick (Ed.), *English language education policy in Asia (Language Policy Book 11)* (pp. 345–361). Springer International Publishing.

Kirmizi, O. (2012). Research trends in MA ELT programs in Turkey. *Procedia-Social and Behavioral Sciences*, *46*, 4687–4691. https://doi.org/10.1016/j.sbspro.2012.06.319

Li, B., & Tin, T. B. (2013). Exploring the expectations and perceptions of non-native English-speaking students in master's level TESOL programs. *New Zealand Studies in Applied Linguistics*, *19*(2), 21–35.

Lin, L., & Cheng, C. (2010). Research trends in selected MA TESOL programs in Taiwan: A preliminary of content analysis of master's theses from 2003–2007. *Asian EFL Journal*, *12*(4), 126–139.

McGrath, I. (2016). *Materials evaluation and design for language teaching* (2nd ed.). Edinburgh University Press.

Phakiti, A., & Li, L. (2011). General academic difficulties and reading and writing difficulties among Asian ESL postgraduate students in TESOL at an Australian university. *RELC Journal*, *42*(3), 227–264. https://doi.org/10.1177/0033688211421417

Richards, J. C. (2006). Materials development and research–Making the connection. *RELC Journal*, *37*(1), 5–26. https://doi.org/10.1177/0033688206063470

Richards, J. C. (2008). Second language teacher education today. *RELC Journal, 39*, 158–177. https://doi.org/10.1177/0033688208092182

Schartner, A., & Young, T. (2015). Culture shock or love at first sight? Exploring the 'Honeymoon' stage of the international student sojourn. In A. H. Fabricius & B. Preisler (Eds.), *Transcultural interaction and linguistic diversity in higher education* (pp. 12–33). Palgrave Macmillan.

Stapleton, P., & Shao, Q. (2018). A worldwide survey of MATESOL programs in 2014: Patterns and perspectives.

Language Teaching Research, *22*(1), 10–28. https://doi.org/10.1177/1362168816659681
Stapleton, P., & Shao, Q. (2017). Research in language teaching over two decades: A retrospective of the first 20 volumes of Language Teaching Research. *Language Teaching Research*, *22*(3), 350–369. https://doi.org/10.1177/1362168816688929
Thumvichit, A. (2020). Why and which ELT master's program? Factors influencing student enrollment decisions. *The International Journal of Learning in Higher Education*, *28*(1), 21–35. https://doi.org/10.18848/2327-7955/CGP/v28i01/21-35
Tomlinson, B. (Ed.). (2008). *English language teaching materials: A critical review*. Continuum.
Tomlinson, B., & Masuhara, H. (2017). *The complete guide to the theory and practice of materials development for language learning.* Wiley Blackwell.
Tseng, C. (2013). *Literacy experiences and disciplinary socialisation of second language students in an MA TESOL program* [Unpublished doctoral dissertation]. University of California. https://escholarship.org/uc/item/3fg1w7pw
Tůma, F., & Pišova, M. (2013). Trends in foreign language didactics research: A thematic analysis of PhD dissertations from the Czech Republic and abroad (2006–2012). *The New Educational Review*, *34*, 125–137.
Wenfeng, W., & Gao, X. (2008). English language education in China: A review of selected research. *Journal of Multilingual and Multicultural Development*, *29*(5), 380–399. https://doi.org/10.1080/01434630802147908

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Vaghei, S., Esfahani, F. R., & Shafiee, S. (2020). The Relationship Between Language Mindsets and Feedback Preferences in L2 Writings of EFL Learners. *Journal of Language and Education, 6*(4), 183-203. https://doi.org/10.17323/jle.2020.10792

The Relationship Between Language Mindsets and Feedback Preferences in L2 Writings of EFL Learners

Shima Vaghei, Fariba Rahimi Esfahani, Sajad Shafiee
Islamic Azad University

Correspondence concerning this article should be addressed to Fariba Rahimi Esfahani, Department of English,Faculty of Humanities, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran.
Email: rahimi_fariba@yahoo.com

The present study intended to investigate intermediate Iranian EFL learners' language mindsets and examine the possible relationships between language mindsets and feedback preferences in L2 writing. To achieve these aims, 150 EFL volunteer learners were recruited from several language institutes in Isfahan, Iran, and their language proficiency level was determined through a proficiency test. The learners were then given the Language Mindsets Questionnaire and the Feedback Preferences Scale to fill out. Frequency counts, mean scores, one-sample t-tests, and SEM in Smart PLS were employed to analyze the collected data. The results of the study indicated that for the three subcomponents of the mindsets questionnaire (i.e., general language intelligence beliefs, second language aptitude beliefs, and age sensitivity beliefs about language learning), the learners did not agree (though not significantly) with the entity items, while they expressed significant agreement with the incremental items. The SEM results also revealed that the model (examining the relationship between entity/incremental mindsets and feedback preferences) indicated that entity mindsets were a significant predictor of feedback preferences, yet the incremental mindsets failed to do so. Among the feedback types, EFL learners' preferences, in a deductive order, were found to be for commentary and conferencing significantly, and then for peer correction, prompts, self-correction to a non-significant extent. The significance of the obtained results are presented and the implications of the study are enumerated.

Keywords: feedback preferences, L2 writing, language mindsets

Introduction

Corrective feedback (CF) has been basically conceptualized as reactions and responses to students' utterances that include errors (Ellis, 2006). It is also a compound phenomenon with several functions (Chaudron, 1988). As noted by Ellis, Loewen, and Erlam (2006), corrective feedback takes the form of answers to learner utterances that have an error. The answers can contain an indication that an error has been made, the delivery of the correct target language item, metalinguistic information about the essence of the error, or any mixture of these.

The aim of presenting feedback is to help second language (L2) students recognize a problem in their production, resulting in the correct use of the form. Corrective feedback plays a crucial role in the kind of scaffolding that teachers need to create for language learners to promote their L2 (Lyster & Ranta, 1997; Lyster, Saito, & Sato, 2013). Theoretical perspectives, from cognitive to social-oriented ones, suggest corrective feedback is not only helpful but also necessary for pushing students forward in their L2 growth (Lyster et al., 2013). Past research has shown that giving feedback effectively contributes to grammatical, morphological, and phonological development (Lyster et al., 2013).

CF is by no means an under-researched area of investigation in SLA as past research in the realm of feedback has focused on grammar (e.g., Ebrahimzadeh & Khodareza, 2016; Mir & Ghornavi, 2017; Zohdi Rad & Ghafournia, 2017), vocabulary (e.g., Noori Khaneghah, 2016; Sippel, 2019), pronunciation (e.g., Karami & Heidari Darani, 2018; Naeimi, Saeidi, & Behnam, 2017), writing (e.g., Abbaspour, Atai, & Maftoon, 2020; Gonzales, Tejeda, Krous, & Vasquez, 2018; Rezaei, Izadpanah, & Shahnavaz, 2017), and speaking (e.g., Farrokhi, Zohrabi, & Chehr azad, 2018; Ramadhani, 2019; Tesnim, 2019; Zhai & Gao, 2018) among other things. Besides, the ways through which learners prefer to receive feedback have attracted a great many scholars' attention as studies on written corrective feedback (e.g., Cohen & Cavalcanti, 1990; Diab, 2005; Halimi, 2008; Lee, 2004; Leki, 1991; Rennie, 2000; Saito, 1994; Wang, 2010) and on the importance of feedback preferences by L2 learners (e.g., Diab, 2005; Ferris, 1999, 2004; Hedgcock & Lefkowitz. 1994; Lee, 2005; Leki, 1991; Truscott, 1996, 1999) abound.

There is no escaping the fact that second language learners have different learning styles and preferences for teaching and learning (Reid, 1997; Katayama, 2007). Some scholars align with this assumption and argue that learners' undesirable feedback might not be effective for developing their writing skills as it may frustrate and demotivate them. Conversely, learners' preferred feedback can positively affect their learning and the development of their writing skills (Hedgcock & Lefkowitz, 1996; Lee, 2005; Katayama, 2007; Schulz, 1996).

What is now missing from the colossal body of scholarly research on CF is whether feedback preferences for L2 written production vary according to the learners' personal attributes and language-learning beliefs, styles, and strategies. As an attempt to demystify this riddle, at least in part, and shed light on this rather unnoticed sphere of CF, the present study sought to examine the L2 writing feedback preferences of EFL learners who held incremental vs. entity mindsets about language learning.

The concept of mindset (MS), or implicit theories, dates back to Kelly's (1955) work, which presented some lay theories of the way people perceive the self and others. Recently, in the area of education, more concepts have been engendered via the construct of MS, in which the assumptions and beliefs of certain human traits have largely been connected to Carol Dweck and her associates' works (e.g., Blackwell, Trzesniewski, & Dweck, 2007; Chiu, Hong, & Dweck, 1997; Dweck, 2006; Dweck & Molden, 2007; Dweck, Chiu, & Hong, 1995; Hong, Chiu, & Dweck, 1999). Additionally, mindsets have a close association with a number of theoretical and empirical studies on second language learning (e.g., Barcelos, 2003; Benson & Lor, 1999; Cotterall, 1999; Horwitz, 1998; White, 2008) where the relationships between mindsets and language-learning behavior have been investigated.

Within the context of language learning, the fixed MS (alternatively referred to as entity MS) belief in language learning depends on a fixed and inborn talent, while the growth MS (alternatively called incremental MS) belief depends on controllable factors such as hard work and continuous training. In the area of foreign language teaching and learning, it is common for people to possess the belief that some people are born with a special talent in a certain domain (Mercer & Ryan, 2010), which is called aptitude, and its development in language learning is significant (Robinson, 2005; Sternberg, 2002). People with an entity MS may hold the belief that having a 'gift' for languages is important when learning a language; therefore, it is a waste of effort to attempt to improve the language in question since it is impossible for poor language learners to develop as a linguist by any means (Mercer, 2012). In this regard, the aptitude of individuals with an entity or incremental MS could be observed to have different learning outcomes.

The current study addressed language mindsets as a new addition to the literature of second language acquisition, examining the links between foreign language (FL) learners' different mindsets and their different feedback preferences in relation to L2 writing.

Literature Review

Language Mindsets

The concept of mindsets, or implicit theories, generally refers to people's tacit beliefs about whether human attributes like intelligence, personality, capabilities, and so on are fixed or malleable (Dweck, 2007). Fixed mindsets (alternatively referred to as entity theories) are the beliefs that a person's abilities or attributes are predetermined and hence cannot be subject to change, while growth mindsets (alternatively called incremental theories) are the beliefs that a person's qualities are apt to be cultivated by means of hard work and the application of strategies. Such mindsets can help students set reasonable goals and take appropriate actions in different learning situations, thus guiding them to different achievement outcomes (Dweck & Leggett, 1988). More specifically, students who strongly hold an entity mindset tend to set more performance goals and prove their ability via performance. They see criticism as evidence of a lack of capability and view challenges as threats to their sense of achievement. In order to preclude their self-esteem from negative performance and feedback, they shun demanding tasks, even tasks that are likely to ameliorate their ability. On the contrary, chances are that students who have a strong incremental mindset set learning goals and concentrate on their improvement and learning processes. They see overcoming problems and challenges as a way of developing their talents and use negative feedback as a means to improve. In comparison with the students with strong fixed mindset, those students who have a strong growth mindset are more likely to welcome challenges and deal with failure situations more positively. Consequently, growth mindsets are arguably a key to students' perseverance and long-term achievement (Yeager & Dweck, 2012).

The trait of mindsets and their impacts have been extensively studied over the past two decades in many educational areas like math and science, which provided far-reaching implications for academic achievement and success in education (see Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013 for a meta-analysis). Recent studies in the domain of language learning have started to examine the conceptualization and motivational impacts of language mindsets (see Lou & Noels, 2016; Mercer & Ryan, 2010). Given that mindsets are domain-specific beliefs, language mindsets, or the beliefs about whether the ability to learn a second language is fixed or can be developed, are shown to be only moderately correlated with mindsets in other areas, such as abilities to learn math and/or sports (see Lou & Noels, 2017). Such findings imply that most students probably hold a mixture of fixed and growth mindsets in different domains (Dweck, 2019). As the mindsets that people hold orient their approach to the learning task at hand (Lou, 2014), different types of mindsets are likely to cause EFL learners to perceive learning-related issues differently and employ certain types of learning styles and strategies; hence, they are arguably apt to cause language learners to welcome certain types of tendencies and preferences, including corrective feedback preferences in general and feedback on L2 writing in particular.

Corrective Feedback

Corrective feedback could be defined as "responses to learner utterances containing an error" (Ellis, 2006, p. 26); they represent a complex phenomenon that is used to perform several functions (Chaudron, 1988). The motive behind providing feedback is to help L2 students recognize a problem in their linguistic production and facilitate in the correct use of that structure. Corrective feedback has a central role in the kind of support that teachers need to provide for language learners to improve their L2 (Lyster & Ranta, 1997; Lyster, et al., 2013).

Corrective feedback in an EFL classroom context could differ in terms of being implicit or explicit (Lyster et al., 2013). Students may be exposed to explicit corrective feedback, more often than not, in the form of metalinguistic explanation as well as explicit correction. Furthermore, implicit corrective feedback may take the form of recasts. Past research has indicated that implicit (e.g., Lyster & Ranta, 1997; McDonough & Mackey, 2006) and explicit corrective feedback (e.g., Carroll, Swain, & Roberge, 1992; Ellis, 2012) have positive impacts on the acquisition of different grammatical, phonological, and morphosyntactic forms.

Empirical Background

As an example of a study on feedback preferences and their relationships with individual attributes of L2 learners, Roy's (2019) study investigated multilingual writers' preferences for audio and/or written feedback, and examined the effect of feedback format on their revision processes. To meet the aims of the study, eight multilingual writers were chosen to be interviewed and their first drafts as well as revised drafts of the final research papers for which they had received audio and written feedback were compared by means of the 'Compare' option in Microsoft Word. Additionally, an early-semester participatory survey and reflection survey were administered among the multilingual writers of the composition course. The obtained results revealed that multilingual writers expected to receive directive explicit feedback from their course instructors. They mentioned that audio feedback was far better for global commentary while written feedback was a better option better for local commentary. The participants' perceptions regarding the effectiveness of audio vs. written feedback on their revision process differed depending on their self-efficacy. No significant effect was observed for the feedback format on the participants' revision process. A positive correlation was found, however, between their preferred type of feedback and their self-perceived English listening proficiency. Those who enjoyed higher self-efficacy and were confident about their English listening proficiency preferred to be given audio feedback rather than written feedback. The obtained results could have immediate implications for L2 writing instructors. This study proposed that L2 writing instructors ought to adopt the approach of providing an amalgamation of feedback formats consisting of both audio and written feedback in order to promote multilingual writers' overall writing skills. It could also be suggested that L2 writing instructors ought to think about providing some relevant grammar lessons for their students in L2 writing courses. Finally, this study recommended further investigation of the potential of audio feedback in arousing student-teacher connections and relationships, especially in online composition courses.

Roy's (2019) study can provide a spark for the idea that feedback preferences could have relationships with other individual attributes (e.g., language-learning mindsets among other things). With respect to the promising results of mindset interventions in education in general, setting up interventions is a crucial direction for language education. Nonetheless, even with recent interest in mindsets in the field of SLA (Mercer & Ryan, 2009; Lou & Noels, 2017), there has been scant research on the nature of altering students' language mindsets and examining their relationship with other language-learning variables. Several double-blind randomized experiments in this area revealed promising results, showing that students who jotted a reflection down after reading a paper that supported the nurturing of language-learning ability confirmed growth language mindsets more strongly than those students who read and wrote about a paper that was not pertinent to language learning (e.g., Lou & Noels, 2019a). Encouraging growth language mindsets can also end in stronger learning goals, more positive effort beliefs, and reduced levels of anxiety among ESL students in a laboratory context (Lou & Noels, 2016, 2019a). A recent experimental study revealed that prompting growth language mindsets could decrease ESL students' feelings of rejection and trigger them to use an L2, especially among those who had lower L2 competence (Lou & Noels, 2019c).

Lou (2014) utilized Dweck's (1999) implicit theories framework in an L2 context to figure out the causal relationships among L2 students' mindsets, goal orientations (that is, learning goals, performance-approach goals, and performance-avoidance goals), and responses to failure situations (that is, mastery responses, helpless responses, anxiety, and fear of failure). A total of 150 L2 university students were randomly assigned to two experimental conditions in which different mindsets were induced; afterwards, they filled out a questionnaire on their L2 goals and responses in failure situations. The results of the study indicated that after being primed for an incremental mindset, irrespective of their perceived L2 ability, participants set higher learning goals and, in turn, expressed more mastery-orientated responses in failure situations. On the other hand, L2 learners who were exposed to priming for an entity mindset, in cases where they perceived of themselves as highly proficient L2 users, set higher performance-approach goals and, in turn, were more fearful of failure situations.

In much the same way the present study intended to cast light on the still-enigmatic issue of CF through the lens of the individual attribute of language mindsets, Albalawi's (2017) study intended to examine the complex dynamism of L2 demotivation. This study was an attempt to modify and reform previous conceptualizations about demotivation and move L2 demotivation mainstream research into a new phase that underscored the complexity of its process as well as its development. The demotivational, motivational, and remotivational triad of L2 learners was studied through the lens of several key psychological and theoretical constructs, which included language mindsets, personality hardiness, learner helplessness, as well as an L2 motivational self-system. The study was conducted in two phases that examined the demotivation of female L2 university students in Saudi Arabia by using a range of research methodologies such as qualitative in-depth interviews, quantitative surveys, and SEM.

First, an exploratory qualitative study was conducted, which intended to examine the L2 learners' different perspectives on their language-learning experiences and their perceptions of different demotivating factors. Qualitative semi-structured interviews were planned and conducted with 13 English language learners and an analysis of the qualitative data unearthed that language-learning mindsets played an important part in the language learner's motivation, demotivation, remotivation, and in their resilience/vulnerability.

In the second phase of the study, confirmatory quantitative analysis was performed aiming to examine the impacts of holding a particular language-learning mindset on the learners' L2 demotivation. While making use of the key variables that emerged in the qualitative phase of the study, the researcher designed and administered a questionnaire to 2044 foundation-year university students. In fact, several tests were conducted to investigate (a) the relationships between the variables, (b) the possible differences between the growth mindset of L2 learners and the fixed mindset L2 learners, and (c) the putative differences between resilient and vulnerable L2 learners. The results obtained from the quantitative phase of the study confirmed all the hypothesized relationships and served to establish an empirical link between L2 learning mindset and both L2 demotivation and L2 resilience.

Finally, a structural model that posited that L2 demotivation could be predicted by the fixed language-learning mindset was hypothesized. Structural equation modeling (SEM) was conducted to empirically test and examine the hypothesized model. A series of causal relationships was simultaneously examined in this model. The SEM analysis confirmed all the hypothesized causal relationships and indicated that L2 demotivation could be accounted for and predicted positively and directly by the fixed L2 learning mindset. The results also revealed that the fixed L2 learning mindset could indirectly end in L2 demotivation by lowering the ability to manifest a positive ideal L2 self and augmenting L2 disappointment.

Albalawi's (2017) study also implied that relationships could exist between language mindsets and other language-learning variables, one of which is preferences for corrective feedback. To capture the nature of Iranian EFL learners' language mindsets and explore the possible relationship between mindset and feedback preferences, the following research questions were formulated: (a) What are Iranian EFL learners' mindsets about language learning? and (b) How do students' language mindsets relate to their feedback preferences?

Methodology

Participants

The population of the study included intermediate EFL learners in Isfahan, Iran, from which a sample of 150 learners was recruited as the participants in this study. The researchers accessed and selected EFL learners from several private language institutes in Isfahan, Iran, who volunteered to take part in the study. They were chosen from among a homogeneous group (in terms of language proficiency) of EFL learners, whose proficiency levels had been assessed through the administration of the Oxford Quick Placement Test (OQPT). Moreover, the participants were native speakers of Persian and included both female (63%) and male (37%) language learners. Their exposure to the English language in language institutes ranged from two to five years (with a mean year of exposure of 3.69), and they were aged between 14 and 22. The sample included both high school students (26%) and university students (74%). Throughout their academic education both at school and language institutes, they had experienced different types of feedback on their oral and written performances.

Instruments

The following instruments were employed in this study: OQPT, a language mindsets inventory, and feedback preferences scale. OQPT is an internationally-recognized, standardized test of language proficiency developed to help place students in different proficiency levels. It includes 60 multiple-choice questions assessing the vocabulary, grammar, and reading comprehension of English language learners. This test has been widely used and its reliability and validity were assured by previous researchers. Based on the OQPT scoring rubric, learners who score between 30 and 37 out of 60 are lower intermediate and those who obtain a score between 40 and 47 could be considered upper intermediate. As intermediate EFL learners were the participants of this study, they included those learners who had received a score ranging from 30 to 47 on the OQPT.

The language mindsets inventory (LMI) employed in the current study was an 18-item survey developed by Lou and Noels (2017), which intended to assess learners' language mindsets. This scale measures language mindsets on a 6-point Likert-type scale ranging from 1 (strongly disagree) to 6 (strongly agree). It includes items related to entity mindsets (e.g., *You have a certain amount of language intelligence, and you can't really do much to change it*) and items related to incremental mindsets (e.g., *No matter what you are, you can significantly change your language intelligence level*) across the three aspects of language mindsets, including general language intelligence beliefs, second language aptitude beliefs, and age sensitivity beliefs about language learning. Thus, each dimension of language mindsets included three incremental and three entity items. The reliability of the questionnaire was checked through Cronbach's formula and it turned out to be .73. To ensure the validity of the questionnaire, it was handed to three experts in the areas of second language learning, education, and psychology. The experts unanimously approved of the validity of the questionnaire.

The feedback preferences scale by Saito (1994) was used to assess the participants' preferences for teacher feedback. The scale assesses feedback preferences across several categories including teacher correction, comments, teacher correction with comments, error identification, peer correction, and correction using prompts. The scale includes 23 items on a Likert-type scale as well as a place for writing further suggestions. The scale is among the most highly used and highly cited scales for measuring learners' preferences for corrective feedback.

Data Collection Procedure

Prior to the commencement of the study, the researcher met a few language institute heads to ask for their permission. In each of the institutes, intermediate classes were listed and the researcher arranged for a placement test to be administered among the intermediate EFL learners who were willing to take part in this study. Among the whole population of EFL learners who were at the intermediate level of language proficiency in those language institutes, which exceeded 800, a total of 150 learners met the criteria of both volunteering and proficiency level (as assess by the OQPT). The learners in the selected sample were given the LMI and the feedback preferences scale to fill out. They were given hard copies of the questionnaires and were asked to return them in a week. Analysis of the data was done through frequency counts, mean scores, and one-sample t-tests for the first research question, and for the other research question, a model of the relations between the language mindsets and feedback preferences was formed and tested using SmartPLS software.

Results

EFL Learners' Mindsets about Language Learning

The first research question of the current study was intended to investigate Iranian EFL learners' mindsets regarding language learning. The required data were collected through the language mindsets inventory, the results of which are displayed in this section. As mentioned above, the language mindsets inventory comprised the three subscales of general language intelligence beliefs (GLB), second language aptitude beliefs (L2B), and age sensitivity beliefs about language learning (ASB), each consisting of an entity and an incremental subcomponent, giving rise to the six subcomponents of GLBENT, GLBINC, L2BENT, L2BINC, ASBENT, and ASBINC, each with three items. These subcomponents are all separately dealt with here. Inasmuch as each choice in this Likert-scale questionnaire carried a point (*Strongly agree* = 6, *Moderately agree* = 5, *Slightly agree* = 4, *Slightly disagree* = 3, *Moderately disagree* = 2, and *Strongly disagree* = 1), the mean score of each questionnaire item was compared with the average value of the choices (i.e. 3.50). This would mean that if the mean score of a questionnaire item was greater than 3.50, the respondents tended to agree with that statement in the item. On the other hand, a mean score less than 3.50 indicated the respondents' inclination to disagree with a statement mentioned in a questionnaire item.

Table 1

GBLENT Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
1	You have a certain amount of language intelligence, and you can't really do much to change it.	8	20	44	28	24	26	3.21	1.45
2	Your language intelligence is something about you that you can't change very much.	15	43	12	25	43	12	3.506	1.57
3	To be honest, you can't really change your language intelligence.	28	28	16	35	18	25	3.58	1.73

In the GBLENT subsection of the questionnaire, the mean score of the first item (M = 3.21) was lower than 3.50, which means that the surveyed EFL learners in this study disagreed that you cannot do much to change your language intelligence. On the other hand, in items # 2 (M = 3.506) and 3 (M = 3.58), they agreed that your language intelligence is something you cannot change (although the two mean scores for these two items were only slightly higher than the average value of the choices). The total mean score for this part of the questionnaire was 3.43, implying that, on the whole, the learners disagreed that you cannot change your language intelligence. Whether this extent of disagreement was of statistical significance or not could be determined by one-sample t-test results, shown and discussed in Table 7 and in the following pages. For the time being, the results of the GBLINC subcomponent of the questionnaire are presented in Table 2.

Table 2
GBLINC Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
4	No matter who you are, you can significantly change your language intelligence level.	26	36	39	19	21	9	4.00	1.47
5	You can always substantially change your language intelligence.	22	45	28	27	22	6	4.00	1.42
6	No matter how much language intelligence you have, you can always change it quite a bit.	38	31	25	21	21	14	4.01	1.65

As shown in Table 2, the mean scores of all of the three items related to the GBLINC subcomponent of the questionnaire turned out to be more than 3.50, indicating that the surveyed EFL learners believed in the incremental nature of language intelligence and concurred that you can change your language intelligence as it is not always fixed. The significance or non-significance of this extent of agreement with the incremental nature of language intelligence will be determined in Table 7. The results of the L2BENT subcomponent of the questionnaire are presented in Table 3.

Table 3
L2BENT Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
7	To a large extent, a person's biological factors (e.g., brain structures) determine his or her abilities to learn new languages.	18	23	33	36	19	21	3.48	1.52
8	It is difficult to change how good you are at foreign languages.	6	34	23	30	36	21	3.20	1.65
9	Many people can never do well in foreign languages even if they try hard because they lack natural language intelligence.	21	30	27	30	16	26	3.54	1.65

Regarding the L2BENT subsection of the questionnaire, two of the mean scores of the questionnaire items were not higher than the average value of the choices; in fact, the EFL learners did not agree that a person's biological factors determine their abilities to learn a new language (item # 7, M = 3.48), and that it is difficult to change how good you are at foreign languages (item # 8, M = 3.20). Nevertheless, the learners agreed (through to a very small extent) that many people cannot do well in foreign languages because they lack the natural language intelligence (item # 9, M = 3.54). Whether this amount of disagreement with the fixed nature of L2 learning is of statistical significance will be determined in Table 7. Results of the L2BINC subcomponent of the questionnaire are shown in Table 4 below.

Table 4

L2BINC Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
10	You can always change your foreign language ability.	21	40	33	22	18	16	3.84	1.54
11	In learning a foreign language, if you work hard at it, you will always get better.	27	48	20	20	28	7	4.03	1.53
12	How good you are at using a foreign language will always improve if you really work at it.	29	34	28	32	20	7	3.99	1.47

All the mean scores for the questionnaire items in the L2BINC subsection of the questionnaire represented values greater than the average value of the choices. Hence, the surveyed EFL learners expressed their agreement with the contention that you can always improve your language-learning ability if you work hard at it. Results of the ASBENT subcomponent of the questionnaire are shown in Table 5.

Table 5

ASBENT Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
13	How well a person speaks a foreign language depends on how early in life at he/she learned it.	24	33	36	27	14	16	3.85	1.53
14	People can't really learn a new language well after they reach adulthood.	10	28	27	25	36	24	3.19	1.53
15	Even if you try, the skill level you achieve in a foreign language will advance very little if you learn it when you are an adult.	18	26	26	24	28	28	3.32	1.64

The first questionnaire item in the ASBENT subsection of the questionnaire had a mean score higher than 3.50; through this item, the respondents expressed their agreement with the claim that language proficiency hinges upon the age when a learner learns a foreign language (item # 13, *M* = 3.85). However, for the other two items here, the surveyed learners disagreed that a language cannot really be learned well in adulthood (item # 14, *M* = 3.19), and that language skills improve very little in adulthood (item # 15, *M* = 3.32). The last questionnaire subsection, i.e., that of the ASBINC dimension, and the students' responses are reproduced in Table 6.

Table 6

ASBINC Subsection of the Language Mindsets Inventory

No.	Statements	Strongly agree	Moderately agree	Slightly agree	Slightly disagree	Moderately disagree	Strongly disagree	Mean	SD
16	Everyone could do well in foreign language if they try hard, whether they are young or old.	24	36	32	27	25	6	3.92	1.44
17	How well a person learns a foreign language does not depend on age; anyone who works hard can be a fluent speaker in that language.	25	43	21	22	27	12	3.87	1.59
18	Regardless of the age at which they start, people can learn another language well.	33	24	28	27	26	12	3.83	1.61

The three questionnaire items in the last subsection of the questionnaire received mean scores above the average value of the choices. More exactly, the surveyed EFL learners believed that everyone, irrespective of their age, can do well at foreign language learning if they try hard (item # 16, *M* = 3.92), that success in learning another language depends more on how much you try more than it does on your age (item # 17, *M* = 3.87), and that people can learn an L2 regardless of the age at which they start (item # 18, *M* = 3.83). The overall mean of the ASBINC subsection of the questionnaire equaled 3.87, which indicates the respondents' agreement with the incremental nature of language learning regardless of the age at which one starts learning a foreign language. The significance or non-significance of the results obtained for the GBLENT, EBLINC, L2BENT, L2BINC, ASBENT, and ASBINC subcomponents of the questionnaire are determined in Table 7 below.

Table 7
One-sample T-test Results for the Language Mindsets Inventory

	Test Value = 3					
	t	df	Overall Mean	Sig. (2-tailed)	95% Confidence Interval of the Difference	
					Lower	Upper
GBL Entity	-.60	2	3.43	.60	-.55	.41
GBL Incremental	151.00	2	4.003	.00	.48	.51
L2B Entity	-.89	2	3.40	.46	-.54	.35
L2B Incremental	7.83	2	3.95	.01	.20	.70
ASB Entity	-.23	2	3.45	.83	-.91	.82
ASB Incremental	14.34	2	3.87	.00	.26	.48
Entity Mindsets	-.94	8	3.43	.37	-.23	.10
Incremental Mindsets	16.82	8	3.94	.00	.38	.50

Table 7 shows that for the GBL, L2B, and ASB subcomponents of the language mindsets questionnaire, the mean scores for the entity items were (though not significantly) lower than 3.50, while the ones for the incremental items were significantly higher than the average value of the choices. Moreover, the entity-mindset beliefs (as a whole) had a lower-than-average mean score that did not reach statistical significance, but the incremental-mindset beliefs received a mean score of 3.94, which was found to be significantly higher than the average value of the choices.

Relationship Between Mindsets and Feedback Preferences

The second research question of the study was formulated to find out the possible relationships between the language mindsets (entity and incremental in this case) of Iranian EFL learners and their feedback preferences in writing. Structural equation modeling (SEM) in SmartPLS was run to find answers to this question. Before presenting the results of this SEM analysis, it is useful to note that the criteria mentioned in Table 8 should be kept in mind and the results of the analysis should be compared with these criteria.

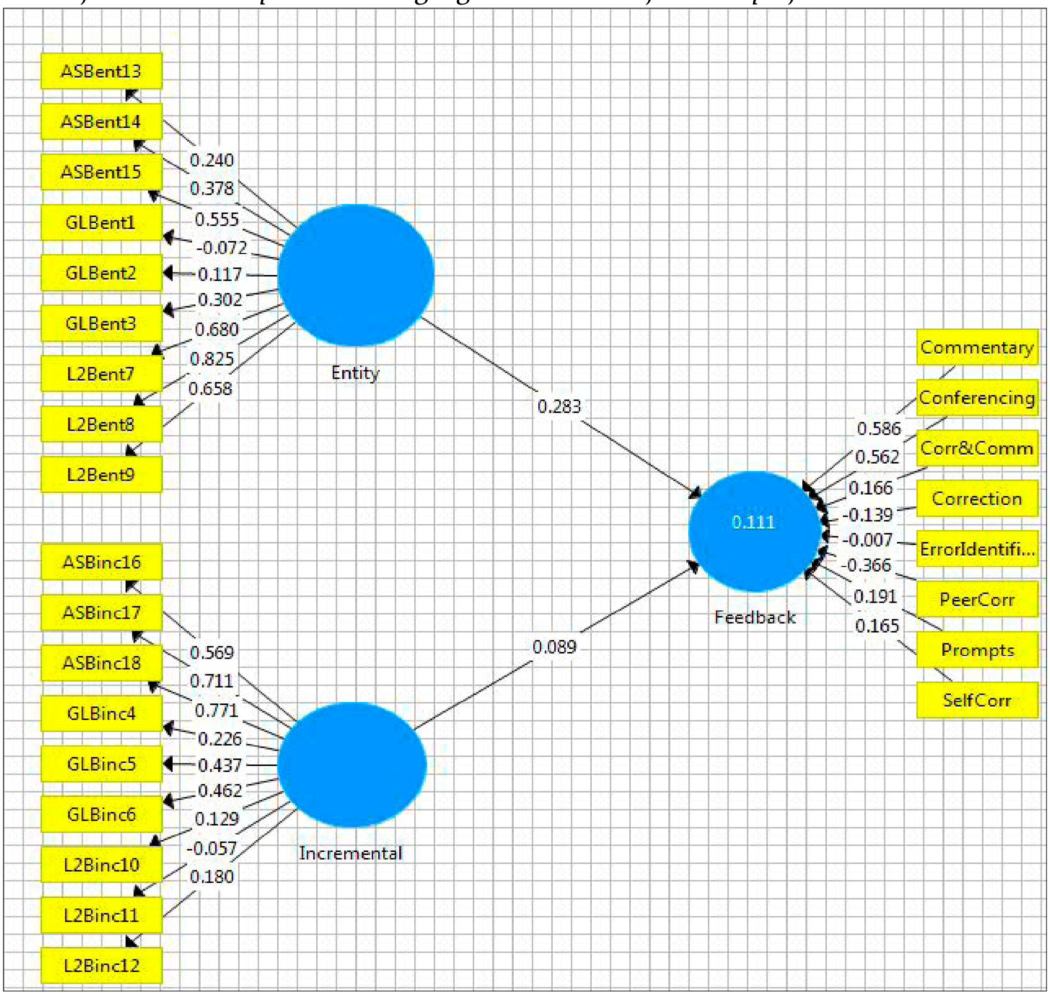
Table 8
Criteria for Evaluating the Variables and Indicators in SmartPLS SEM

Model	Evaluation	Criterion	Range
Reflective Measurement Model	Convergent Validity	Outer Loadings	> .40
		Average Variance Extracted (AVE)	> .50
	Discriminant Validity	Cross Loadings	Higher Coefficients of Indicators
		Fornell-Larcker	Higher Coefficients for AVE Squared
		HTMT	< .90
	Reliability	Cronbach's Alpha	> .70
		Composite Reliability	> .70
Reflective Measurement Model	Validity	Convergent Validity (Redundancy Analysis)	> .70
		Collinearity of Indicators	VIF < 5 Tolerance > .20
		Significance and Relationships of Outer Weights	t > 1.96

Model	Evaluation	Criterion	Range
Structural Model	Testing Hypotheses	Magnitude and Significance of Path Coefficients	t > 11.96 .025 = weak, .50 = moderate, .75 = strong
	Model Evaluation Quality	R^2	.19 = weak, .33 = moderate, .67 = strong
		F^2	.02 = weak, .15 = moderate, .35 = strong
		Q^2	.02 = weak, .15 = moderate, .35 = strong
	Goodness-of-fit	SRMR	< .80
		GOF	.10 = weak, .25 = moderate, .35 = strong

According to the criteria portrayed above, the measurement models (i.e., the relationships between indicators/items and the latent variables) as well as the structural model (i.e., the relationships between the latent independent and dependent variables) could be examined. The model tested in the current study considered as its independent variables the two types of language mindsets (i.e., entity mindset and incremental mindset), while the dependent variable was feedback preferences, as shown in Figure 1 below.

Figure 1
The initial model of the relationship between language mindsets and feedback preferences

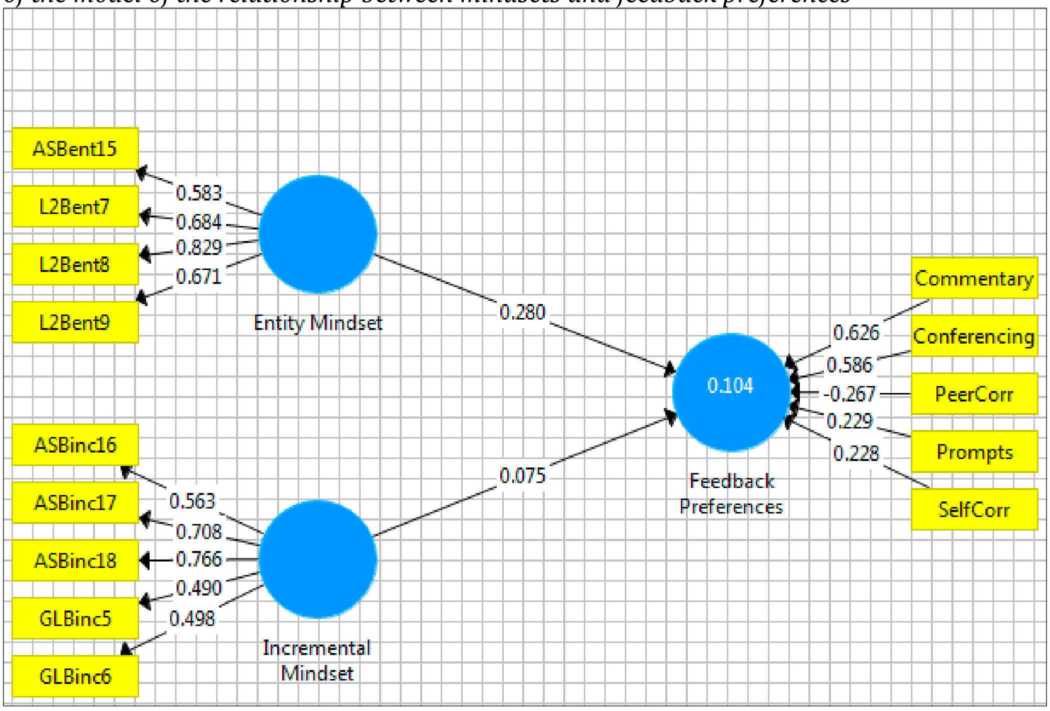


In this initial model, the coefficients of the relationship between the latent variables of the study (i.e., between the independent and dependent variables) are shown. The relationship between these variables is called the structural model. The coefficients of the relationships between indicators/items of each latent variable and that given variable are also displayed. These relationships are referred to as the measurement models.

The first criterion in Table 8 is related to the outer weights in the reflective measurement models. It should be noted that the measurement models for entity and incremental mindsets are reflective (since the arrows are drawn from the latent variables towards the indicators) and the measurement model for feedback is formative (as the arrows are drawn from the indicators to the latent variable). Now, the outer weightings of the measurement models should not fall below .40. Thus, indicators 13, 14, 1, 2, and 3 in the entity model and indicators 4, 10, 11, and 12 in the incremental model were removed from the models. Correction, correction & commentary, and error identification were also removed from the feedback preferences model since they had very low outer loadings. The whole model was run again and the resulting model is presented in Figure 2 below.

Figure 2

Second run of the model of the relationship between mindsets and feedback preferences



As shown in Figure 2 and Table 9, all the outer loadings for the two measurement models of entity mindset and incremental mindset were greater than .40, which indicates acceptable outer loadings for the measurement models.

Table 9

Outer Loadings of the Reflective Measurement Models

Indicator	Entity Mindset	Incremental Mindset
Ent07	.70	
Ent08	.85	
Ent09	.68	
Ent15	.53	
Inc05		.47
Inc06		.49
Inc16		.56
Inc17		.71
Inc18		.77

Table 10 shows the results of convergent validity through AVE.

Table 10

AVE for the Variables of the Study

Variables	Entity Mindset
Entity Mindset	.48
Incremental Mindset	.37

The AVE values for the entity and incremental mindsets were found to be .48 and .37, respectively, which were not very far below acceptable (i.e., .50). Discriminant validity was examined through cross loadings, the results of which are represented in Table 11 below.

Table 11

Cross Loadings of the Indicators

Indicator	Entity Mindset	Incremental Mindset
Ent07	.68	.34
Ent08	.82	.43
Ent09	.67	.16
Ent15	.58	.32
Inc05	.21	.49
Inc06	.38	.49
Inc16	.27	.56
Inc17	.26	.70
Inc18	.30	.76

Discriminant validity is achieved if the values for the loadings of the indicators of a given variable are greater than that given variable, compared to the loadings of its indicators with another variable. In Table 11, it can be seen that the coefficients for the relationship between Ent07, Ent08, Ent09, and Ent15 with the entity mindset are greater than the coefficients for the relationship between these indicators and the incremental mindset. On the other hand, the relationships between Inc05, Inc06, Inc16, Inc17, and Inc18 and incremental mindset are stronger than the relationships between these indicators and the entity mindset. Discriminant validity was also checked via the Fornell-Larcker criterion, as presented in Table 12.

Table 12

Results for the Fornell-Larcker Criterion

Variables	Entity Mindset	Feedback Preferences	Incremental Mindset
Entity Mindset	.69		
Feedback Preferences	.31		
Incremental Mindset	.48	.21	.61

Since the AVE squared value for the entity mindset is .69, which is greater than its relationship with feedback preferences (.31) and with incremental mindset (.48), it could be claimed that the Fornell-Larcker criterion for the discriminant validity was also met. Another index of discriminant validity was the heterotrait-monotrait (HTMT) ratio, the results of which are provided in Table 13.

Table 13

Results for the HTMT

Variables	Entity Mindset	Incremental Mindset
Entity Mindset		.65
Incremental Mindset	.65	

An HTMT ratio of .90 or lower could be deemed as a harbinger of acceptable discriminant validity (Hair, Hult, Ringle, & Sarstedt, 2017). As the HTMT ratio turned out to be .65, it could be concluded that this criterion of discriminant validity was also met. Regarding the reliability analysis, the results for Cronbach's alpha and composite reliability are shown in Table 14.

Table 14
Results for Reliability of the Variables

Variables	Number of Indicators	Cronbach's Alpha	Composite Reliability
Entity Mindset	4	.65	.78
Incremental Mindset	5	.58	.74

The results for Cronbach's reliability did not appear to be acceptable since they were lower than .70, but the results for composite reliability, which is a more dependable criterion and does not suffer from the limitations of Cronbach's alpha's reliability, did turn out to be acceptable as they were higher than .70. In what follows, the results for analyzing the formative measurement model are presented (in Tables 15 to 17), followed by the results for the whole structural model.

Table 15
Collinearity Analysis of the Indicators

Indicators	Outer VIF Values
Ent07	1.76
Ent08	2.45
Ent09	1.61
Ent15	1.04
Inc05	1.51
Inc06	1.56
Inc16	1.70
Inc17	2.74
Inc18	2.15
Commentary	1.07
Conferencing	1.06
Peer Correction	1.07
Prompts	1.07
Self-correction	1.03

Because all of the outer VIF values in Table 15 were less than 5.00, and especially the values for commentary, conferencing, peer correction, prompts, and self-correction were far less than 5.00, it could be argued that there was no collinearity among the indicators in the formative measurement model. The formative measurement model was also evaluated by means of examining the outer weights, as portrayed in Table 16.

Table 16
Outer Weights of the Indicators

Indicators	Outer Weights
Commentary	.62
Conferencing	.58
Peer Correction	-.26
Prompts	.22
Self-correction	.22

The values for the outer weights of the indicators of the formative measurement model are presented above. To examine the significance of these outer weights, the Table 17 should be consulted.

Table 17
Significance of the Outer Weights of the Indicators

Indicators	t Statistic	p Value
Commentary	2.72	.02
Conferencing	2.22	.02
Peer Correction	.90	.36
Prompts	.90	.36
Self-correction	.83	.40

Acceptable values for the significance of the outer weights are t values greater than 1.96 (at .05 level of significance) and p values smaller than .05. Among the indicators lined up in Table 17, commentary and conferencing had acceptable t and p values, but there are good reasons (based on the literature on feedback types and regarding the content validity of feedback preferences) to include all other types of feedback listed in the table above as indicators of the variable 'feedback preferences.' The structural model of the study is examined in the following tables.

Table 18
Collinearity of the Variables in the Structural Model

Variables	Feedback Preferences
Entity Mindset	1.30
Incremental Mindset	1.30

The inner VIF values for entity mindset and incremental mindset equaled 1.30, which is lower than 5.00, indicating that there was no collinearity between the two predictor variables of the model and they could be considered as two separate independent variables. The R^2 in a model refers to the amount of variance in a dependent variable explained by the variances in the independent variables. This statistic is presented in Table 19.

Table 19
R Square and Adjusted R Square

Variable	R Square	R Square Adjusted
Feedback Preferences	.10	.09

It was previously mentioned in Table 8 that R square is weak if it is .19, moderate when it equals .33, and strong at .67. The R square in the table above was .10, which is weak, indicating that entity and incremental mindsets were weak predictors of feedback preferences. The effect size of either of the independent variables on the dependent variable of the study is shown by F^2 in the table below.

Table 20
F Square Results

Variables	Feedback Preferences
Entity Mindset	.067
Incremental Mindset	.005

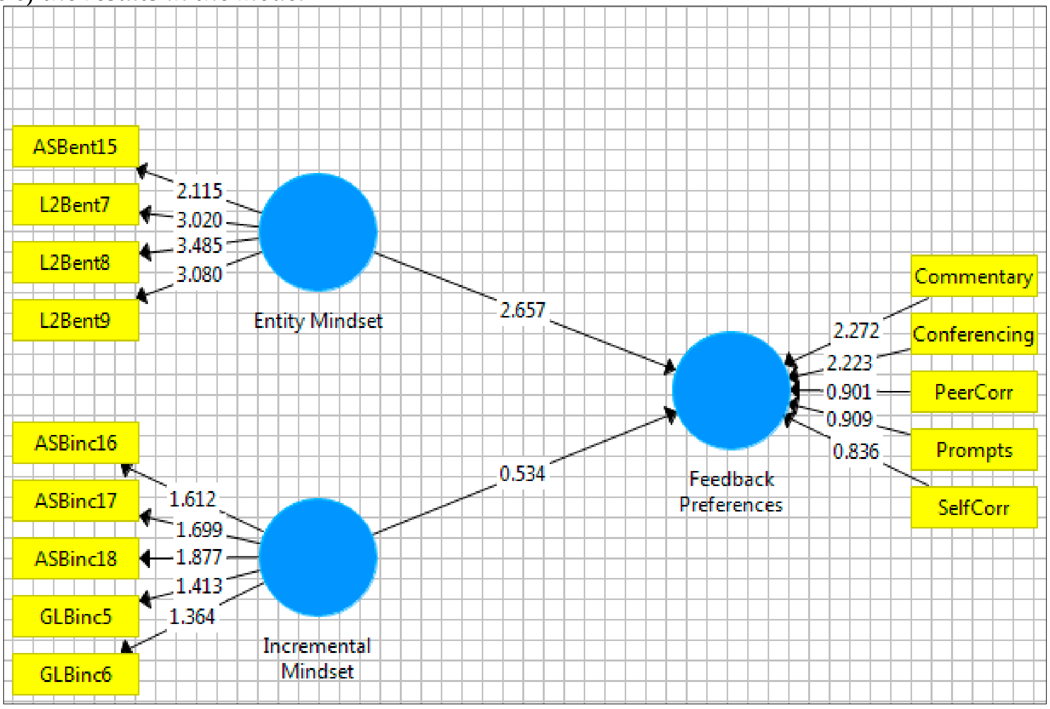
The *F* square value for entity mindset and incremental mindset were .067 and .005, respectively. For *F* square, .02 is weak, .15 is moderate, and .35 is strong. Thus, the effect size of the entity mindset on feedback preferences was rather weak, and for the incremental mindset, it was very weak. Path coefficients and their significance are examined in Table 21.

Table 21
Path Coefficients

Variables	Coefficients	t Statistic	p Value
Entity Mindset	.28	2.65	.008
Incremental Mindset	.07	.53	.593

For path coefficients, .25 is weak, .50 is moderate, and .75 is strong. Table 21 shows that the entity mindset (.28) and the incremental mindset (.07) were both weak predictors of feedback preferences, but the entity mindset could be a better predictor of feedback preferences than the incremental mindset, and significantly predicted feedback preferences ($t > 1.96$, $p < .05$). However, incremental mindsets could not significantly predict feedback preferences ($t < 1.96$, $p > .05$). The *t* values and the significance/non-significance of the effects could be graphically shown through the bootstrapped graphical representation of the model, as illustrated in Figure 3.

Figure 3
Significance of the results in the model



It was mentioned above that *t* values greater than 1.96 imply a significant effect/relationship. Figure 3 shows that the entity mindset ($t = 2.657$) was a significant predictor of feedback preferences, yet the incremental mindset ($t = .534$) failed to do so. Finally, the goodness-of-fit of the model was evaluated through the SRMR criterion, and its value was found to be .13; since SRMR values can range between zero and 1.00, and values lower than .08 imply goodness-of-fit of the model, the present model lacked a satisfactory goodness-of-fit.

Discussion

The results of the first research question of the study unearthed that the surveyed EFL learners agreed with the incremental-mindset beliefs to a significant extent, and disagreed (though not significantly) with the entity-mindset beliefs. One reason why the learners in the present study enjoyed more incremental than entity mindsets could be related to the finding of Lou’s (2014) study, in which the students who were primed for an incremental mindset, irrespective of their perceived L2 ability, set higher learning goals, expressed more mastery-orientated responses in failure situations, and reported stronger intentions to continue learning the target language (Lou & Noels, 2016). In the present study, the participants were selected from among those who volunteered to take part in this study and this sampling procedure might have affected the data obtained from them, as those who volunteer to do something may be more interested in it and thus possess attributes that make them different from a typical L2 learner.

However, the results of the first research question contradict the findings of Lou and Noels (2019b) with regard to the effect of language ideology and acceptance of multilingualism in a society on language mindsets. In their study, they concluded that fixed language mindsets may be less common in more multilingual societies. In an environment where most people have access to learning and speaking multiple languages, people might be less likely to believe that language-learning ability is fixed. Although the EFL learners of this study do not live in a multilingual society and the only language they speak is Persian, they enjoyed incremental rather than entity mindsets. The reason why this result was obtained could be attributed to the fact that the participants in this study were youngsters aged between 14 and 22, who are among the people labeled Generation Z (alternatively called neo-digital natives), and are characterized by lives integrated with the use of technology and social media. They are usually well-informed, updated, and always connected. Given their immense exposure to global culture and the English language through social media, the internet, satellites, and so forth, even though they are considered monolinguals (in Iran), they are in fact bilingual multicultural beings who have developed open minds capable of capturing the realities behind many matters of which older generations were quite incognizant.

The results are in line with those of Claro et al. (2016) and Destin et al. (2019) in that students from high socioeconomic status backgrounds were more likely to hold incremental mindsets. Since the participants in this study were language learners in private language institutes, they had the economic advantage to afford more opportunities to learn and more access to resources for learning and hence were more likely to develop growth mindsets about their own ability.

At the cultural level, mindsets and achievement-goal orientations may vary across socio-cultural and ethnolinguistic environments (e.g., Dweck et al., 1995; Lou & Li, 2017; Ryan & Mercer, 2011). Hence, it is possible that a growth-mindset system is prevalent in countries such as Iran because Iranian culture emphasizes self-criticism and the importance of self-improvement.

The findings of this study also support Lou and Noels' (2019b) belief about the relationship between cognitive capacities and one's mindset. According to Lou and Noels, individual/intrapersonal systems (e.g., cognitive/affective traits and personal experience) can impact students' mindsets. It is possible that learners' current level of cognitive capacities and the learning environment interactively predict one's mindsets. For example, students with low aptitude levels may be more likely to experience repeated failures, especially in a less supportive environment, and thus are more likely to endorse fixed mindsets. Since the EFL learners in this study were at an intermediate level of proficiency, it can be assumed that they had been partially successful in learning an L2 and hence developed a growth mindset. Differently put, they had climbed several rungs of success in foreign language learning and they probably assumed they could successfully climb the remaining rungs up to the top of the ladder.

Through the second research question of the study, it was found that entity mindsets were a significant predictor of feedback preferences, yet the incremental mindsets failed to do so. Among the feedback types, EFL learners' preferences, in a deductive order, were found to be for commentary and conferencing significantly, and then for peer correction, prompts, and self-correction to a non-significant extent. Besides, the whole model (positing the relationship between entity/incremental mindsets and feedback preferences) lacked satisfactory goodness-of-fit, indicating that the two independent variables (i.e., entity and incremental language mindsets) could not, on the whole, account for the EFL learners' feedback preferences in L2 writing.

Previous research has shown that language mindsets have (causal) relationships with a number of language-learning variables. As a case in point, Albalawi (2017) found that mindsets played an important part in language learners' motivation, demotivation, remotivation, and resilience/vulnerability. Similarly, Ryan and Mercer (2011) argued that learners' beliefs about the role of natural talent can affect their motivation and their ability to develop a positive identity as a self-directed language learner. Moreover, it has been shown in the literature on language mindsets that mindsets have relationships with other motivational factors such as attribution (Hong et al. 1999), effort beliefs (Lou & Noel, 2016, 2017), achievement goals (Papi, Rios, Pelt, & Ozdemir, 2019), self-regulatory tendency in the face of adversity (Lou & Noel, 2016; Papi, Bondarenko, Wawire, Jiang, & Zhou, 2019), and competence-based emotional tendencies in the face of adversity (Lou & Noels, 2019), but it does not mean that language mindsets are bound to have roles and relationships with all other language-learning variables (such as feedback preferences) for at least two reasons: first, although mindsets are often described as relatively stable beliefs that vary across individuals (Robins & Pals, 2002), they can vary within an individual over time and across situations; in fact, mindsets are context-specific and may vary according to the context/purpose for which they are used. Second, feedback preferences may be under the influence of other factors that could have a more powerful bearing, compared to language mindsets, on those feedback preferences. Some of these more related influences are mentioned in the following paragraphs.

One reason why the EFL learners in this study did not prefer teacher correction, or corrections and comments as preferred ways of receiving feedback, may lie in a finding that was brought up in a study by Johansson (2018), who found that his research participants perceived feedback that is not commonly used by teachers or was used differently by different teachers difficult to interpret. That is, students struggle with interpreting feedback when there are variations in teachers' practices in applying it.

In a study by Lee (2013), advanced ESL learners reportedly received recasts more than any other type of corrective feedback, but they mostly preferred explicit and immediate corrections in the middle of their interchanges and during the teacher-student interactions. The learners in the current study did receive prompts and were shown to welcome this type of feedback, although not significantly. The reason why they would welcome prompts could have its roots in the fact that prompts are less face-threatening than explicit feedback types and reserved learners would prefer to receive prompts than any other type of feedback that could be a threat to their face, as cultural differences may be at work when students choose how to be instructed and treated.

The fact that different groups of students prefer different types of feedback could be attributed to a number of other factors (such as level of proficiency, learning context, and gender, to name a few). Among these factors one can refer to the results of Witt (2019), who found differences in feedback preferences between L1 graduates and L2 graduates, and also between undergraduate students and graduate students. As another example, Roy (2019) found that multilingual writers expected directive explicit feedback from their instructors. They pointed out that audio feedback was better for global commentary and written feedback was better for local commentary. Their perceptions regarding the effectivities of audio and written feedback on their revision process varied depending on their own self-efficacy.

Conclusion

The present study addressed a niche in the domain of second language-learning mindsets, i.e., the roles of mindsets and their relationships with feedback preferences in writing. The data elicited from 150 intermediate EFL learners unveiled that the participants of this study enjoyed more incremental than entity mindsets in the three areas of general language intelligence beliefs, second language aptitude beliefs, and age sensitivity beliefs about language learning, and their agreements with incremental-mindset items was of statistical significance, while their discord with the entity-mindset items was not statistically significant. It was also found that, on the whole, entity/incremental mindsets could not significantly account for feedback preferences in L2 writing, although entity mindset per se was a significant predictor of feedback preferences. This may have far-fetched implications for language teachers as it is recommended that they seek influential factors (like the ones mentioned above in the Discussion section) other than language mindsets if they intend to work out what makes EFL learners prefer one type of feedback over another. Last but still important, it is recommended that EFL teachers use an amalgamation of different corrective feedback types to appeal to the different needs, styles, and preferences of EFL learners.

Conflict of interests

The authors declare that they have no conflict of interest.

References

Abbaspour, E., Atai, M.R., & Maftoon, P. (2020). The effect of scaffolded written corrective feedback on Iranian EFL learners' writing quality: An activity theory perspective. *International Journal of Foreign Language Teaching and Research, 8*(30), 177-196. https://doi.org/10.32038/ltrq.2020.16.01

Albalawi, F. H. E. (2017). *L2 demotivation among Saudi learners of English: The role of language learning mindsets* [Unpublished doctoral dissertation]. University of Nottingham.

Barcelos, A. M. F. (2003). Researching beliefs about SLA: A critical review. In P. Kalaja & A. M. F. Barcelos (Eds.), *Beliefs about SLA: New research approaches* (pp. 7–33). Dordrecht: Kluwer.

Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System*, 27(4), 459–472.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, *78*(1), 246-263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological bulletin*, *139*(3), 655 -701. https://doi.org/10.1037/a0029531

Carroll, S., Swain, M., & Roberge, Y. (1992). The role of feedback in adult second language acquisition: Error correction and morphological generalizations. *Applied Psycholinguistics, 13*(02), 173-198. https://doi.org/10.1016/j.sbspro.2014.03.637

Chaudron, C. (1988). *Second language classrooms*. Cambridge University Press.

Chiu, C. -Y., Hong, Y. -Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. J*ournal of Personality and Social Psychology, 73*(1), 19–30. https://psycnet.apa.org/doi/10.1037/0022-3514.73.1.19

Cimpian, A., Arce, H. M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*(4), 314-316. https://doi.org/10.1111%2Fj.1467-9280.2007.01896.x

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Mindset equals income as a predictor of achievement. *Proceedings of the National Academy of Sciences of the United States of America, 113*(31), 8664–8668. https://doi.org/10.1073/pnas.1608207113

Cohen, A. D., & Cavalcanti, M. C. (1990). Feedback on compositions: Teacher and student verbal reports. *Second language writing: Research insights for the classroom*, *13*(2), 155-177. https://doi.org/10.1017/S027226310000992X

Cotterall, S. (1999). Key variables in language learning: What do learners believe about them? *System, 27*(4), 493–513. https://doi.org/10.1016/S0346-251X(99)00047-0

Destin, M., Hanselman, P., Buontempo, J., Tipton, E., & Yeager, D. S. (2019). Do student mindsets differ by socioeconomic status and explain disparities in academic achievement in the United States? *AERA Open*, *5*(3). https://doi.org/10.1177%2F2332858419857706

Diab, R. L. (2005). EFL university students' preferences for error correction and teacher feedback on writing. *TESL Reporter, 38*(1), 27-51. http://hdl.handle.net/10725/2796

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Dweck, C. S. (2006). Mindset: The new psychology of success. Random House. https://doi.org/10.4236/jep.2011.24038

Dweck, C. S. (2007). The perils and promises of praise. *Kaleidoscope, Contemporary and Classic Readings in Education*, *12*, 34-39.

Dweck, C. S. (2019). The choice to make a difference. *Perspectives on Psychological Science*, *14*(1), 21-25. https://doi.org/10.1177%2F1745691618804180

Dweck, C. S., Chiu, C. -Y., & Hong, Y. -Y. (1995). Implicit theories and their role in judgements and reactions: A world from two perspectives. *Psychological Inquiry, 6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*(2), 256. http://dx.doi.org/10.1037/h0037130

Dweck, C. S., & Molden, D. C. (2007). Self theories: Their impact on competence motivation and acquisition. In A. J. Elliot & C. S. Dweck (Eds.), Handbook of competence and motivation (pp. 122–140). The Guilford Press.

Ebrahimzade, M. & Khodareza, M.R. (2016). The effect of hedged-form feedback vs. uncoded feedback on grammatical accuracy of Iranian intermediate EFL learners. *International Journal of Foreign Language Teaching and Research, 4*(15), 25-33.

Ellis, R. (2006). Researching the effects of form-focused instruction on L2 acquisition. *AILA Review, 19*, 18–41.

https://doi.org/10.1075/aila.19.04ell

Ellis, R. (2012). *Language teaching research and language pedagogy*. Wiley Blackwell.

Ellis, R., S. Loewen & R. Erlam (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*, 339–368.

Farrokhi, F., Zohrabi, M., & Chehr Azad, M.H. (2017). The effect of corrective feedback on Iranian EFL learners' speaking accuracy and breakdown fluency. *Journal of Language Horizons, 1*(2), 107-129.

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes: a response to Truscott (1996). *Journal of Second Language Writing*, *8*(1), 1-11. https://doi.org/10.1016/S1060-3743(99)80110-6

Ferris, D. R. (2004). The «grammar correction» debate in L2 writing: where are we, and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing*, *13*, 49-62. https://doi.org/10.1016/j.jslw.2004.04.005

Gonzales, M.L.L., de Tejeda, M.L., Krous, J.T., & Vasquez, D. (2018). EFL teachers' corrective feedback and students' revision in a Peruvian university: A descriptive study. *International Journal of Foreign Language Teaching and Research,* 6(23), 11-22.

Hair, J.F., Hult, G.T.M., Ringle, C.M., & Sarstedt, M. (2017). *A premier on partial least squares structural equation modeling* (2nd ed.). Sage.

Halimi, S. S. (2008). Indonesian teachers' and students' preferences for error correction. *Wacana, 10*(1), 50-71. https://doi.org/10.5539/efl.v4n3p128

Hedgcock, J., & Lefkowitz, N. (1994). Feedback on feedback: Assessing learner receptivity to teacher response in L2 composition. *Journal of Second Language Writing*, *3*(2), 141-163. https://doi.org/10.1016/1060-3743(94)90012-4

Hedgcock, J., & Lefkowitz, N. (1996). Some input on input: Two analyses of student response to expert feedback in L2 writing. *The Modern Language Journal*, *80*(3), 287-308. https://doi.org/10.1111/j.1540-4781.1996.tb01612.x

Hong, Y. Y., Chiu, C. Y., Dweck, C. S., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology, 77*(3), 588-599. https://psycnet.apa.org/doi/10.1037/0022-3514.77.3.588

Horwitz, E. K. (1998). The beliefs about language learning of beginning foreign language students. T*he Modem Language Journal, 72*(3), 283–294. https://doi.org/10.1111/j.1540-4781.1988.tb04190.x

Johansson, M. (2018). *Formative assessment: Students' attitudes and preferences in Swedish upper secondary school* [Unpublished master's thesis]. Orebro University.

Karami, M., & Heidari Darani, L. (2018). Impact of prompts as corrective feedback strategy on teaching /θ/ and /ð/ among Iranian intermediate EFL learners. *International Journal of Foreign Language Teaching and Research,* 6(24), 67-82.

Katayama, A. (2007). Japanese EFL students' preferences toward correction of classroom oral errors. *Asian EFL Journal, 9*, 289–305.

Kelly, G. A. (1955). *The Psychology of personal constructs*. Norton. http://dx.doi.org/10.1037/11175-008

Lee, E. J. E. (2013). Corrective feedback preferences and learner repair among advanced ESL students. *System, 41,* 217-230. https://doi.org/10.1016/J.SYSTEM.2013.01.022

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing, 13*(4), 285-312. https://doi.org/10.1016/j.jslw.2004.08.001

Lee, I. (2005). Error correction in the L2 writing classroom: What do students think? *TESL Canada journal*, 1-16. https://doi.org/10.18806/tesl.v22i2.84

Leki, I. (1991). The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals, 24*(3), 203-218. https://doi.org/10.1111/j.1944-9720.1991.tb00464.x

Lou, M. T. (2014). *Changing language learning mindsets: The role of implicit theories of L2 intelligence for goal orientations and responses to failure* [Unpublished master's thesis]. University of Alberts.

Lou, N. M., & Li, L. M. W. (2017). Interpersonal relationship mindsets and rejection sensitivity across cultures: The role of relational mobility. *Personality and Individual Differences*, *108*, 200-206. https://psycnet.apa.org/doi/10.1016/j.paid.2016.12.004

Lou, N. M., & Noels, K. A. (2016). Changing language mindsets: Implications for goal orientations and responses to failure in and outside the second language classroom. *Contemporary Educational Psychology*, *46*, 22-33. https://psycnet.apa.org/doi/10.1016/j.cedpsych.2016.03.004

Lou, N. M., & Noels, K. A. (2017). Measuring language mindsets and modeling their relations with goal orientations and emotional and behavioral responses in failure situations. *The Modern Language Journal*, *101*(1), 214-243. https://doi.org/10.1016/j.paid.2016.12.004

Lou, N. M., & Noels, K. A. (2019a). Sensitivity to language-based rejection in intercultural communication: The role of language mindsets and implications for migrants' cross-cultural adaptation. *Applied Linguistics*, *40*(3), 478-505. https://doi.org/10.1093/APPLIN%2FAMX047

Lou, N. M., & Noels, K. (2019b). Does my teacher believe I can improve? Meta-perceptions of potential mediate the effect of teachers' feedback on ESL students' mindsets and need satisfaction. Paper presented at the *American educational research association* (AERA) conference, Toronto, ON.

Lou, N. M., & Noels, K. (2019c). Breaking the vicious cycle of language barriers: Growth language-mindsets improve migrant university students' communication experiences [Unpublished doctoral dissertation]. University of Alberta. https://doi.org/10.7939/r3-tnzx-4736

Lou, N. M., & Noels, K. A. (2020). Breaking the vicious cycle of language anxiety: Growth language mindsets improve lower-competence ESL students' intercultural interactions. *Contemporary Educational Psychology*, *61*, 101847. https://doi.org/10.1016/j.cedpsych.2020.101847

Lyster, R. & L. Ranta (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*, 37–66. https://doi.org/10.1017/S0272263197001034

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, *46*(01), 1-40. https://doi.org/10.1017/S0261444812000365

McDonough, K. & Mackey, A. (2000). Communicative tasks, conversational interaction, and linguistic form: An empirical study of Thai. *Foreign Language Annals, 33*, 82–91. https://doi.org/10.1111/j.1944-9720.2000.tb00893.x

Mercer, S. (2012). Dispelling the myth of the natural-born linguist. *ELT Journal*, *66*(1), 22-29. https://doi.org/10.1093/elt/ccr022

Mercer, S., & Ryan, S. (2010). A mindset for EFL: Learners' beliefs about the role of natural talent. *ELT Journal*, *64*(4), 436-444. https://doi.org/10.1093/elt/ccp083

Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Sage publications.

Mir, F., & Ghornavi, K. (2017). Effect of corrective feedback on the acquisition of English prepositions of movement and place in third-grade high school EFL learners' grammar performance. *International Journal of Foreign Language Teaching and Research, 5*(20), 125-140.

Naeimi, A., Saeidi, M., & Behnam, B. (2017). Effect of oral corrective feedback on Iranian EFL learners' phonological uptake and retention. *International Journal of Foreign Language Teaching and Research,* 5(17), 57-75.

Noori Khaneghah, R. (2016). The efficiency of different types of corrective feedback on vocabulary development of Iranian EFL learners: A comparative study. *International Journal of Humanities and Cultural Studies,* Special January Issue, *1*(1), 2534-2548.

Papi, M., Bondarenko, A. V., Wawire, B., Jiang, C., & Zhou, S. (2019). Feedback-seeking behavior in second language writing: Motivational mechanisms. *Reading and Writing*, *33*, 485-505. https://doi.org/10.1007/s11145-019-09971-6

Papi, M., Rios, A., Pelt, H., & Ozdemir, E. (2019). Feedback-seeking behavior in language learning: Basic components and motivational antecedents. *The Modern Language Journal, 103*(1), 205-226. https://doi.org/10.1111/modl.12538

Ramadhani, S.A.F. (2019). Investigating corrective feedback in speaking practice: Students' preferences. *Advances in Social Science, Education, and Humanities Research, 424,* 187-192.

Reid, J.M. (1997). *Understanding learning styles in the second language classroom*. Prentice Hall.

Rennie, C. (2000). *Error feedback in ESL writing classes: What do students really want* [Unpublished Master's thesis]. California State University.

Rissanen, I., Kuusisto, E., Tuominen, M., & Tirri, K. (2019). In search of a growth mindset pedagogy: A case study of one teacher's classroom practices in a Finnish elementary school. *Teaching and Teacher Education*, *77*, 204-213. https://doi.org/10.1016/j.tate.2018.10.002

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity, 1*(4), 313-336. https://doi.org/10.1080/15298860290106805

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46–73. https://doi.org/10.1017/S0267190505000036

Roy, S. (2019). *Exploring multilingual writers' preference between audio and written feedback, and the impact of feedback format on their revision process in a U.S. composition class* [Unpublished doctoral dissertation]. Indiana University of Pennsylvania.

Ryan, S., & Mercer, S. (2011). Natural talent, natural acquisition and abroad: Learner attributions of agency in language learning. In G. Murray, X. Gao & T. Lamb (Eds.), *Identity, motivation and autonomy in language learning* (pp. 160-176). Multilingual Matters.
Saito, H. (1994). Teachers' practices and students' preferences for feedback on second language writing: A case study of adult ESL learners. *TESL Canada Journal, 11*(2), 46-70. https://doi.org/10.18806/tesl.v11i2.633
Schulz, R. A. (1996). Focus on form in the foreign language classroom: Students' and teachers' views on error correction and the role of grammar. *Foreign Language Annals, 29,* 343-364. https://doi.org/10.1111/j.1944-9720.1996.tb01247.x
Sippel, L. (2019). The impact of peer corrective feedback on vocabulary development. *Foreign Language Annals, 52*(3), 595-611. https://doi.org/10.1111/flan.12416
Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language aptitude testing. In P. Robinson (Ed.), *Individual differences and instructed language learning*. John Benjamins Publishing.
Tesnim, O. (2019). Oral corrective feedback and its impacts on learners' speaking skills: Tunisian ESL students as a case study. *International Journal of Language and Linguistics, 7*(3), 138-149. https://doi.org/10.11648/j.ijll.20190703.15
Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, *46*(2), 327-369. https://doi.org/10.1111/j.1467-1770.1996.tb01238.x
Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, *8*(2), 111-122. https://doi.org/10.1016/S1060-3743(99)80124-6
Wang, P. (2010). Dealing with English majors' written errors in Chinese universities. *Journal of Language Teaching and Research, 1*(3), 194-205. https://doi.org/10.4304/jltr.1.3.194-205
White, C. (2008). Beliefs and good language learners. In C. Griffiths (Ed.), Lessons from good language learners (pp. 121–130). Cambridge University Press.
Witt, J. (2019). *Disciplinary writing feedback practices: Comparing preference, expectations, and practices.* [Unpublished doctoral dissertation]. University of Arizona.
Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302-314. https://doi.org/10.1080/00461520.2012.722805
Zhai, K., & Gao, X. (2018). Effects of corrective feedback on EFL speaking task complexity in China's university classroom. *Cogent Education, 5*(1), 1-13. https://doi.org/10.1080/2331186X.2018.1485472
Zohi Rad, M., & Ghafournia, N. (2016). The effect of focused corrective feedback and attitude on grammatical accuracy: A study of Iranian EFL learners. *International Journal of Foreign Language Teaching and Research, 4*(15), 75-85.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Murtisari, E. T., Puspitasari, D., & Setiamunadi, A. A. (2020). Peer Teaching to Assist Tertiary EFL Grammar Learning: Indonesian Tutors' Perceptions of Challenges and Strategies. *Journal of Language and Education, 6*(4), 204-215. https://doi.org/10.17323/jle.2020.10763

Peer Teaching to Assist Tertiary EFL Grammar Learning: Indonesian Tutors' Perceptions of Challenges and Strategies

Elisabet Titik Murtisari, Dewi Puspitasari, Antonina A. Setiamunadi
Universitas Kristen Satya Wacana

Correspondence concerning this article should be addressed to Elisabet Titik Murtisari, English Language Education Program, Faculty of Language and Arts, Universitas Kristen Satya Wacana (UKSW), Jl. Kartini No.15-17, Salatiga 50714, Indonesia. Email: etmurtisari@gmail.com

As an essential language component, grammar plays a crucial role in communication. However, with the need to master various L2 forms within several years and an L1 that has a very different grammatical system from English, many tertiary EFL students find grammar learning challenging. To solve this issue, peer teaching/tutoring can be a very effective way to assist students in improving their grammar. While isolated grammar teaching has its downsides, it is superior in clarifying complex concepts and promoting accuracy. It may help increase students' grammar competence when combined with other methods as an eclectic approach used in a communicative language program. As grammar tutors play a crucial role in helping their peers, this study examines the challenges they perceive in peer tutoring at the tertiary EFL education level and their strategies for overcoming difficulties. Using interviews to collect data from ten EFL grammar tutors, this qualitative study revealed some issues the tutors faced. Besides identifying problems such as motivating students, preparing the materials, and dealing with less/more proficient students, this research demonstrated that peer teaching might lead to 'cognitive dependence' among the lower-level learners. To deal with the various issues, the peer tutors applied practical strategies they had developed mainly from intuition and experience.

Keywords: peer teaching/tutoring, grammar, tertiary EFL students, challenges, strategies

Introduction

An essential element in meaning-making, grammar plays a crucial role in effective communication. However, with the need to master various L2 forms within several years and an L1 that often has a very different grammatical system from English, many tertiary EFL students find grammar learning challenging. To overcome this issue, peer teaching/tutoring may provide an effective support system to help such students (Knight, 2013; Mynard & Almarzouqi, 2006; Mulatsih, 2018; Watcharapunyawong, 2018; Won, Change, & Kang, 2017), as it is difficult to simply rely on teachers' assistance from regular class meetings. Allowing a more friendly environment, peer teaching may help students improve their understanding of the subject matter. Despite this, research on peer teaching that specifically addresses EFL grammar is still scarce to date. Instead, grammar has mainly been examined as part of tutoring in language skills, especially writing (e.g., Kim, 2018; Snyder, Nielson, & Kurzer, 2016; Winder, Kathpalia, & Koo, 2016). As a lower-order issue in writing, grammar does not usually receive adequate attention in such peer tutoring. As the language element is a crucial aid for communication, this highlights the need for peer tutoring that specifically addresses grammar for students who need such support. Due to its ability to illuminate complex concepts and foster formal accuracy, isolated grammar teaching still has its place for cognitively mature EFL learners in communicative language programs (Murtisari, Hastuti, & Arsari, 2019; Murtisari, Salvadora, & Hastuti, 2020). With all these issues in mind and the central role of the tutors on the success of a peer tutoring program, this study aims to investigate the challenges they face in assisting their peers or near peers and their strategies to overcome difficulties. This research is significant for assisting tertiary language programs in their endeavor to improve their students' grammar mastery, which is an integral part of functional language competence.

Literature Review

Peer teaching/tutoring

Peer teaching or peer tutoring, which has been implemented for decades in higher education (Williams & Reddy, 2016), is a form of learning where a student teaches another student of a similar status. According to Gaies (1985), peer teaching may take the form of same-level or cross-level partnerships. In the same-level type, both the tutors and tutees come from the same class and level of proficiency. In cross-level tutoring, on the other hand, the tutors are more competent and may be the same age as their tutees (same-age tutoring) or older (intergrade tutoring). Cross-level peer tutoring is commonly applied in tertiary institutions to help students to pass a particular course or reduce the drop-out rate.

Peer teaching has been reported to benefit both tutors and tutees. With peer tutoring, tutees may understand the subject matter better because there are more opportunities to respond, more time-on-task and practice with the academic content, and more accessible feedback (Bowman-Perrott, deMarín, Mahadevan, & Etchells, 2016; Mynard & Almarzouqi, 2006). It may also enhance their motivation and confidence, enable them to gain access to a learning "role model", and permit them to have a safer learning environment (Menezes & Premnath, 2016, p.160). Bohórquez, Rodríguez, and González's (2019) research also showed that peer tutoring may promote learners' autonomy, which refers to learners' (and tutors') capacity to "play an active and responsible role in designing learning agendas rooted in learners' needs" (para. 9). For peer tutors, on the other hand, peer teaching may serve as an avenue for both academic and non-academic self-improvement. "[S]hift[ing] from being students as recipients to being productive teachers," they will need to have a deeper comprehension of the subject matter to be able to teach effectively (Stigmar, 2016, p.125). This will allow them to grow cognitively as they also become more critical of what they are learning. Non-academically, peer teaching has also been shown to improve tutors' skills in communication, leadership, group facilitation (Menezes & Premnath, 2016), multi-task management (Adams, 2011), and teaching (Naeger, Conrad, Nguyen, Kohi, & Webb, 2013).

Despite the benefits, there are common problems in peer tutoring. One issue relates to the tutors' ability in teaching. There are concerns that tutors may not be able to deliver quality content because they are still inexperienced compared to faculty members (Menezes & Premnath, 2016). A lack of preparation to equip tutors has been identified as one issue that may complicate this problem. According to Irvine, Williams, and McKenna (2018), in referring to several studies, this includes a lack of time to prepare for teaching sessions and a lack of preparation to help them adjust to gaps of learning styles among themselves, the tutees, and the faculty's course instructors/lecturers. Another preparation issue the researchers mentioned is difficulties in preparing teaching materials, which was also indicated in Stanley's study (1998). Other challenges tutors may face in peer teaching are a lack of confidence about their knowledge and playing the role as a teacher (Christiansen, Bjørk, Havnes, & Hessevaagbakke, 2011), anxiety before teaching (McKenna & French, 2011), and dealing with critical students who may weaken the tutors' power as a facilitator (Irvine et al., 2018). Tutors, however, have also been reported to be concerned that they may sound arrogant when sharing their knowledge (Christiansen et al., 2011). On top of this, cultural issues may also affect peer tutoring. Takeuchi (2015), for instance, reported that the hierarchical relationship between Japanese tutors and tutees may prevent collaborative learning. As tutees look up to tutors, they tend to be reluctant to ask questions even though they do not understand. The relationship also tends to become formal with tutors assuming more power.

Apart from the above problems, research has also identified issues related to tutees. One study showed that peer teaching may cause dependence among tutees. As many as 60% of Mynard and Almarzouqi's (2006) student-teacher participants reported that their tutees developed a "high degree of dependence" on them (p. 16). According to the tutors, this was indicated by the tutees' requests to do their homework and extra sessions before English tests. However, the research on peer tutoring for English and grammar for female foundational year students in the UAE did not clearly describe the issue of 'dependence'. While asking for more tutoring to get prepared for tests may be a problem on its own, it does not necessarily comprise dependence. Irrespective of this, Menezes and Premnath (2016) also suggested that tutees tend to use peer tutoring to help them to perform well on exams rather than equip them with the necessary knowledge and skills that can help them in their careers later. It was suggested that exam-driven and teacher-centered learning have contributed to such behavior. This appears to show that students tend to be pragmatic when it comes to their studies.

Regardless of its potential challenges, peer education may be linked to two pedagogical paradigms: social constructivism and cognitive congruence. Based on the idea that learning occurs in a social context, social constructivism views knowledge as being constructed through interactions with other people. Through peer teaching, students collaborate in a context where "a less able peer is able to enter a new area of potential development through problem solving with someone more able" (Asghar, 2010, p.406). This process involves "active questioning, explaining, monitoring, and regulating in the learning process" (Stigmar, 2016, p.131) based on equal power, mutual respect, and openness to others' ideas (O'Sullivan & Cleary, 2014). Furthermore, in light of the cognitive congruence theory, while the tutors are more competent, the knowledge gap between them and their peers is much smaller compared to faculty educators. With this similarity of experience, peer tutors may explain concepts in ways that are more comprehensible to tutees, which could enhance knowledge construction during peer teaching. This may explain why the method is considered useful for promoting the learning of complex concepts (Karpicke & Blunt, 2011).

However, peer tutors' teaching styles also shape the nature of learning in peer tutoring. According to Berghmans, Michiels, Salmon, Dochy, and Struyven (2014), tutors may be directive or facilitative in their teaching. With a directive style, a tutor gives explicit directions and explanations to tutees' queries, focusing on the provision of essential content and information. Some strategies that fall within this style are informing, giving direct answers and feedback, explaining, and summarizing (Berghmans et al., 2014). Within a facilitative environment, on the other hand, a tutee is more actively engaged in the learning process by initiating more questioning and challenging their own knowledge construction (Berghmans et al., 2014). A facilitative tutor may apply methods such as hinting, questioning, prompting, guiding, filling-in-the-blank, and stimulating self-feedback. Each type of strategy has its own strengths and downsides, but in the study by Berghmans et al., the tutees discovered that the directive style offered more clarity while the facilitative one allowed more in-depth learning and hence more knowledge gains. However, their categories of strategies/methods do not seem to be mutually exclusive, as tutors may apply different combinations of strategies. Nevertheless, the research does not show if the use of teaching methods may be part of the tutors' strategies for dealing with specific types of tutees, which will be crucial to discover if this is the case.

Isolated Grammar Teaching

Isolated grammar teaching may be defined as "separate/isolated instruction especially devoted to focus[ing] on discrete grammatical items by using primarily explicit techniques" (Murtisari et al., 2020, p. 19). In tertiary EFL language programs, this traditional approach is typically carried out in independent courses based on a structure-based syllabus, which seems to be a common tradition in Indonesia. Allowing a closer focus on grammatical forms, isolated grammar teaching may effectively enhance students' understanding of L2 forms, especially of complex concepts, and promote accuracy (Murtisari et al., 2019). According to Spada and Lightbown (2008), the approach may assist students with learning grammatical items that are "rare or absent in the language they are exposed to via CLT [Communicative Language Teaching] or CBI [Content Based Instruction] classes". It is also more effective to teach students the explicit knowledge of grammar (Umeda, Snape, Yusa, & Wiltshier, 2017) that is crucial for tertiary students intending to pursue EFL teaching career paths.

Despite the advantages, isolated grammar teaching has been criticized for not being able to teach students to use grammar to communicate due to its frequently limited communicative content (Larsen-Freeman, 2015; Long, 2016). However, this weakness can be mitigated by giving students more exposure to authentic L2 use and co-implementing an integrative grammar teaching approach (Focus on Form). In this way, the explicit knowledge obtained from isolated teaching can be recycled and reinforced for communicative use. As DeKeyser (1998, 2008, 2015) pointed out, explicit knowledge can become implicit or automatized through practice during which students may internalize the rules. However, teaching grammar integrally does not guarantee automaticity either. While it introduces forms in communicative contexts, it is ultimately students' practice and exposure to the L2 that will help them develop implicit language skills. Sheen (2005) discovered that his experimental study's participants taught by the Focus on Form approach continued to produce inaccurate forms. Therefore, isolated and integrative grammar teaching deliveries are complementary rather than "oppositional" (Ellis, 2015, p.10) as each has its own merits and pitfalls. With its particular strengths, isolated grammar teaching remains a highly appropriate approach for EFL language programs, especially those expecting students to learn various grammatical forms within a relatively short time as well as the explicit knowledge of the L2.

Previous Studies

Research on peer tutoring that focuses on grammar at the tertiary level is still limited. Although the language element has frequently been addressed in studies on peer teaching, it has largely been examined only as part of the tutoring of language skills, especially writing (e.g., Winder, Kathpalia, & Koo, 2016; Kim, 2018; Snyder, Nielson, & Kurzer, 2016). In addition, while it still tends to receive substantial attention in such tutoring (Snyder et al., 2016), it has usually been addressed only as part of proofreading or editing (Bell & Elledge, 2008; La Clare & Franz, 2013). The few studies that have investigated peer teaching focusing on grammar are mainly concerned with the benefits of such a program for tutors and tutees. Hidayah (2014), for instance, found that the approach may improve high school students' understanding of the simple past tense in his quasi-experimental study on peer tutoring's effectiveness for grammar learning. By examining EFL students in a Korean tertiary context, Won, Change, and Kang (2017) also discovered that peer tutoring may improve less proficient learners' grammar. Both tutors and tutees were reported to benefit from the program. Corroborating this study, Mulatsih's research (2018) showed that over 84% of the participants believed that peer teaching helped them to study grammar more intensively and enhanced their understanding of L2 forms and skills when using them. Two advantages tutees pointed out were that the tutoring enabled them to deal with the grammatical problems of individual learners and access explanations they could understand more easily due to the more open communication channels with their tutors. In reporting about the online peer tutoring for grammar, Watcharapunyawong (2018) also demonstrated very high satisfaction with the use of the method among EFL learners. The study also revealed an enhanced sense of responsibility among the tutors, as they had to prepare well to be able to assist their peers on a specific grammatical item. However, despite all the positive results on peer grammar teaching, none discussed peer tutors' challenges and their strategies to deal with the problems. Therefore, as tutors play an important role in the success of a peer teaching program, it is crucial to examine the issues in research.

Methodology

This preliminary qualitative descriptive research was conducted in a respected undergraduate EFL language education (teacher training) program in Central Java, Indonesia. English was used as the medium for instruction for around 90% of its courses, but students used Indonesian or their local language outside classes. The peer tutoring was carried out as part of two grammar courses to help students improve their grammar competence to the faculty's standards. The grammar courses were prerequisites for academic writing courses so they were high-stake subjects. In order to help students succeed, they had been divided into classes according to their levels of English competence to allow different paces of learning. In addition, one introductory session of the Basic Grammar course was devoted to addressing the importance of grammar to motivate students to learn the language aspect. Teachers were also encouraged to keep motivating their students to learn it. Furthermore, based on a structural syllabus, the grammar courses implemented isolated grammar teaching, which relies heavily on explicit instruction. Here, it is worth noting that grammar was also integrated into skill courses in the EFL program. The isolated and integrative methods were considered to be complementary in order to develop students' grammar competence.

In line with the two grammar courses' objectives, the tutoring program aimed to help the English major students understand the meaning of specific grammatical items, enhance their ability to use them accurately, and improve their fluency in using grammar in communication through more applicative practice. The subjects taught basic linguistic concepts and various grammatical features such as articles, verbs and adjective phrases, imperatives, the passive voice, tenses, clauses, and reduced clauses. Furthermore, the assessment was conducted through written tests focusing on the accuracy and appropriateness of the use of grammatical items both at the sentence and discourse levels. The tutoring was deemed necessary as around 60% of the language program's new students tended to have low grammar competence, which was partly reflected by their low PBT TOEFL scores. Grouping students into different classes based on their English proficiency was considered insufficient as individual students had different problems in their grammar learning.

With the country's EFL setting and limited hours for the English subject at school, grammar had rarely received adequate attention at the previous levels of education. As a result, many students were not able to communicate in English with proper grammar by the time they entered the university. Many were also ignorant of basic English grammatical rules, such as the use of the auxiliary verbs "be", "do", and "have". As grammar competence and explicit grammar knowledge are essential for the program's graduates to be qualified English teachers in the country, students in need were provided with a generous number of hours for tutoring. The tutoring was not compulsory, but each student had access to one tutor that had been assigned to help him/her and was encouraged to have regular meetings. It was conducted on an appointment basis and usually involved two to four students at the most. At the lecturer's request, one-to-one peer teaching was also conducted for students who were less proficient.

Situated in the above context and consideration that research on peer teachers's challenges and the strategies to deal with them is scarce, this study sought to answer the following research questions:

1. What perceived challenges do peer tutors face in peer teaching to assist students of isolated grammar courses?
2. What strategies do they self-reportedly apply in order to deal with the challenges?

It is worth noting that this preliminary research was focused on describing the problems from the viewpoints of the tutors and, therefore, was only based on their narrative reflections from interviews.

Participants

Ten tutors were involved in this study. They were selected out of a total of 12 tutors employed during the study to represent gender and seniority (year of study). To become a tutor, they had to obtain an 'A' in all the grammar courses offered by the EFL program and have a GPA of at least 3.5 (out of 4). They also had to possess relevant soft skills, such as having a strong sense of responsibility, sociability, and ability to communicate. All the tutors received an initial briefing on their administrative duties and how to assist their peers, such as how to be accommodating and encouraging. No specific strategies were given on how to teach their peers.

Table 1
Participants' Background

Tutor	Gender	Age	Length of experience of being a tutor (trimester)
A	F	21	1
B	F	20	2
C	F	21	5
D	F	22	6
E	M	19	2
F	F	21	3
G	M	19	2
H	F	20	2
I	F	21	4
J	F	20	2

Data Collection Instrument and Procedures

In order to obtain more in-depth information from the participants, the data was collected using semi-structured interviews. The main interviews were all administered by one member of the research team. The audio-recorded interviews took around 20-30 minutes for each participant and were conducted in the Indonesian language to help the participants to express themselves more comfortably (see Tsang, 1998). To clarify or confirm the responses obtained in the interviews, another team member contacted several tutors to get more detailed information.

Data Analysis

All the results of the recorded interviews were transcribed and analyzed using a thematic analysis, which is used for "identifying, analyzing, and reporting patterns (themes) within data" (Braun & Clarke, 2006, p. 79). Developing themes through close and frequent reading, this method may generate "a thematic map" (p. 87) and provide a detailed account of the data. Following Braun and Clarke's method, the analysis of the data required several steps, which were: repeated careful reading to identify meaningful data features relevant to the research issues, coding the data, and then collating it to establish the themes. After the classification was conducted by one research team member, it was crosschecked by a second member to ensure the accuracy of the themes. Different perspectives on the data interpretation were resolved through mutual reviews.

Results and Discussion

The data analysis showed some perceived challenges during the peer teaching, which included motivation, getting teaching materials, instructional/knowledge construction issues, and culture-related issues. More concerns were raised on instructional matters. As non-professional teachers, the peer tutors were often faced with teaching problems arising from the various characteristics of students. The peer teachers, however, were shown to be able to find practical strategies to overcome the difficulties mainly by intuition and experience. The results of this study are discussed further below.

Challenges of Grammar Peer Teaching in EFL Learning

Low Motivation to Learn Grammar

The most common problem tutors mentioned in the peer teaching was tutees' "low motivation" to learn grammar, which had adversely affected their attitudes towards the peer tutoring (Tutors A, B, D, E, G, H, I, and J). According to the tutors, this was primarily shown by many tutees' reluctance to review materials and their relatively limited appointments for the non-compulsory peer tutoring program. While the tutoring program was overall relatively successful and improved many students' grades with approximately a doubled number of students passing the courses, tutors reported that the tutees did not have regular sessions as encouraged. They instead tended just to request tutoring to get prepared for a grammar test, which echoes findings in Mynard and Almarzouqi's (2006) and Menezes and Premnath's (2016) research. Tutor G said that concentrated sessions before tests were more challenging as he had to review several topics at once, which could make the meetings less effective. The assessment-oriented behavior was observed by Tutor D, the most senior tutor in the program:

> Students sometimes do not have enough motivation because they see grammar tutoring as being useful only to help them pass the course, whereas it is actually done to help them improve [their grammar] proficiency.

There were several causes that may have contributed to the tutees' reluctance to have more regular sessions. Firstly, many seemed to have low awareness of the essential role of grammar in communication, so they did not prioritize grammar learning in their study. Secondly, as many, if not most, students were not used to studying grammar analytically in school, they may have found dealing with linguistic concepts and focusing on details for accuracy daunting or boring. Thirdly, many also did not seem to understand that they needed regular study and practice to create automaticity in grammar production. With all these possible factors, many tutees seemed to have been driven by a short-term learning goal just to pass the grammar course they were taking.

Furthermore, less engagement was reported to occur among course repeaters, which could be partly expected as low motivation is common among low achievers (Linnenbrink & Pintrich, 2002). Commenting on such tutees, Tutor I said:

> It is not compulsory for students to join grammar tutoring, but the level of motivation this year is much lower, very different from the previous year. [...] We have many repeater students this year. We must motivate them to join tutoring, but I don't feel comfortable pushing them. We merely suggest that if there are any class materials they haven't understood, there is [tutoring] support available for them.

As an essential determinant of students' learning engagement (Saeed & Zyngier, 2012) and hence academic success (Wigfield & Cambria, 2010; Steinmayr, Weidinger, Schwinger, & Spinath, 2019), motivation is a critical issue to address. To deal with this, tutors applied several strategies and took up the role as motivators, as encouraged by the tutoring program coordinator. Working at the intrinsic and extrinsic levels, they tried to promote awareness of the importance of grammar for their language learning, study, and future careers among their tutees. To increase intrinsic motivation, tutors created a conducive learning atmosphere to make tutees feel comfortable during the learning sessions and develop positive associations with the grammar subject. This was done, for instance, by being friendly and very patient, and pacing the learning sessions according to the tutees' abilities.

On the other hand, in order to develop extrinsic motivation, the tutors, for example, explained to tutees how grammar may assist them in communication, increase their grades, and improve their opportunities for better careers, which the tutors were encouraged to do by the peer teaching program coordinator. Tutors also reported giving extrinsic motivations to their tutees in the form of a lot of encouragement and praise for their efforts during peer teaching sessions. This strategy may create more rewarding learning experiences which may motivate the tutees to learn more of the subject (Opdecam, & Maussen, 2017; Hancock, 2010). Going beyond this, one tutor (Tutor I) motivated her peer students by giving small presents if they obtained a good grade. While extrinsic motivation is not as desirable as intrinsic, they both enhance students' engagement in learning (Everaert et al., 2017; Saeed & Zyngier, 2012). Thus, the tutors' strategies to deal with both were appropriate. Working at a more personalized level with learners, peer tutors may be able to address motivational issues more deeply through 'mate to mate' talks. This demonstrates how the task of developing students' motivation can be carried out through partnership between teachers and peer tutors, as the work usually takes much effort, especially among low achievers. However, to reap this benefit, it is essential to select tutors with excellent people skills so they can act as good mentors.

Teaching Materials

Obtaining teaching materials was usually not an issue for student-teachers, but several tutors (Tutors G, C, and D) mentioned that finding proper materials was not easy, as the tutees also needed a lot of practice. Although there were faculty resources and the internet also abounded with grammar exercises, materials that could be readily used were limited. In the context of teaching individual tenses before mixing them in exercises, Tutor D pointed out that:

> "[…] the materials on the internet are not the same as the faculty's [courses']. For example, the materials on tenses on the internet are mixed tenses, while those from the faculty's courses are taught one by one in sequence. So, we had to adjust the materials from the internet".

Tutor C added that she needed to find additional materials that were appropriate to her tutees' capacity to understand. She added that:

> The challenges were more on finding exercises that could help the students understand [the course materials] easily, and I had to think very hard to find ways to help students with weak memory to be able to really absorb the materials and always remember them, and I am still struggling with this.

This problem demonstrated the need for the faculty's support to provide exercises that are directly aligned to the course content for students with different levels of proficiency. This will require a significant effort such as hiring professionals or qualified senior students to develop the materials. However, since these types of materials are central for creating a link with the grammar course the peer tutoring supports, the faculty's assistance should be made available.

Instructional/Knowledge Construction Issues

There were several instructional issues identified in this study, which reflected the different characteristics of the tutees they had to assist. Tutees with low language proficiencies were shown to have affected the tutors most during the peer teaching. Here, students' competence was shown to be a significant factor determining the tutors' use of directive or facilitative strategies.

Dealing with Students with Low L2 Proficiency. Seven tutors (Tutors A, C, G, E, F, I, and J) mentioned that they had difficulty explaining the materials to students with low English competence. As around 30% of the grammar courses' students fell far below the language program's standard of L2 proficiency, they had to work harder to explain things. This shows that while it is assumed there is more cognitive congruence in peer teaching, tutors may still face a significant knowledge gap. Commenting on this, Tutor F said:

> Most of the repeater students lacked basic knowledge of English grammar [...]. Rather than teaching the course's materials, we [tutors] had to teach [more basic] materials that should have been taught at school. I asked them about [personal] pronouns, but they had no idea and did not know how to use them [...].

Tutors reported having implemented some common strategies, which tended to be very directive, when dealing with less competent students, such as exercising extra patience, explaining things slowly, and frequently checking students' understanding. Another strategy the tutors (Tutors E, F, I, and J) mentioned utilizing when assisting the less proficient tutees was repetition by re-explaining concepts and reviewing them over and over to help them remember the grammatical concepts, many of which were relatively new to them. However, by giving tutees more practice through exercises, the tutors also played a facilitative role.

Another strategy tutors (C, F, and J) mentioned for helping weak students was making notes and summaries for the tutees. As many concepts were relatively complex or new to their tutees, the strategy was helpful. Tutor C said:

> For teaching about tenses, I usually make notes that are easy for them to understand. Sometimes learning about tenses from the [course] book makes students lazy because it is thick, so I make notes to help them understand [the concepts] more easily.

According to Tutor C, finding more creative but simple ways to explain grammatical concepts to students with low English proficiency, especially slow learners, was part of her task as a tutor. This kind of an attitude is essential to the success of a peer tutoring program and the course it supports, as substantial individual knowledge gaps cannot be addressed in the grammar courses' regular meetings. What the tutors did to help the low-level students understand the subject matter represents another essential advantage of peer tutoring. With its focused and personalized attention, peer tutoring can greatly benefit weak learners, which was shown by significantly increased test scores for many of them when they were repeating the course. This finding corroborates Green, Alderman, and Liechty's (2004) research finding on the use of peer teaching for at-risk second grade readers.

Dealing with Critical, Highly Proficient Students. Two tutors (Tutors J and G) who were assigned highly proficient students found it challenging to assist them in learning grammar. Being more critical, these students often raised questions that needed more expertise in order to construct a more comprehensive understanding of the grammatical items taught. This sometimes led to difficult situations as the tutors themselves often still had knowledge gaps about the subject matter.

To overcome this problem, Tutor J did not do any special preparation for a tutoring session with the tutees. However, when she did not know the answer to her tutees' question(s), she would tell them that she needed to find out the answer first and consult a more able tutor or search the internet to help her. This showed that a tutors' network to provide peer support could be very helpful, and, therefore, may be necessary. Unlike Tutor J, Tutor G reported getting himself more prepared by "rereading the materials until [he] fully understood the concepts [he] was going to explain to the tutees". He also prepared more advanced examples for uses of the relevant L2 forms in real contexts. He pointed out how working with the most able tutees had benefited him for his own learning:

> [...] these students could help me to reflect on my own grammar knowledge – how much I had understood the concepts and uses of those grammatical items.

Tutors J and G reported having taken a more facilitative role for the proficient tutees as they were actively engaged and took the initiative more often. Tutor G said, "things just flowed as they started asking questions". Despite this, Tutor G said he still prepared a general structure, which he wrote down on a note pad to make his sessions more organized. However, although the tutors tended to act as facilitators for more proficient students, the finding demonstrates that such tutees need tutors with a high level of competence so that they can handle the tutees' questions.

Dependence on Peer Tutors. Tutors E and G, on the other hand, raised the issue of dependence among some lower-level students. According to the tutors, instead of utilizing the peer sessions to support their own studies, they relied on their peer teachers to help them learn the grammar course materials. During the peer tutoring sessions, tutors had to be more directive in helping such tutees. Tutor G pointed out, "[…] they did not want to read the materials themselves," so he had to explain a lot to them. Commenting on this issue, Tutor E said that such students did not seem to be very motivated to study independently because they found learning by themselves challenging. So rather than reading the course materials by themselves, they just relied on the peer teachers to explain everything to them. This nature of reliance is different from what was reported in Mynard and Almarzouqi's (2006) research, where tutees asked for more sessions and assistance to complete their homework. The tutees' reluctance to be engaged more cognitively supports previous findings that learners do not always favor constructivist-oriented ways of learning in which they will take more responsibility (e.g., Struyven, Dochy, Janssens, & Gielen, 2008).

The tutees' 'cognitive dependence' on tutors in the present study highlights a crucial possible downside of peer-assisted learning that is not generally reported in peer teaching programs. Rather than enhancing learners' autonomy among tutees (Bohórquez et al., 2019; Stigmar, 2016), it may also create an easy shortcut for them to fulfill their study tasks. Students' overreliance seemed to have been heightened by the traditional knowledge transmission strategies commonly applied in Indonesia, in which they play a more passive role in learning. Another factor that seemed to have contributed to this issue was the perceived complexity of the subject matter. Involving the discussion of complex grammar forms and linguistic concepts in detail, learning in isolated grammar courses can be complicated and demanding.

To tackle students' overreliance, the tutors who raised the dependence issue devised a strategy to give the tutees exercises in conjunction with their sessions to do at home. In this way, they stimulated the students to collaborate more in the knowledge construction of the tutoring. According to Tutor G, this strategy was helpful because he could not just remind them to study grammar.

Culture-Related Issues

There were two culture-related issues raised in the interviews. The first was concerned with the Indonesian collective culture. Within such a culture, people prefer to work in a group to maintain togetherness and ensure that support is available to all the members. Because of this, while having different levels of competence, some students did not want to be grouped in different sessions because they wanted to do peer tutoring with their close classmates. Working in such a group may lessen one's anxiety, but this may present difficulties when the tutees have different levels of abilities. Tutor F said this was not always easy to deal with but the course's teacher or program coordinator could intervene to assist tutors by establishing more appropriate groupings.

The second problem was concerned with tutor-tutee relations, which may be linked to local (Javanese) cultural values for relating to other people. Tutors E, G, and H reported feeling uncomfortable teaching a fellow student from the same batch or a more senior one. This seemed to stem from the value of respect ('hormat') that entails humility ('andhap asor') and empathy ('tepa selira'). Respect is a crucial for maintaining social harmony, which is a central value in Javanese culture, and it is of great importance especially when one interacts with people with a higher status, including those who are older (Geertz, 1960; Irawanto, Ramsey, & Ryan, 2011). As teaching is usually associated with superiority in knowledge, experience, and/or skills, peer teachers in a cultural context may be concerned that they could make older students or those of the same status feel like they are inferior or that they, as student-teachers, may appear patronizing. Such an issue may present particular challenges in peer tutoring as it may affect the process of knowledge transfer/knowledge construction (see Takeuchi, 2015). Tutor G pointed out that initially he did not find it easy to teach a tutee who happened to be a student from the same year but by using a 'mate-to-mate' approach he believed he could solve the problem. He added:

> I found it a bit awkward to teach my own mate […]. [W]ell, it was my own mate, so I tried to be as natural as possible [just like doing it to a mate when explaining things], while maintaining mutual respect […]. He was also very open and humble [so it was fine].

Tutor H, who was tutoring a more senior student repeating a grammar course, also faced a similar problem. Tutor H felt 'sungkan', an Indonesian word referring to the feeling of awkwardness and reluctance to take actions toward someone due to some reason (such as the person having higher status) when teaching the tutee.

She did not want to sound condescending or show that she was smarter than the senior student. To overcome this problem, she consulted the program coordinator. After receiving some encouragement and advice, she felt more confident. However, her own strategy was to build a cordial rapport with the tutee so she could apply a friend-to-friend method for tutoring her. Based on an equal relation, this strategy allowed her to cooperate in the knowledge construction effectively and help the tutee improve her understanding of the subject matter. Similar to Tutor G, she said, "I explained things to her like talking to a friend." She also reported being not too rigid with the learning agenda, and having small talks during the tutoring session were helpful for creating a more relaxed atmosphere to encourage the senior tutee to ask questions and talk about her problems. The use of social talk and tutees' verbalization of problems confirms Madaio, Ogan, and Casell's (2016) research results on friendship-based tutor-tutee interactions in peer teaching.

Conclusion

The purpose of this study was to identify peer teachers' perceived challenges when tutoring their peers and the strategies they self-reportedly applied to overcome them. The difficulties were due to factors such as low motivation, the limited availability of supporting teaching materials, different levels of students' L2 competence, and the tutees' socio-cultural status. However, in general, the problems seem to have mainly stemmed from the dynamics of dealing with the different characteristics of students, which they had yet to learn to manage as non-professional teachers. In spite of this, with the peer teaching program receiving positive assessments from the faculty, the tutors had demonstrated their success as students' partners in grammar learning. Working at a more personalized level, the tutors could address individual problems in different areas, which would not have been able to be sufficiently dealt with in the isolated grammar courses' regular classes. With students' improved grades, this research shows that cross-level peer education may provide essential support to improve students' grammar in EFL contexts where the language element has not received adequate attention.

This study also highlights the importance of employing specific criteria in tutors' recruitment as their success in peer tutoring is inseparable from their expertise and relevant soft skills. However, while the student-teachers were shown to be able to adequately deal with the difficulties they reportedly faced, it was demonstrated that the faculty's support remained crucial for the tutors in the partnership. Working together to develop students' motivation to learn grammar and provide appropriate materials for practice were shown to be two areas the faculty could assist the most. The provision of materials that allows scaffolding could be essential to help those with low language aptitude.

Finally, since this is a small-scale study, the present research is not generalizable. More investigations are necessary to examine grammar tutors' challenges and their strategies to tackle them in other contexts involving more participants, including tutees and faculty members. As this study only relied on limited data from interviews, future research also needs to include data from other sources, such as observations of the peer teaching and examinations of the supplemental materials the tutors use. Despite these limitations, the present study has provided some crucial insights into the issues tutors face in grammar peer-teaching and the strategies they implement to overcome them. Such information is indispensable for providing an effective learning support system for grammar learning through peer tutoring.

Acknowledgments

The authors would like to thank our anonymous reviewers for their invaluable suggestions and comments and language editors who have polished our English.

Conflict of interests

The authors declare that they have no conflict of interest.

References

Adams, J. (2011). Investigating peer tutoring for academic writing support in a UK university. *Journal of Academic Writing*, *1*(1), 110-119. https://doi.org/10.18552/joaw.v1i1.1

Asghar, A. (2010). Reciprocal peer coaching and its use as a formative assessment strategy for first-year students. *Assessment & Evaluation in Higher Education*, *35*, 403-417. https://doi.org/10.1080/02602930902862834

Bell, D. C., & Elledge, S. R. (2008). Dominance and peer tutoring sessions with English language learners. *Learning Assistance Review*, *13*(1), 17-30.

Berghmans, I., Michiels, L., Salmon, S., Dochy, F., & Struyven, K. (2014). Directive versus facilitative peer tutoring? A view on students' appraisal, reported learning gains and experiences within two differently-tutored learning environments. *Learning Environ Res*, *17*, 437-45. https://doi.org/10.1007/s10984-014-9168-8

Bohórquez, L. I., Rodríguez, J. D., & González, J. J. (2019). Online peer-tutoring: A renewed impetus for autonomous English learning. *HOW Journal*, *26*(2), 12-31. https://doi.org/10.19183/how.26.2.503

Bowman-Perrott, L., deMarín, S., Mahadevan, L., & Etchells, M. (2016). Assessing the academic, social, and language production outcomes of English language learners engaged in peer tutoring: A systematic review. *Education and Treatment of Children*, *39*(3), 359-388. https://doi.org/10.1353/etc.2016.0016

Braun, V., & Clarke, V. (2006). Using a thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Christiansen, B., Bjørk, I.T., Havnes, A., & Hessevaagbakke, E. (2011). Developing supervision skills through peer learning partnership. *Nurse Education in Practice*, *11*(2), 104-108. https://doi.org/10.1016/j.nepr.2010.11.007

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom language acquisition* (pp. 42-63). Cambridge University Press.

DeKeyser, R.M. (2008). Introduction: Situating the concept of practice. In R.M. DeKeyser (Ed.), *Practice in second language: Perspectives from Applied Linguistics and Cognitive Psychology* (2nd ed., pp. 1-18). CUP.

DeKeyser, R.M. (2015). Skill Acquisition theory. In B. Van Patten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 94-112). Routledge.

Ellis, R. (2015). The importance of focus on form in communicative language teaching. *Eurasian Journal of Applied Linguistics, 1*(2), 1-12. https:/doi.org/10.32601/ejal.460611

Everaert, P., Opdecam, E., & Maussen, S. (2017). The relationship between motivation, learning approaches, academic performance and time spent. *Accounting Education*, *26*(1), 78-107. https:/doi.org/10.3200/PSFL.49.1.11-18

Gaies, J. S. (1985). *Peer involvement in language learning*. Prentice Hall, Inc.

Geertz, C. (1960). *The religion of Java*. Chicago: The University of Chicago Press.

Green, S. K., Alderman, G., & Liechty, A. (2004). Peer tutoring, individualized intervention, and progress monitoring with at-risk second-grade readers. *Preventing School Failure: Alternative Education for Children and Youth, 49*(1), 11-17, https://doi.org/10.3200/PSFL.49.1.11-18

Hidayah, N. (2014). *The effectiveness of peer tutoring to students' understanding in using Simple Past Tense* [Unpublished thesis]. Universitas Islam Negeri Syarif Hidayatullah. http://repository.uinjkt.ac.id/dspace/handle/123456789/1897

Irawanto, D. W., Ramsey, P.L., & Ryan, J.C. (2011). Challenge of leading in Javanese culture. *Asian Ethnicity, 12*(2), 125-139. https://doi.org/10.1080/14631369.2011.571829

Irvine, S., Williams, B., & McKenna, L. (2018). Near-peer teaching in undergraduate nurse education: An integrative review. *Nurse Education Today*, *70*, 60-68. https://doi.org/10.1016/j.nedt.2018.08.009

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772-775. https://doi.org/10.1126/science.1199327

Kim, J. (2018). Could you check my grammar? The forbidden request in the university writing center. *International Journal of Research Studies in Language Learning*, *7*(4), 21-37. https://doi.org/10.5861/ijrsll.2017.2000

LaClare, E., & Franz, T. (2013). Writing centers: Who are they for? What are they for? *Studies in Self-Access Learning Journal*, *4*(1), 5-16.

Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching, 48*(2), 263-280. https://doi.org/10.1017/S0261444814000408

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, *31*(3), 313-327.https://doi.org/10.1080/02796015.2002.12086158

Long, M. (2016). In defense of tasks and TBLT: Non-issues and real issues. *Annual Review of Applied Linguistics, 36*, 5-33. https://doi.org/10.1017/S0267190515000057

Madaio, M. A., Ogan, A, & Cassell, J. (2016). The effect of friendship and tutoring roles on reciprocal peer tutoring strategies. In A. Micarelli, J. Stamper & K. Panourgia (Eds.), *International conference on intelligent tutoring systems* (pp.423-429). Springer International Publishing.

McKenna, L., & French, J. (2011). A step ahead: Teaching undergraduate students to be peer teachers. *Nurse Education in Practice*, *11*(2), 141-145. https://doi.org/10.1016/j.nepr.2010.10.003

Menezes, S.D., & Premnath, D. (2016). Near-peer education: A novel teaching program. *International Journal of Medical Education*, *7*, 160-167. https://doi.og/10.5116/ijme.5738.3c28

Mulatsih, M.V.E. (2018). The effectiveness of the grammar tutoring program based on students' feedback in batch 2016 ELESP. *LLT Journal*, *21*(2), 175-182. https://doi.org/10.24071/llt.2018.2102016

Murtisari, E. T., Hastuti, G., & Arsari, R. W. (2019). EFL students' perception of grammar teaching: Isolated or integrated in skill courses? *Asian EFL Journal*, *26*(6), 207-226.

Murtisari, E. T., Salvadora, L., & Hastuti, G. (2020). Isolated and integrated grammar teaching in a tertiary EFL context: Indonesian teachers' beliefs. *SAGA: Journal of English Language Teaching and Applied Linguistics, 1*(1), 17-30. https://doi.org/10.21460/saga.2020.11.9

Mynard, J., & Almarzouqi, I. (2006). Investigating peer tutoring. *ELT Journal*, *60*(1), 13-22. https://doi.org/10.1093/elt/ccio77

Naeger, D. M., Conrad, M., Nguyen, J., Kohi, M., & Webb, E. (2013). Students teaching students: Evaluation of "near peer" teaching experience. *Acad Radiol*, *20*(9), 1177-82. https://doi.org/10.1016/j.acra.2013.04.004

O'Sullivan, I., & Cleary, L. (2014). Peer-tutoring in academic writing: The infectious nature of engagement. *Journal of Academic Writing*, *4*(1), 52-65. https://doi.org/10.18552/joaw.v4i1.72

Saeed, S., & Zyngier, D. (2012). How motivation influences student engagement: A qualitative case study. *Journal of Education and Learning*, *1*(2), 252-267. https://doi.org/10.5539/jel.v1n2p252

Sheen, R. (2005). Focus on forms as a means of improving accurate oral production. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 271-310). Mouton de Gruyter.

Snyder, D. W., Nielson, R. P., & Kurzer, K. (2016). Foreign language writing fellows programs: A model for improving advanced writing skills. *Foreign Language Annals, 49*(4), 750–771. https://doi.org/10.1111/flan.12231

Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, *42*(2), 181-207. https://doi.org/10.1002/j.1545-7249.2008.tb00115.x

Stanley, C. (1998). A framework for teacher reflectivity. *TESOL Quarterly*, *32*(3), 584-591. https://doi.org/10.2307/3588129

Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The importance of students' motivation for their academic achievement: Replicating and extending previous findings. *Frontiers in Psychology, 10*(1730), 1-11. https://doi.org/10.3389/fpsyg.2019.01730

Stigmar, M. (2016). Peer-to-peer teaching in higher education: A critical literature review. *Mentoring &* Tutoring*: Partnership in* Learning, *24*(20), 124-136. https://doi.org/10.1080/13611267.2016.1178963

Struyven, K., Dochy, F., Janssens, S., & Gielen, S. (2008). Students' experiences with contrasting learning environments: The added value of students' perceptions. *Learning Environments Research, 11*, 83-109. https://doi.org/10.1007/s10984-008-9041-8

Takeuchi, H. (2015). Peer tutoring in Japan: A new approach for a unique educational system. *Studies in Self-Access Learning Journal*, *6*(1), 112-119. https://doi.org/10.37237/060108

Tsang, E.W.K. (1998). Inside story: Mind your identity when conducting cross-national research. *Organization Studies*, *19*(3), 511-515. https://doi.org/10.1177/017084069801900307

Umeda, M., Snape, N., Yusa, N., & Wiltshier, J. (2017). The long-term effect of instruction on learners' knowledge of English articles. *Language Teaching Research, 23*(2), 1-21. https://doi.org/10.1177/1362168817739648

Watcharapunyawong, S. (2018). The effects of online peer tutoring on first year undergraduate students' English grammar achievement. *Indonesian Journal of Informatics Education, 2*(2), 69-76. https://doi.org/10.20961/ijie.v%vi%i.24382

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interests: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1–35. https://doi.org/10.1016/j.dr.2009.12.001

Williams, B., & Reddy, P. (2016). Does peer-assisted learning improve academic performance? A scoping review. *Nurse Education Today*, *42*, 23-29. https://doi.org/10.1016/j.nedt.2016.03.024

Winder, R., Kathpalia, S. S., & Koo, S. L. (2016). Writing centre tutoring sessions: Addressing students' concerns. *Educational Studies*, *42*(4), 1-17. https://doi.org/10.1080/03055698.2016.1193476

Won, J., Change, H., & Kang, M.K. (2017). The effect of peer learning on college EFL learners in Korea. *Information, 20*(12), 8309-8316.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 4, 2020

Štubňa, P. (2020). Beyond «Listen and Repeat»: Investigating English Pronunciation Instruction at the Upper Secondary School Level in Slovakia: A Book Review. *Journal of Language and Education, 6*(3), 216-220. https://doi.org/10.17323/jle.2020.10919

Beyond «Listen and Repeat»: Investigating English Pronunciation Instruction at the Upper Secondary School Level in Slovakia by R. Metruk: A Book Review

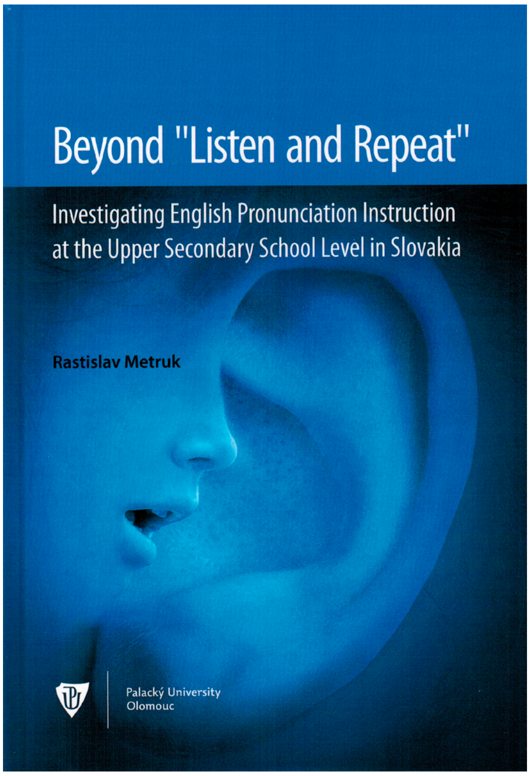
Pavol Štubňa

Comenius University

Correspondence concerning this article should be addressed to Pavol Štubňa, Department of Romance Studies, Comenius University, Gondova 2, Bratislava, 811 02, Slovak Republic.
E-mail: pavol.stubna@uniba.sk

Metruk, R. (2020). *Beyond "Listen and Repeat": Investigating English pronunciation instruction at the upper secondary school level in Slovakia*. Palacký University Olomouc.

The author of the monograph, Dr. Rastislav Metruk (a senior lecturer at the Department of English Language and Literature, Faculty of Humanities, University of Žilina, Slovakia), is a prolific academic scholar with extensive experience in teaching, linguistic and didactic research, and publishing activities. His expertise in the field of phonetics and phonology is confirmed by several books and remarkable research articles, for instance 'Pronunciation of English Dental Fricatives by Slovak University EFL Students' in the *International Journal of English Linguistics* (Metruk, 2017a), "West" or "Vest"? Pronunciation of English of English Consonants [w] and [v] in the Utterances of Slovak EFL Speakers' in the *Journal of Language and Education* (Metruk, 2018a), or 'Comparing Holistic and Analytic Ways of Scoring in the Assessment of Speaking Skills' (Metruk, 2018b). In his books, *A Concise Introduction to General American Pronunciation: Segmental Features* (Metruk, 2017b) and *A Concise Introduction to General American Pronunciation: Suprasegmental Features* (Metruk, 2019), he deals with segmental and suprasegmental phenomena not only from the phonological point of view, but also from the perspective of effective interlingual communication and EFL (English as a foreign language) didactics.



The purpose of *Beyond «Listen and Repeat»* is to investigate pronunciation teaching practice at the upper-secondary school level in Slovakia from segmental and suprasegmental points of view. Traditionally, fluency and accuracy are acknowledged as two major components of speaking ability in the process of teaching a foreign language (Ekbatani, 2010). While fluency usually refers to the ease and speed of a speech flow, accuracy predominantly concerns the accurate usage of language systems such as grammar, vocabulary, and pronunciation. Analogically, English pronunciation can be approached from the same standpoints: fluency and accuracy. While segmental features correspond more closely with accuracy, the prosodic features relate more directly to fluency.

To briefly summarize the contents of the monograph, the opening chapter deals with the current pronunciation teaching practices in Slovakia, i.e. objectives, models, activities, technology affordance, etc. Moreover, the segmental versus suprasegmental debate is introduced. Chapter two analyzes the selected factors that predominantly affect EFL pronunciation, e.g. age, motivation, exposure to the target language, learner's native language interference. The third chapter introduces original research carried out at the Slovak upper-secondary school level. Chapter four defines the methodology of the research and chapters five to seven shed light on research methods and the research results are analyzed. Finally, in the eighth chapter, conclusions are formulated and recommendations provided.

Currently, the teaching of (English) pronunciation is still often neglected by educational institutions and teachers worldwide (Kılıçkaya, 2011; Szpyra-Kozłowska, 2015). Traditionally, instructors concentrate primarily on grammar, vocabulary, and the four language skills without incorporating pronunciation adequately into the teaching and learning process. Some teachers even assume that students can learn correct pronunciation without explicit training and instructions (Gilakjani, 2016).

However, the author of the monograph argues that the EFL teachers should be more aware of the importance of pronunciation and pay more attention to it in their curricula. The primary aim of communication is not only to understand what others say, but also to make yourself understood. In fact, grammar and vocabulary can become useless, if the speaker cannot pronounce the required language elements accurately. On the other hand, if a learner's pronunciation is appropriate or at least acceptable within oral speech, other speakers can understand him/her despite grammatical errors. In other words, correct pronunciation plays a significant role in efficient language production.

In order to develop appropriate pronunciation skills, English teachers are supposed to, first, carry out pronunciation diagnostic analysis and select aspects that can have a great effect on changing the speech of learners towards increased comprehensibility. For this reason, it is important to set both short-term and long-term objectives, and finally, to promote and incorporate new (creative) instructional modes and modules (e.g. pair-work, small-group work, whole-class activity, audio/video recordings, specialized computer software, self-study) (Morley, 1991). In fulfilling the task of achieving the proper phonological control, pronunciation instruction plays a huge role. Through incorporating proper and/or particular pronunciation instruction into the teaching and learning process, students do not only improve their speaking skills but also increase self-confidence (a positive self-image) and develop speech-monitoring and speech-modification strategies (Gilakjani, 2012).

EFL teachers should constantly look for efficient new methods and forms of indicating, practicing, and providing feedback on English pronunciation. The methodologies they adopt are supposed to be appropriate, efficient, and captivating for learners (teachers usually think that pronunciation practice is too monotonous for students). Other works on pronunciation that attempt to draw a broader picture on this complex language system and could be beneficial to EFL educators and researchers include Derwing and Munro (2015), Isbell, Park, and Lee (2019), and Mompean and Fouz-González (2015).

As mentioned above, in the second chapter of the monograph, the author deals with factors influencing the acquisition of proper or acceptable pronunciation: individual phonetic abilities, motivation (positive and negative), and specific personality traits (openness, mental flexibility, curiosity, responsibility, creativity, persistence, etc.). In most cases, the aim of teaching (English) pronunciation is not to sound like native speakers but to reach a state of comfortable intelligibility, i.e. that the speaker's English is easily understandable. However, for some teachers (or learners), their aim is to achieve a native-like accent but this is often an unattainable goal.

As often confirmed by experience, strongly accented speech or an accent that is too different from the norm can become confusing and cause negative judgements on the speaker's personality and competences (Morley, 1994). We must be aware of the fact that appropriate pronunciation is also a significant metacommunicative factor as it not only affects our overall ability to understand a foreign language's speech flow, but it subconsciously classifies the speaker at several socio-cultural levels (country of origin, education achieved, linguistic or other cognitive abilities), and thus subconsciously determines our response to what a person says.

On the contrary, having appropriate pronunciation may not automatically guarantee the correct use of grammar and vocabulary, although a native-like accent automatically evokes in the listener the notion (and feelings) of an excellent command of language, particularly if the listener is a person who does not demonstrate a sufficient mastery of the language. On the other hand, poor pronunciation skills typically create an impression of linguistic insufficiency and they may cause certain problems on the side of the listener.

Therefore, it is suggested that the learner's accent become as close as possible to the language standards. Teachers ought to raise awareness in learners of both the American and British varieties of English and students should be able to understand them both (Mompean, 2004). Other models have been also proposed, including one for intelligible pronunciation such as Lingua Franca Core (Jenkins, 2000), or even a non-native speaker who is intelligible and comprehensible (Murphy, 2014). It should be also noted that, traditionally, a single model should be chosen; yet it has been emphasized that no single model ought to be forced upon students (Archer, 2018). Concurrently, not only learners, but also teachers – serving as examples to the students – should be accurately trained to improve their pronunciation, which could lead to the subsequent improvement of their pupils' pronunciation.

As far as research methodology is concerned, empirical research in the field of natural sciences usually differs considerably from that of social sciences and humanities in several respects. Its research methods and tools are not only precise, measurable, reliable, but also independent of the person who makes the observations, i.e. they are impartial and objective. A valid scientific method must be «logical, confirmable, repeatable and scrutinizable. Any branch of inquiry that does not allow the scientific method to test its basic laws or theories cannot be called *science*» (Bhattacherjee, 2012, p. 6).

The current methodologies utilized in social sciences and the humanities include a vast range of approaches, tools, and techniques – both qualitative and quantitative – to make valid observations, interpret results correctly, and generalize them. In any case, they necessarily imply the use of mental abstractions like the unit of analysis, proposition, hypothesis, construct, (independent and dependent) variable, model, paradigm, and so forth.

Moreover, a solid (e.g. functionalist) research methodology must certainly comprise the phases of exploration (research questions, literature review, theory), research design (operationalization, research method, sampling strategy), and research execution (pilot testing, data collection, data analysis) since all scientific research is an iterative process of observation, rationalization, and validation (Bhattacherjee, 2012, p. 20). To increase the validity of evaluations and research findings, a «triangulation» strategy is often used in social sciences and the humanities that effectively integrates the advantages of qualitative and quantitative approaches so «by combining multiple observers, theories, methods, and empirical materials, researchers can hope to overcome the weakness of intrinsic biases and the problems that come from single-method, single observer and single-theory studies» (Denzin, 1978, p. 307).

In accordance with the above-mentioned triangulation strategy and functionalistic research methodology, the author of *Beyond «Listen and Repeat»* applied the following exploration tools and methods: a) current textbooks' *content analysis*, b) an inquiry of EFL teachers on related aspects by means of a *questionnaire*, and c) *interview*ing students on their attitudes towards pronunciation teaching practice.

The results of the textbook *analysis* (a) revealed that, predominantly, suprasegmental features are practiced and the drill technique is prevalent during teaching at the Slovak upper-secondary school level.

The data obtained by the *questionnaire* (b) suggest that although teachers regard pronunciation as an important component, a larger amount of pronunciation work should be incorporated in the course textbooks; particularly, suprasegmental categories appear to be more problematic for them to deal with.

The *interview* (c) data analysis revealed that students recognize and realize the importance of pronunciation instruction – both segmentals and suprasegmentals. However, stress, intonation, and rhythm seem to cause greater problems for the students, mainly because of the fact that the stress patterns of the English language differ fundamentally from Slovak.

It is surprising that several students reported that prosody is often not taught at the upper-secondary school level in Slovakia, which is, without a doubt, an issue that needs to be dealt with. The author further maintains that the importance of both segmental and suprasegmental features in EFL teaching is undeniable and the neglect of either of the two phenomena could result in pronunciation problems of EFL learners as well as lower (overall) speaking ability.

In his final chapter, the author suggests that pronunciation instruction at the upper-secondary levels in Slovakia ought to be given more time. Further, prospective teachers need to undergo a quality pronunciation training, predominantly because the English phonological system is normally considered to be particularly difficult for many EFL learners. Finally, the author calls for further examination in this area by conducting additional content analyses and interviews, and distributing questionnaires to more participants.

To conclude this review, it seems that the author's aims have been successfully fulfilled. The contents of the book appear accurate, clear, relevant, logically organized, and appropriate within and among singular units of study. The research procedure, the interpretation of results, and the final discussion are rigorous, consistent, and pertinent to the topic. The theoretical and practical recommendations provided by the author of the monograph appear exceedingly useful, motivating, and revealing. The need to develop a new textbook (or an original set of textbooks) with more pronunciation activities that would be more suitable for the Slovak learners of English has been confirmed.

References

Archer, G. (2018). *Pronunciation models in regional environments: a comparison and assessment of RP and SSE* [Unpublished MRes thesis]. University of Glasgow.

Bhattacherjee, A. (2012). *Social science research: Principles, methods and practices*. University of South Florida.

Denzin, N. (1978). *The research act*. Prentice Hall.

Derwing, T., & Munro, J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Ekbatani, J. (2010). *Measurement and evaluation in post-secondary ESL*. Routledge.

Gilakjani, A. P. (2016) English pronunciation instruction: A literature review. *International Journal of Research in English Education, 1*(1), 1-6. http://ijreeonline.com/article-1-21-en.pdf

Isbell, D., Park, O., & Lee, K. (2019). Learning Korean pronunciation: Effects of instruction, proficiency, and L1. *Journal of Second Language Pronunciation, 5*(1), 13–48. https://doi.org/10.1075/jslp.17010.isb

Jenkins, J. (2000). *The phonology of English as an international language.* Oxford University Press.

Kılıçkaya, F. (2011). Improving pronunciation via accent reduction and text-to-speech software. In M. Levy, F. Blin, C. Siskin & O. Takeuchi (Eds.), *WorldCALL. International perspectives on computer-assisted language learning* (pp. 85-96). Routledge.

Metruk, R. (2017a). Pronunciation of English dental fricatives by Slovak university EFL students. *International Journal of English Linguistics, 7*(3), 11–16. http://dx.doi.org/10.5539/ijel.v7n3p11

Metruk, R. (2017b). *A concise introduction to general American pronunciation: Segmental features*. Palacký University.

Metruk, R. (2018a). "West" or "vest"? Pronunciation of English consonants [w] and [v]in the utterances of Slovak EFL speakers. *Journal of Language Education, 4*(2), 24-29. https://doi.org/10.17323/2411-7390-2018-4-2-24-29

Metruk, R. (2018b). Comparing holistic and analytic ways of scoring in the assessment of speaking skills. *The Journal of Teaching English for Specific and Academic Purposes, 6*(1), 179-189. https://doi.org/10.22190/JTESAP1801179M

Metruk, R. (2019). *A concise introduction to general American pronunciation: Suprasegmental features*. Palacký University.

Mompean, J. (2004). Options and criteria for the choice of an English pronunciation model in Spain. In J. Oro et al. (Eds.), *Linguistic perspectives from the classroom: Language teaching in a multicultural Europe* (pp. 1043–1059). Servizo de Publicacións e Intercambio Científico da USC.

Mompean, J., & Fouz-González, J. (Eds.). (2015). *Investigating English pronunciation trends and directions*.

Palgrave Macmillan.
Morley, J. (1991) The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481-520. https://doi.org/10.2307/3586981
Morley, J. (1994). A multidimensional curriculum design for speech-pronunciation instruction. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New views, new directions* (pp. 64-91). TESOL.
Murphy, J. (2014) Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System, 42*, 258–269. https://doi.org/10.1016/j.system.2013.12.007
Szpyra-Kozlowska, J. (2015). *Pronunciation in EFL instruction. A research-based approach*. Multilingual Matters.

Thanking our reviewers, 2020

The editors of Journal of Language and Education would like to express their gratitude to all the reviewers who helped us work on each issue.

You helped us take JLE from an idea 6 years ago to a full-fledged quality publication indexed in Scopus, Web of Science, DOAJ, EBSCO, etc.

The reviewers' expertise and benevolence contributed greatly to this success. We value and truly appreciate the time and effort of our reviewers who provided insightful analysis of the manuscripts and helped us to assure quality of the journal.

Your reviews became the foundation of every decision the editorial board made regarding the submissions. At the same time, the reviewers' input proved incredibly helpful to the authors who took their manuscripts to the next level of publication quality by responding to your comments and suggestions.

Our reviewers are an essential part of the JLE family, so we would like to take this opportunity to thank you and to express our sincere appreciation of your contribution!

Aaron Kozbelt
Aaron Olaf Batty
Abd. Basith
Abdorreza Tahriri
Abdulghani Muthanna
Abu Bakar Mohamed Razali
Abu Nawas
Abuelgasim Sabah Elsaid Mohammed
Adam Butt
Adeola Folasade Akinyemi
Aditi Bhatia
Adrian K.C. Lee
Agota Kun
Agurtzane Azkarai Garai
Agustín Vicente
Ahmad Mohammad Ahmad Al-Harahsheh
Ahmar Mahboob
Ahmed Abdulateef Sabti
Ahmed Ahmed Shoeib
Ahmed Alduais
Ahmed Masrai
Ainhoa Aparicio Fenoll
Aintzane Doiz
Aiyoub Jodairi Pineh
Aki Tsunemoto
Akifumi Yanagisawa
Akiko Nagao
Al Tiyb S Al Khaiyali
Alan K. Goodboy
Alejandro Daniel Domínguez González
Alessandra Molino
Alexander Nikolaou
Ali Abdulhussein Sagban
Ali Dabbagh
Ali Derakhshan
Ali H. Al-Hoorie
Ali Roohani
Alison Fox
Alka S. Kelkar
Amanda Godley
Amanda K. Kibler
Amare Tesfie Birhan
Amir Mahdavi-Zafarghandi
Amjjad Sulaimani
Ana Cristina Lahuerta Martinez
Ana Maria Ferreira Barcelos
Anastasia Gkaintartzi
Anastasios Mikropoulos
Andrea Revesz
Andrea Schalley
Andrés Canga Alonso
Andrew D. Cohen
Andrew Schenck
Andrew Sewell
Andy Xuesong Gao
Ang Leng Hong
Angel Arias
Angela Smith
Angelica Carlet
Angelica Galante
Anie Attan
Anita Habók
Anita Lie
Anna Åhlund
Anna Czura Anna Czura
Anna Dziemianko
Anna Tristram
Anne Edstrom
Anne Schippling
Annie Tremblay
Anouk Ticheloven
Anssi Roiha
Anthony Leddicoat
Antonia Estrela
Antonina Laposhina
Antonio Di Paolo
Arif Ahmed Mohammed Hassan Al-Ahdal
Arif Husein Lubis
Arnie Cann
Asama Tasanameelarp
Ashish Kumar Chaubey
Asier Calzada
Atefeh Rezanejad
Athip Thumvichit
Atikhom Thienthong
Aubrey Neil Leveridge
Aya Matsuda
Ayca Uner
Ayhan Kahraman
Ayse Taskiran
Azza A. M. Abdelrahim
Bachir Bouhania
Badreddine Hamma
Bahattin Şimşek
Bahman Gorjian
Bahram Behin
Barbara Skinner
Barry Lee Reynolds
Batia Laufer
Beatriz Santana Perera
Bedrettin Yazan
Benjamin Moorhouse
Benjamin S. Selznick
Bethany Gray
Betsy Gilliland
Beverly Baker
Bilal Genc
Biljana Radić-Bojanić
Biook Behnam
Björn Rasch
Blanka Klimova
Bogdana Huma
Boning Lyu
Bonnie Wing-Yin Chow

Booyuel Kim
Brad Perks
Bram Bulté
Branka Drljača Margić
Brent Allan Kelsen
Brian Paltridge
Budi Setyono
Buratin Khampirat
Camino Bueno-Alastuey
Canay Karci Aktaş
Carie S. Tucker King
Carlo Consani
Carmen Argondizzo
Caroline Barratt-Pugh
Caroline Hahnel
Caroline Rowland
Carolyn Blume
Carrie Symons
Carrie Yea-Huey Chang
Cassi L. Liardét
Cassie J. Brownell
Cassie Smith-Christmas
Caterina Fiorilli
Catherine Barbour
Cecilia Andorno
Cecilia Guanfang Zhao
Cecilio Lapresta Rey
Celia Thompson
Cemil Gokhan Karacan
Chan Suet Fong
Chan Swee Heng
Chantarath Hongboontri
Charles Clifton
Chatree Faikhamta
Cheng-Huan Chen
Chibueze Maureen
Chinh Duc Nguyen
Christian Krekeler
Christina Gkonou
Christine C. M. Goh
Christopher G. Thompson
Chunchun Yeh
Chunlin Yao
Ciler Hatipoğlu
Claire Martinot
Claire Rodway
Clarence Gerald Green
Colin Milligan
Conchita M. Temporal
Coral Hunt-Gómez
Corinne A. Seals
Cristina Suárez-Gómez
Cristine Mcmartin-Miller
Csaba Kálmán
Cüneyt Demir
Cynthia S Wiseman
Dan Melzer
Dana Di Pardo Leon-Henri
Dana Ferris
Dang H. Tri
Daniel Bosmans
Daniel Corral
Daniel Isbell
Danielle Joulia
Darío Luis Banegas
Darren Joseph Cunningham
Darrin Griffin
David A. M. Peterson
David Banks
David Duran
David Lasagabaster
Deborah Dubiner
Dedi Turmudi
Denchai Prabjandee
Denisa Šukolová
Dennis Duncan
Dennis E. Clayson
Derek Denis
Desirée Schmuck
Diler Aba
Dimitrios Vlachopoulos
Dimitris Serafis
Dina Abdel Salam El-Dakhs
Donita Joy Shaw
Dorothy Chun
Durk Gorter
Dyah Rochmawati
Dzintra Iliško,
Edgar Emmanuell Garcia-Ponce
Ehsan Mehri
Ehsan Namaziandost
Eija Aalto
Eileen Mcbride
Ela Sjolie
Eleanore Hargreaves
Elena Cotos
Elena Domínguez
Elena Golubovskaya
Elena Martín-Monje
Elena Volodina
Elena Zaitseva
Eleonora Esposito
Eleonora Rossi
Elisabet Arnó-Macià
Elisabet Pladevall-Ballester
Elisabet Titik Murtisari
Elizabeth M. Kissling
Elizabeth R. Miller
Elke Peters
Ellen Hurst
Elvira Barrios
Emilie Vayre
Emma R. Britton
Emma Trentman
Emrah Cinkara
Enikō Csomay
Erasmos Charamba
Erhan Aslan
Eri Kurniawan
Eric Araka
Eric Hauser
Ericson Alieto
Erin N. O'Reilly
Esma Çolak
Esmat Babaii
Esther Lee
Ettien Koffi
Eunjeong Choi
Eva Jakupčević
Eva Wilden
Evgeny Chukharev-Hudilainen
Faezeh Mehrang
Faisal Mustafa
Faisal S. Al-Maamari
Faizah Mohamad Nor
Fan Fang
Fatemeh Soltanpour
Fatih Gungor
Fatihyavuz
Fatmawati Djafri
Federica Barbieri
Federico Navarro
Feifei Han
Ferit Kılıçkaya
Fien De Smedt
Flora Komlosi-Ferdinand
Folkert Kuiken
Forough Kasiri
Forough Rahimi
François Randour
Frankie Subon
Frederik Cornillie
Fryad Hama Najib Muhammad
Fumiyo Nakatsuhara
Gabriela Portugal
Gaby Semaan
Galina A. Khorokhorina
Gareth Carrol
Gavin Melles
Ge Lan
George E.K. Whitehead
Georgios Neokleous
Gertrud Tarp
Gholam Hassan Khajavy
Glenn S Levine
Glenn Stockwell
Glenn Toh
Gökhan Öztürk

Gonc¸Alo Paiva Dias
Goudarz Alibakhshi
Greg Poarch
Guangwei Hu
Guler Duman
Gulnara K. Gimaletdinova
Gulten Genc
Gunseli Orhon
Guy Cook
Gwojen Hwang
Hacer Hande Uysal
Hadar Netz
Hakan Sarıçam
Hamdan Bin Said
Hamid Marashi
Hangyan Lu
Hanna Sundari
Hanne Roothooft
Harold Castañeda-Peña
Harriet Jisa
Hassan Mohebbi
Hayder Hamdany
Hayriye Kayi-Aydar
Hazita Azman
Heidi Byrnes
Heidi Diefes-Dux
Hend S. Al-Khalifa
Heng-Tsung Danny Huang
Herménégilde Rwantabagu
Hesham Suleiman Alyousef
Hezi Y. Brosh
Hilary Nesi
Hilda Borko
Hiram H. Maxim
Hiwa Weisi
Hlaviso A Motlhaka
Hojjat Jodaei
Holger Gaertner
Honglin Chen
Hooman Saeli
Houssine Soussi
Hsiang-Yung Feng
Huei-Ling Lai
Hugo Horta
Hülya İpek
Hülya Tuncer
Hussein Meihami
Hussein Taha Assaggaf
Hyeyoon Cho
Ian Moodie
Ibrahim Farouck
Icy Lee
I-Jung Chen
Ikhfi Imaniah
Ikuya Aizawa
Imam Suyitno
Iman Alizadeh
Imma Miralpeix
Ines Wilhelm
Ingrida Vaňková
Iosif Keselman
Irina Golubeva
Irina I. Klimova
Irini Mavrou
Isabel Negro
Isabel Verdaguer Clavera
Ismaeil Fazel
Jack Burston
Jagoda Topalov
Jamile Forcelini
Jan L. Plass
Jane Jackson
Jane Sunderland
Janin Roessel
Jaya Nur Iman
Jayakumar Jayaraman
Jayne C. Lammers
Jelena Bobkina
Jennifer Herriman
Jennifer Jenkins
Jenny Waycott
Jeongyeon Park
Jeroen Janssen
Jessica To
Jessie S. Barrot
Jesus M. Carro
Jie Chi Yang
Jihyun Lee
Jijung Yao
Jim Mckinley
Jim Ranalli
Jin Kyoung Hwang
Jinrong Li
Jo Mynard
Joanna Nijakowska
Johanna Kallo
John Heritage
John Sweller
Jonás Fouz-González
Jongbong Lee
Jörg Matthes
José Antonio Mompean
José Goris
José M. Mestre
Joseph A. Foley
Joshua James Zwisler
Joshua Matthews
Joshua Wilson
Judy C.R. Tseng
Juhee Lee
Juhi Kim
Julia Miller
Julia Molinari
Julia Valeiras-Jurado
Julie M. Sykes
Jun Zhao
Jurgen Schmude
Justine Dandy
Ka Long Roy Chan
Kamran Janfeshan
Karen Smith
Karin Aronsson
Karwan Kakabra Kakamad
Katarzyna Ożańska-Ponikwia
Kathrin Karsay
Kathryn E. Chaffee
Kathy Escamilla
Kathy Minhye
Katie Dhingra
Katsuhiro Chiba
Kay Hammond
Kay L. Ritchie
Kazuyoshi Sato
Keigo Niji
Keith C Radley
Kemal Yildirim
Ken Hyland
Kenia Lara Da Silva
Kevin Ford
Kevin M. Wong
Kevin Wai Ho Yung
Kevin Werbach
Khaled Moradi
Kia Gluschkoff
Konstantinos Alexakos
Kornwipa Poonpon
Krista Uibu
Kristina Turner
Kukuh Dwi Pamuji
Kundharu Saddhono
Kutay Uzun
Laili Etika Rahmawati
Lama Nassif
Lanchukorn Sriwimon
Lara Bryfonski
Lara Lomicka
Laura Gurney
Laura Hamman-Ortiz
Lauri Linask
Leanna M. Archambault
Lee Jin Choi
Lee Mccallum
Leena Eskola
Leila Dobakhti
Leslie Rush
Liang Li
Lieselotte Sippel
Lijing Shi

Lilia Raitskaya
Liljana Skopinskaja
Ling Gan
Linlin Xu
Liss Kerstin Sylvén
Long Nguyen
Lorena Llosa
Lori Mckee
Lourdes Ortega
Luis Fernando Cuartas Alvarez
Luke Plonsky
Lulus Irawati
Lydia Mitits
Mable Chan
Mack C. Shelley
Made Suastra
Magdalena Trinder
Mahdi Alosh
Mahnoosh Hamavandi
Mahsa Alinasab
Mahsa Kazempour
Majid Elahi Shirvan
Mandie B. Dunn
Mandy R. Menke
Manex Aguirrezabal Zabaleta
Manjet Kaur Mehar Singh
Manuel Palomo-Duarte
Marcela Ruiz-Funes
Marco Cancino
Marcos Garcia
Marcus Callies
Margaret Robertson
Maria Adelina Ianos
Maria Alfredo Moreira
María Ángeles Hidalgo
Maria Del Mar Suárez
María Fernández Agüero
María Gimena San Martín
María José Luzón
Maria Kaltsa
Maria Lebedeva
María Martínez Lirola
Maria Pilar Agustín-Llach
Maria Victoria Lopez Perez
Marie Luce Bourguet
Marija Mijušković
Marina A. Kosycheva
Marion Heron
Marisa Mendonça Carneiro
Mariusz Kruk
Mark D. Johnson
Marta Kopinska
Martin East
Martin H. Daumiller
Mary Frances Litzler
Maryam Asoodar
Maryam Nasseri
Masakazu Iino
Masanori Yamada
Masoumeh Bahman
Matthew Ryan Lavery
Matthias Wilde
Mauricio Véliz-Campos
Maxim Mozgovoy
Md Al Amin
Mehdi Riazi
Mehrak Rahimi
Mehrbakhsh Nilashi
Mehrdad Vasheghani Farahani
Mei Lee Ng
Meihua Liu
Melanie Jones Gast
Meliha R. Simsek
Melinda Reichelt
Melissa Bond
Melissa M. Patchan
Melody M. Terras
Melvin Jebaraj C
Mercedes Querol-Julian
Michael Higgins
Michiko Toyama
Mieke Van Houtte
Mikyung Kim Wolf
Milada Walková
Milica Savić
Milica Vuković Stamatović
Min Jung Jee
Mine Gündüz Kartal
Minoo Alemi
Miriam A. Locher
Miriam Taverniers
Mishima Masakazu
Misook Kim
Mohamad Yahya Abdullah
Mohamed Ridha Ben Maad
Mohammad Ali Nasrollahi
Mohammad Amin Mozaheb
Mohammad Amiryousefi
Mohammad Hamed Hoomanfard
Mohammad Mosiur Rahman
Mohammad Reza Ebrahimi
Mohammad Zohrabi
Mohammadreza Nasr
Mohammed Ali Mohsen
Mona Tabatabaee-Yazdi
Mónica Lourenço
Monique Volman
Moses Kopong Tokan
Moshe Machlev
Mostafa Janebi Enayat
Mostafa Morady Moghaddam
Mostafa Papi
Mostafa Taghizadeh Langari
Muhammad Waleed Shehzad
Murat Balkis
Mustafa Dolmaci
Mustafa Zülküf Altan
Nadežda Silaški
Nadia Abid
Nancy H. Hornberger
Nantje Stiensmeier-Pelster
Nasser Fallah Nasser Fallah
Natalia Dolgova
Natalia Judith Laso Martin
Natalia Popova
Nawal Fadhil Abbas
Nebojša Vladisavljević
Ngee Thai Yap
Nguyen Gia Viet
Ni Nyoman Padmadewi
Nicholas A. Wright
Nicola Fraschini
Nicolas Gillet
Nicole Deh´E
Nigel A. Caplan
Nik Aloesnita Nik Mohd Alwi
Nikol Rummel
Nina Daskalovska
Nina Lazarević
Nina Vyatkina
Ninke Stukker
Noermanzah Noermanzah
Nooreen Noordin
Nuanpan Piengkes
Nurdan Atamtürk
Oğuzhan Atabek
Oi Yee Kwong
Oksana Vorobel
Olcay Sert
Olexandra Kostenko
Olga Bernikova
Olga Dontcheva-Navratilova
Omid Noroozi
Omid Saed
Orphée De Clercq
Ozkan Eren
Pamela Gunning
Pamela Humphreys
Paola Vettorel
Parisa Abdolrezapour
Patricia Bou-Franch
Pattamawan Jimarkon Zilli
Paul Ashwin
Paul Baker
Paul H. Thibodeau
Paul Kei Matsuda
Paul Leeming
Paula Andrea Echeverri-Sucerquia

Paula J. Liendo
Paula López Rúa
Paula Rautionaho
Pengchong Zhang
Peter D. Wiens
Péter Medgyes
Peter Sayer
Peter Seow
Peter Siemund
Phirunkhana Phichiensathien
Phu Hung Bui
Pia Resnik
Pia Sundqvist
Pilar Moreno-Crespo
Pramela Krish N Krishnasamy
Putu Mas Dewantara
Qing Shao
Quan-Hoang Vuong
Rachael Ruegg
Rachel Kraut
Rachel Shanks
Rachelle Esterhazy
Raees Calafato
Rafael Forteza Fernandez
Ragnhild Elisabeth Lund
Rana Namouz
Randy Appel
Raphaela Porsch
Raquel Sanchez Ruiz
Rasoul Mohammad Hosseinpur
Rastislav Metruk
Rebecca Callahan
Reza Kafipour
Reza Rezvani
Richard Kern
Richard Watson Todd
Rick Vogel
Rita Anita Forssten
Rob Waring
Robert Reynolds
Rod Roscoe
Roddy Theobald
Rodrigo Aragão
Ronny Meyer
Rosa Lorés-Sanz
Rosa M. Jiménez Catalán
Rose K. Hendricks
Roslyn Appleby
Rosukhon Swatevacharkul
Roswati Abdul Rashid
Ruilin Wu
Ruiying Niu
Ryan Spring
Saadiya Wudaa Alquraishya
Sabina Valente
Sabine Glock
Sadia Anjum
Safnil Arsyad
Salikoko S. Mufwene
Sally Wiggins
Salvador Montaner-Villalba
Samane Naderi
Samantha Curle
Sandra I. Musanti
Sara Hillman
Sarah Benesch
Sarah Stein
Saskia Van Viegen
Scott Brown
Scott Grant
Scott Sterling
Scott Thornbury
Seftirina Evina Sinambela
Senad Bećirović
Senkamalam Periyasamy Dhanavel
Seran Dogançay-Aktuna
Serge Bibauw
Sergio Mérida-López
Seyed Reza Dashtestani
Seyeed Mohammad Reza Adel
Seyyed Ehsan Golparvar
Shahab Moradkhani
Shenghai Dai
Shengyu Fan
Shim Lew
Shima Ghahari
Shivasakthy Manivasakan
Siân Preece
Siavash Zokaeieh
Sidney Martin
Sílvia Melo-Pfeifer
Silvia Silvia
Sin Wang Chong
Singhanat Nomnian
Siowai Lo
Siros Izadpanah
Somchai Watcharapunyawong
Sompatu Vungthong
Sonca Vo
Sonia López-Serrano
Stacey Brockman
Stacey Wallen
Stefan D. Keller
Stefan Rathert
Stefan Th. Gries
Stefanie A. Wind
Stephanie Link
Stéphanie Roussel
Stephanie Rudwick
Stephen Flusberg
Steven Coats
Stuart Webb
Suchada Sanonguthai
Sue Garton
Sukardi Weda
Suleyman Davut Göker
Sunaina Shenoy
Sungmook Choi
Susan Gass
Susan Goldin-Meadow
Susan Marandi
Suzanne R. Jongman
Svenja Vieluf
Svetlana A. Dubinko
Svetlana Georgieva Dimitrova-Gyuzeleva
Syafi'Ul Anam
Takayuki Nakanishi
Talip Gonulal
Tanju Deveci
Tássia Galvã
Tatjana Đurović
Teresa Guasch
Terry Lamb
Tessa Mearns
Tetsuo Harada
Thi Hoang Hoa Chau
Thi Nhu Ngoc Truong
Thi Thuy Loan Nguyen
Thomas Hoffmann
Thuan Pham
Tian Luo
Timothy Lewis
Tineke Brunfaut
Tobias Bernaisch
Todd Reynolds
Touran Ahour
Tove Larsson
Trude Heift
Tuba Yilmaz
Tuçe Öztürk Karata
Turan Paker
Tzipora Rakedzon
Uludag Pakize
Urarat Parnrod
Urip Sulistiyo
Van Dat Tran
Vera Leier
Veronica Gracia-Ibáñez
Vicent Beltran-Planques
Victoria Tuzlukova
Vildana Dubravac
Violeta Jurkovič
Virginia H.Y. Kwok
Vivian Yenika-Agbaw
Viviana Cortes
Vivien Xiaowei Zhou
Volkmar Paul Engerer
Wahiza Wahi

Wendy Green
Wendy Reinke
Wendy S. Francis
Wenhsien Yang
Will Baker
William S. Pearson
Winai Wongsurawat
Xabier San Isidro
Xi Lian
Xiangdong Li
Xiaofei Lu
Xing Wei
Xiqin Liu
Xue Zhang
Xuyan Qiu
Y. Fan
Yalun Zhou
Ye He
Yeji Han
Yeşim Sevinç
Yew Jin Lee
Yi Song
Yi-Chen Chen
Yi-Hsin Chen
Ying-Chun Shih
Yo Hamada
Yucel Yilmaz
Yudhi Arifani
Yunjuan He
Yutaka Fujieda
Yu-Ting Kao
Yves Bestgen
Zachary Ward
Zahra Akbari
Zahra Fakher Ajabshir
Zahra Shurooq Fakhir Abdul
Zdena Kralova
Zhe Wang
Zoe Kuehn

Contents