**Original Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 1-6*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2022.2268

# "I Treat Them All the Same": Reflecting on Classroom Communicative and Interactional Competence

Thomas S.C. Farrell[1]

[1]Professor in Applied Linguistics, Brock University, Canada

**ARTICLE HISTORY**

Received: 27 August 2022
Revised: 27 October 2022
Accepted: 24 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: tfarrell@brocku.ca

**ABSTRACT**

I am delighted to be invited to write a paper for the inaugural edition of the journal *Applied Linguistics Inquiry*. I am also honored to be invited to serve on its editorial board. In this paper, I will first describe the seminal study reported by Biggs and Edwards (1991) to serve as a backdrop for a discussion of classroom communicative and interactional competence. The paper outlines how teachers can reflect on such competence by exploring teacher questions, teacher feedback, grouping and nonverbal communication, and classroom communicative and interactional competence. The paper ends with a discussion of evidence-based reflective practice as professional development as well as the place of emotions in reflection.

**KEYWORDS:** Reflection; Communicative competence; Interactional competence

## "I Treat Them All the Same"

I still remember the day, many years ago, when I saw the provocative title "I Treat Them All the Same" (Biggs & Edwards, 1991). It prompted me to ask myself: "Do I treat them all the same?" It was the first time I really looked at *all* my students and tried to explore how I set up communications and interactions in my classes. Until then, I took everything for granted and assumed that *all* my students fully understood what was required of them during activities and in my classroom in general. Then I had this sudden realization that all may not be as routine as I had imagined, and I may even have been unwittingly blocking opportunities for my students to learn with the communication patterns I had established or has allowed to be established in my classroom. I realized that I would need to explore and reflect on the communication and interactional patterns in my classroom and see if these are really what I had wanted and were providing opportunities rather than blocking them for my students to learn.

The seminal Biggs and Edwards (1991) study that triggered all these reflections actually took place in England and examined the interactions of five teachers (all white, majority teachers) working within multiethnic classes (of mostly Panjabi children, the largest minority group). Specifically, the researchers were interested in looking into the underachievement of ethnic minority children (EMC) by trying to identify patterns of language behavior that may have placed them at some disadvantage and that could be mediated through language. Generally, they discovered that all five teachers interacted more frequently with majority children than with the EMC in the following categories: general overall interactions, exchanges with students lasting more than 30 seconds, and indeed, most discussions on tasks.

**Cite this article:** Farrell, T.S.C. (2023). "I Treat Them All the Same": Reflecting on classroom communicative and interactional competence. *Applied Linguistics Inquiry*, *1*(1), 1-6. doi: 10.22077/ali.2022.2268

In addition, Biggs and Edwards (1991) discovered that the EMC adopted a topic-associating-style approach to classroom communication, while the majority children adopted a topic-centered style. The teachers found it easier to relate to the majority children because they could expand classroom communications directly through the comments and questions that they used in everyday speech, be it in the classroom or in their homes. However, the same teachers could not understand how the EMC children communicated. This misunderstanding could have stemmed from the fact that these children all came from a cultural background that included informal or experiential learning, usually characterized by little direct verbal interaction, in which skills tended to be acquired through observation and imitation and where immediate verbal feedback (including evaluation and criticism) was rare. In addition, there was little pressure for systematic testing at various stages in the learning process within their cultural backgrounds.

Consequently, I realized that there is an important need for *all* teachers and particularly language to be able to recognize, understand, and reflect on their practice and especially how communication patterns influence (positively or negatively) their students' learning (Farrell, 2015, 2018a,b, 2019, 2021, 2022a). Indeed, language teachers, because they have all of the power and authority to direct classroom interactions, control for the most part who interacts with who, and this unique status in the classroom allows the teacher to arrange interactions that facilitate (or block) learning. Language teachers can provide maximum opportunities for language learning, as Sert (2019) noted, when interactional practices are fine tuned to pedagogical goals, and thus teachers, he notes, should become more aware of the importance of classroom interaction in relation to learning. Thus, language teachers should reflect on how patterns of communication and interaction are set up in their classes, and how these patterns of communication either provide or block opportunities for *all* students to learn and such patterns are present in teacher questions, teacher feedback, grouping and nonverbal communication. All of these eventually lead to classroom communicative competence.

## Reflecting On Teacher Questions

One of the most common ways in which teachers communicate with their students during class is by asking (and answering) questions. Teachers use questioning at the beginning of their classes to establish who controls the interaction, what the topic is for that class, and who is expected to speak. Many teachers also use questions during the course of the class to constantly check their students' understanding of the particular concepts they are teaching at that time. In fact, teachers use questioning as the most frequent means of communication in their classrooms. Forrestal (1990), for example, discovered that almost 60% of the total time a teacher talks in class involves the use of questioning of some sort, and most questions teachers ask their students are those to which the teacher already knows the answers (sometimes called *display-type questions*, such as 2+2—of course, the answer is 4). Walsh (2015) notes that display question, although they can be useful for eliciting responses, checking understanding and/or guide learners towards a 'required' response, they can also shut down learning because they promote more mechanical classroom interactions. Thus, he suggests that teachers ask more genuine questions during lessons, and these are called *referential-type questions*. Such type of open-ended questions according to Walsh (2015), leads to more natural discussions and responses, but ultimately it all depends on the lesson objectives of each lesson. For example, many grammar-focused lessons use display questions to check for quick understanding that the students are following, and for discussions, more referential type questions are more appropriate to generate more practice speaking. If asked though, many teachers have no idea how many questions they ask in each class or what type of questions they favor in their classes, or indeed the function of questions in their lessons. This is problematic because unless teachers become more aware of how and why they use questioning during their classes, this strategy is unlikely to be an effective aid to instruction.

We must also consider what happens after the teacher asks the question and how long he or she waits for a student to answer is also an important reflection for teachers. Good questioning behavior requires allowing students sufficient time to think about and to respond to questions. Rowe (1974) reported that the teachers she observed waited less than *one second* before calling on someone to respond. Furthermore, even after calling on a student, they waited only about *one second* for a response. Such teacher behavior does not make sense because teachers minimize the value of their questions by failing to give students time to think. In language teaching Walsh (2015) notes that wait-time is a very important tool for many teachers, and he suggests this tool should be consciously considered to involve more students in classroom interactions. Walsh (2015) notes that extended wait-time n language lessons can increase the number of learner responses and can also result in more complex answers from learners. Indeed, wait-time becomes even more important in online teaching environments where students will need even more time to provide responses because of the lack of instantaneous interactions that are the norm in face-to-face lessons. As such, language teachers in online environments will have to be even more strategic when asking questions and waiting for their students to respond.

## Reflecting On Teacher Feedback

Feedback is a prominent feature of all classroom communications, and giving feedback is very important for language teachers (Wong & Waring, 2009). By providing feedback to students, teachers are generally communicating one of two things: the teachers like the response, or they do not like the response. In other words, feedback provides students with a measure of their

current progress and tells them whether or not they need to improve in any way. Regardless of their age group, students are very savvy; they quickly learn how to read a teacher and the particular ways in which he or she provides feedback.

Generally speaking, when people provide any kind of feedback in communication, they are giving the speaker the information that they are listening to what is being said, and it can be verbal (such as "OK" or "Really!") or nonverbal (such as a nodding of the head or a smile). This process of giving and receiving feedback is ongoing and can be positive (such as in the examples mentioned previously) or negative; when negative (such as a frown), the intent is to change the behavior or direction of the communication. Not many teachers, however, realize that *OK* can have many different meanings, yet it is the most frequent comment teachers use after student responses in class (Fanselow, 1992). Similarly, a teacher responding with 'very good' can have an unintended effect of closing opportunities for learning because it can signal that the topic has ended and there will be no more discussion (Wong & Waring, 2009). Thus, Wong and Waring (2009) suggest that feedback symbols such as 'very good' be used sparingly, and teachers accept learner's correct responses in less evaluative ways such as saying 'alwright'. If learners provide incorrect answers, they maintain that teachers can ask them 'are you sure' to delay or ask for repetition or pursuing with questions such as: 'Why do you say that?'.

Teachers provide feedback to students with some different purposes in mind, such as providing *information* for both teachers and students, providing *advice* for students, providing students with *motivation* and providing feedback that can lead to *student autonomy* (Lewis, 2002). Feedback is one way for teachers to tell their students what they are doing well and what they need to improve on. Feedback from students tells teachers how both individual students and the class as a whole are doing so that they can adjust their instruction accordingly. Thus, feedback provides an ongoing form of evaluation for both teachers (of their teaching) and students (of their progress), which is in addition to the information provided by the regular end-of-term grading system. Connected to the valuable information that feedback provides for both teachers and students is the follow-up of specifically advising students as to how they can improve their learning if it is deemed a problem. Teachers can follow up their feedback with suggestions for specific learning strategies that the students should incorporate into their learning; if possible, the teacher should model these strategies rather than just explain them. Depending on how it is presented, feedback can provide motivation for students by encouraging them to stretch their abilities to the fullest.

## Reflecting On Grouping

Whole-class grouping is probably the most common classroom learner arrangement—the students usually sit in rows, and the teacher is located at the front of the room. This type of learner grouping has both positive and negative effects on student learning. From a positive perspective, whole-class learning may actually be a more efficient way of instruction in situations where teachers have to deal with very large classes and have a limited time to teach a specific curriculum. Indeed, when the whole class is together, it can promote a sense of security among slower learners because they can rely on the faster learners to provide group answers until the slower learners catch up. Whole class grouping also has some negative effects on student learning.

Harmer (1995: 243) called whole-class grouping "lockstep" learning and explained: "All the students are 'locked into' the same rhythm and pace, the same activity . . . the traditional teaching situation, in other words, where a teacher-controlled session is taking place." In this arrangement, the teacher controls all of the communication and the class, regardless of how many students are in the room, is seen as only one group. In such an arrangement, students have few opportunities to interact with the teacher or with each other because of the amount of control the teacher asserts, not to mention that the seating arrangement (in rows) makes interaction unnatural. When teachers use a whole class grouping arrangement, they are assuming (although I do not think most teachers are consciously aware of these assumptions) that *all* students proceed at the same learning pace; however, we also have students who are slower and have different learning styles and use different learning strategies than the main group. All this may be lost and hidden to the teacher because he or she is probably monitoring the whole group and not individual student reactions.

Another type of interactions that are popular in classrooms are small group or pair work learner arrangements. Such arrangements mean that teachers encourage *learner autonomy* and *collaborative learning* in their classrooms (Farrell & Jacobs, 2020). Learner autonomy involves "learners being aware of their own ways of learning, so as to utilize their strengths and work on their weaknesses" (Farrell & Jacobs, 2020: 10–11). Pair work provides opportunities for learner autonomy in that our students can learn how to access the quality of their own work within the pair or group; this lessens the threat of assessment being based solely on teacher evaluations. Thus, students working in pairs and groups can become self-dependent and self-motivated to learn because the teacher is no longer in total control and no longer solely responsible for student learning. In addition, in order to enhance learning in pairs, students need assistance from their peers that would not be available in whole-class activities, and this also calls for more collaboration with these same peers. Collaborative learning, also known as cooperative learning, consists of learners engaging in group activities that enhance student-student interaction. For this to happen, the teacher must support the pairs rather than abandon them with the hope that meaningful dialogue will just happen if you put the students together.

## Reflecting On Non-Verbal Communication

So far, we have only focused on spoken communications in the classroom and how teachers can reflect on and assess these communications to see if they are providing or blocking opportunities for student learning. Of course, there is another type of communication that occurs in all classrooms is nonverbal communication. As Stevick (1982: 163) put it, "If verbal communication is the pen which spells out details, nonverbal communication provides the surface on which the words are written and against which they must be interpreted." It is important for teachers to be able to control their own nonverbal communication and to be able to read their students' nonverbal signals in the classroom. Nonverbal communication can be focused mainly on kinesics and proxemics. Kinesics deals with gestures, posture, touching behaviors, facial expressions, and eye behaviors. Proxemics deals with space and seating arrangements and examines who interacts and responds more verbally and non-verbally in classrooms and where the teacher is placed during these interactions. We must be cautious when interpreting the meaning of nonverbal behaviors because they may have multiple meanings (as do many of the words we use) depending on where they are used (context), how they are used, and who uses them. For the latter, a detailed knowledge of the person may be required before we can conclude the meaning of the nonverbal behaviors he or she may use.

## Classroom Communicative and Interactional Competence

Some language teachers assume that all the participants in a classroom know how to communicate and interact and that their interactions are mostly smooth. In other words, we teachers assume that all our students have some kind of classroom communicative competence. Johnson (1995: 160) defines classroom communicative competence (CCC) as, "students' knowledge of and competence in the structural, functional, social, and interactional norms that govern classroom communication." Johnson (1995) further suggests that teachers must *define* their students CCC, *establish* it and *extend* it. It is very important for the students to be able to understand established patterns of classroom communications so that they will be able to follow what the teacher expects from them.

In terms of communication interactional competence (CIC), Walsh (2013), suggests it is "teachers' and learners' ability to use interaction as a tool for mediating and assisting learning" (Walsh 2013: 130). As Walsh (2015) notes, when teachers can extend CIC, they provide more learning opportunities for their students. When students know what they are supposed to do each day without having to spend too much time working this out each day, they can better focus what they are supposed to be learning. This focus on classroom interactional competence (CIC) was also highlighted by Walsh (2013) when he maintained that teachers and learners need to develop this if they are to work effectively together.

## Evidence-Based Reflective Practice

The previous sections on exploring and reflecting on various aspects of classroom communication and interaction introduces the need for language teachers to engage in reflective practice as part of their professional development and is based on the belief that teachers can improve their understanding of their own teaching by consciously and systematically reflecting on their teaching experiences (Farrell, 2018a). The key to reflecting on classroom communication patterns in the manner they have been addressed in this paper is that teachers must gather concrete data (evidence) about classroom communications to make informed decisions about their teaching.

As language teachers the only real concrete evidence we have that a lesson has occurred is a recording and transcription of the communication that represents the moment-to-moment communications between the teacher and students and between students themselves that occurred during the lesson. As Walsh (2015) notes, we can only get a real understanding of the complexities of classroom communication and interaction when we have a precise representation of what is really occurring. The most important type of precise or concrete classroom communication data a teacher should collect is a recording of the communications and a record of this recording in the form of a written classroom transcript. If we rely on our memory of classroom communications and events, we may miss some important data because we all have selective memories. Recordings and transcriptions are the best concrete evidence we teachers can get about our work. We can collect this type of concrete data by placing an audio recorder or video recorder in our classroom. If students break up for group or pair work, place the audio recorder in the middle of one of the groups because it may be impossible to record what each group is saying.

Once the data has been collected, the teacher then needs to transcribe the part of the recording that relates to the focus of the investigation. For example, teachers can transcribe only parts of their lesson, such as the opening or the closing, each time they give instructions, each time they ask a question, or whatever part of the lesson they are interested in. After this transcribing, the teacher can analyze and interpret the transcript. After making interpretations about the communications that exist in their classes, teachers can decide if these are the types of communications and interactions that facilitate learning. In the following sections, I discuss the interconnected ways teachers can engage in professional development through reflective practice.

### *Engage in Group Discussions*

One way to begin the reflection process is to get a group of interested teachers to come together to talk about their teaching—especially the communication patterns they see existing in their classes. All the group members have to be equally responsible for keeping the group on track, so the group should negotiate when, where, and how often they want to meet. They should also negotiate an agenda for each meeting and distribute responsibilities evenly between the members (Farrell, 2014).

### *Engage in Classroom Observations/Discussions*

The group of teachers can decide to engage in classroom observations along with audio or video recording their classes so that they can see and hear the exact interactions and communications that take place in specific classes and gauge their effect on learning. Observation can be carried out alone, as in self-observation, pairs (as in critical friendships) can observe each other's classes, or the group can try to observe each member's classes in turn. That said, I suggest that classroom observations should start with the teachers looking at their own classroom communication patterns and interactions. For example, participants can tape their own classes and transcribe the parts of the tape that they are interested in investigating. In this way, the teachers can develop more confidence in describing their own teaching to others (especially to parents and administrators) because they have specific evidence in the form of recordings and transcripts, and they can also bring these to their group discussions with other teachers (Farrell, 2011).

### *Engage in Journal Writing*

Journal writing can also be carried out alone in the form of a diary, in pairs writing to and for each other, or in the group writing to and for each other. I suggest that teaching journals provide teachers with a written record (evidence) of various aspects of their practice, such as classroom events and interaction, and allow teachers to step back for a moment to reflect on these issues (Farrell, 2013). When teachers write regularly in a teaching journal, they can accumulate information that on later review, interpretation, and reflection can assist them in gaining a deeper understanding of the types of communication and interaction that occur in their classes. These journals can then be shared with the members of the teacher group, and the other members can comment orally or in writing. In addition, the group may want to collaborate to write a group journal with all members taking turns adding excerpts about classroom communication and interaction. This type of collaboration may raise more questions about important issues concerning aspects of communication and interaction that may not normally occur if writing a journal alone. Of course, there are many ways teachers can engage in professional development other than the three outlined above, but I have found that these three (group discussions, classroom observations, and journal writing) are a good beginning for teachers and they can be the most productive.

Finally, for most teachers, professional life is hectic as their day begins well before they enter their classrooms, and never really ends given the endless preparation, planning, and grading that they must engage in. Thus, teachers must be wary of being paralyzed as a result of reflecting on their classrooms and as such should also consider the emotional aspect of engaging in all of the activities outlined in this paper. For example, Stanley (1998: 587) has cautioned language teachers that when engage in reflective teaching, they may have some "emotional reactions to what is uncovered through investigation." Because emotions are said to be a the 'core' (Holmes, 2010: 147) of reflective practice in the context of teaching, attending to this affective side of teacher reflections, can help develop a greater awareness and understanding of their emotions, and language teachers should thus be emotionally ready to face what they may discover after they begin their reflections (Farrell, 2022b).

## Conclusion

This paper outlined a seminal study by Biggs and Edwards (1991) that served as a backdrop for a discussion of classroom communicative and interactional competence. The paper outlined how language teachers can reflect on such competence by exploring teacher questions, teacher feedback, grouping and nonverbal communication, and classroom communicative and interactional competence. The paper then discussed the idea of language teachers engaging in evidence-based reflective practice that entails teachers recording, transcribing and interpreting the results as part of their professional development. The paper ended with a discussion of reflective practice as professional development that encouraged teachers to reflect in groups, and/or, classroom observations, and/or journal writing. Regardless of which reflective tools teachers use, I agree with Walsh (2015) who noted that by reflecting on their own classroom communications and interactions, language teachers can not only improve their professional practice but also provide more learning opportunities for their students. In addition, all language teachers should be. In addition, just as the act of teaching is not lacking emotions, so too emotions cannot be detached from the teacher who is reflecting (the reflector), because the moment feelings behind events and behaviors are revealed, thus all teachers will reflect through their personal emotional lens (Farrell, 2022b).

## References

Biggs, A. P., & Edwards, V. (1991). "I treat them all the same": Teacher-pupil talk in multiethnic classrooms. *Language and Education, 3*, 161–176.

Fanselow, J. (1992). *Contrasting conversations: Activities for exploring our beliefs and teaching practices*. New York: Longman.

Farrell, T.S.C. (2011). Keeping SCORE: Classroom observations as reflective practice. *RELC Journal, 42*, 265-272

Farrell, T.S.C. (2013). *Reflective writing for language teachers.* Sheffield, England: Equinox Publishing Ltd.

Farrell, T.S.C. (2014). *Reflective Practice in ESL Teacher Development Groups: From Practices to Principles*. Basingstoke. UK: Palgrave Macmillan.

Farrell, T.S.C. (2015). *Promoting teacher reflection in second language education: A framework for TESOL professionals*. New York: Routledge.

Farrell, T.S.C. (2018a). *Reflective language teaching: Practical applications for TESOL teachers.* Bloomsbury Publishing.

Farrell, T.S.C. (2018b). *Research on reflective practice in TESOL*. Routledge.

Farrell, T.S.C. (2019). *Reflective Practice in ELT. Sheffield*. England: Equinox Publishing Ltd.

Farrell, T.S.C. (2021). *TESOL Teacher Education: A Reflective Approach*. Edinburgh, UK: Edinburgh University Press.

Farrell, T.S.C. (2022a). *Reflective language teaching*. Cambridge, UK: Cambridge University Press

Farrell, T.S.C. (2022b) 'I felt a sense of panic, disorientation and frustration all at the same time': the important role of emotions in reflective practice, *Reflective Practice, 23*, 3, 382-393

Farrell, TS.C. & Jacobs, G (2020). *Essentials For Successful English Language Teaching* (2nd Edition). London, UK: Bloomsbury.

Forrestal, P. (1990). Talking: Toward classroom action. In M. Brubacher, R. Payne, & K. Rickett (Eds.), *Perspectives on small group learning: Theory and practice* (pp. 158–167). Ontario, Canada: Rubicon.

Harmer, J. (1995). Taming the big "I": Teacher performance and student satisfaction. *ELT Journal, 49*, 4, 337–345.

Holmes, M. (2010). The emotionalization of reflexivity. *Sociology, 44*, 1, 139–154.

Johnson, K. (1995). *Understanding communication in second language classrooms*. New York: Cambridge University Press.

Lewis, M. (2002). *Giving feedback in language classes*. Singapore: RELC.

Rowe, M. B. (1974). Wait-time and rewards as instructional variables, their influence in language, logic and fate control. Part 1: wait time. *Journal of Research in Science Teaching, 11*, 81–94.

Sert, O. (2019). Classroom Interaction and Language Teacher Education. In S. Walsh and S. Mann (Eds) *The Routledge Handbook of English Language Teacher Education* (pp. 216-238). London: Routledge.

Stanley, C. (1998). A framework for teacher reflectivity. *TESOL Quarterly, 32*, 3, 584-591

Stevick, E. (1982). *Teaching and learning languages*. New York: Cambridge University Press.

Walsh, S (2013) *Classroom Discourse and Teacher Development*. Edinburgh: Edinburgh University Press.

Walsh, S. (2015). *Classroom interaction for language teachers*. Alexandria, VA: TESOL Press.

Wong, J. & Waring, H. (2009). 'Very good' as a teacher response. *ELT Journal, 63*, 3, 195-203.

**Original Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 7-17*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2022.5561.1003

# A Snippet of an Ongoing Narrative:

# A Non-linear, Fragmented, and Unorthodox Autoethnographic Conversation

Bedrettin Yazan[1]
Ufuk Keleş[2]

[1]Associate Professor, Department of Bicultural-Bilingual Studies, University of Texas at San Antonio, U.S.A.
[2]Assistant Professor, ELT Department, School of Educational Sciences, Bahcesehir University, Turkey

**ARTICLE HISTORY**

Received: 18 August 2022
Revised: 18 November 2022
Accepted: 27 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: bedrettin.yazan@utsa.edu

**Abstract**

In this unorthodox autoethnographic study, we present a dialog between us - two transnational scholars. Throughout the manuscript, we explore several issues, which instantaneously came out during our unstructured, simultaneous, and casual conversations. We first discuss how our dichotomous relationships based on respect as an advisee and an advisor; a mentee and a mentor; and a student and a teacher have transformed into a sincerity-based friendship over time owing to our shared interest in autoethnography. We then move on to a discussion of our beliefs/thoughts/emotions about "home" in accordance with our lived experiences as transnational scholars. We scrutinize how inhabiting, knowing about, and becoming in academia complicated our understanding of where, or more importantly, what "home" means for us. Afterwards, we talk about the affordances and challenges of autoethnographic discourse agreeing that it requires us to practice vulnerability in order for us and our readers to benefit from the therapeutic effect of autoethnography. Throughout the manuscript, we also discuss how using both singular and plural first-person voice provide us with the opportunity to maintain our individual voices in an interpersonal and collaborative relationship while achieving a multivocal tone. We hope that our discussion extends with our readers' critique of, negotiation with, participation in, and/or resistance to our beliefs/thoughts/emotions as stated in our conversation.

**Keywords:** Transnational identity; Mentorship; Friendship; Autoethnography; Multivocality; Home

## Prologue

In this rather unorthodox autoethnographic dialog, we (Bedrettin and Ufuk) take the liberty to free-write about our "academic friendship" built on and around our advisee/advisor, mentee/mentor, and student/teacher relationships… Although we do not deny the existence of these labels in our collective narrative, we refrain from positioning each other dichotomously, and simply define our relationship as a friendship which is more nuanced, equity-based, and sincere. This friendship is built on respect and sincerity we feel towards each other as fellow applied linguists, educational researchers, teacher educators, teachers, and so on. Frankly speaking, I (Ufuk) personally feel that my level of respect to Bedrettin and his work is (and should be) higher than his respect to mine given that he is an older timer in academia, who has produced more work than me, and who gladly shared his

**Cite this article:** Yazan, B., & Keleş, U. (2023). A snippet of an ongoing narrative: A non-linear, fragmented, and unorthodox autoethnographic conversation. *Applied Linguistics Inquiry*, *1*(1), 7-15. doi: 10.22077/ali.2022.5561.1003

experiences with me to support my endeavors first as a doctoral candidate and later as an early career scholar. Yet, our relationship is built more on friendliness than respect.

I see Bedrettin as a friend rather than an advisor/mentor. It is perhaps because he has never imposed his older timer authority on me, and always trusted me in my (post)doctoral work. And I think I have never let him down seriously - at least to my knowledge… In addition to his substantial help with my dissertation, we have worked on multiple publication projects. In all stages of these projects, from designing the study to writing the paper, we brought to the table our time, efforts, knowledge, motivation, and beliefs/thoughts/emotions as equals. I believe we have learned so much from each other as "writing buddies."

Similar to our academic friendship, we have built a close friendship in our private life. We help each other whenever we need a friend's help. We give a ride to each other or pick up from the airport from time to time. For us, our long rides are a perfect opportunity to catch up academically, professionally, and personally. When I was a PhD student, we bought each other lunch (mostly he did, by the way, for obvious reasons). He helped me move into my apartment; I helped him move out of his. We exchanged furniture when we no longer needed them. Supporting each other as friends, in time, strengthened the bonds between us.

Interestingly, despite our close relationships, our interest in autoethnography grew almost separately from each other. Bedrettin had already published an autoethnographic paper in a renowned journal in our field. At the same time, I was introduced to and immediately fascinated by autoethnography while taking an introductory qualitative research course from another professor. By the time Bedrettin used autoethnography as a pedagogical tool in teacher education (Yazan, 2019a) while teaching a graduate course that I took, I had already started reading into autoethnography. Later, when I told him that I was planning to write an autoethnographic dissertation, he immediately stated his support. As my mentor, however, he cautioned me that I would be the first to do that in the university's school of education, so it could be a challenge to persuade other professors to be in my dissertation committee. He also said that it could be difficult to find a job in the US with an autoethnographic dissertation. Nevertheless, I decided to write my dissertation in an autoethnographic format (Keleş, 2020).

Long story short, this manuscript is about our friendship. Yes. But it is not simply that. It is also about how knowing, doing, and living autoethnography have helped us make sense out of our interpersonal relationships along with our mutual experiences in academia that go back as far as 2015 when Bedrettin (an early career assistant professor at the University of Alabama) and I (a doctoral student on a Fulbright grant at the University of Alabama) started negotiating my future PhD plans. We believe that our conversation below may shed light on how autoethnography may help our readers make meaning of their lived experiences (as advisees and advisors; mentees and mentors; newcomers and older timers; students and teachers; and most importantly as two academic friends). Doing so, they may introspect into and redefine their current interpersonal relationships so that they can anticipate and plan their future trajectories.

***

I (Bedrettin) also wanted to add to this wonderfully written prologue. In the title, we have all those adjectives to describe our conversation, but I wanted to unpack "autoethnographic," especially the *ethno* dimension of such an approach. As you said in an earlier paper of yours, autoethnographers frequently refer to the morphological constituents of the term: *auto* (self), *ethno* (culture), and *graphy* (narration) to frame their autoethnographic work (Keleş, 2022a, 2022b). In other words, autoethnography is "writing about the personal and its relationship to culture" (Ellis, 2004, p. 37). Similarly, to me, "the act of autoethnography writing is a concentrated and profound experience of identity negotiation" through self-narrative (Yazan, 2019b). In this narrative, we mostly focused on ourselves, which means that the *auto* component has been the main aspect of our autoethnographic component.

Although we don't explicitly mention which cultures are being examined and critiqued in our conversation, I believe the reader will engage in a critical reflection about cultures/discourses that surround our experiences and identities, as they listen to us talk. For example, when we discuss our transnational identities and experiences, we'd be critiquing how dominant discourses of nation-states operate to construct and maintain the ideological and physical borders. We share how our identities do not fit within those borders.

I also wanted to direct attention to our multivocality individually and collectively in this manuscript. I'd like to apply multivocality to describe our writing at two levels. On one level, like Bakhtin (1986) argues, "any utterance is a link in a very complexly organized chain of other utterances" (p. 69). What Ufuk and I say here is a reflection or a snippet of our ongoing conversation with colleagues and friends and with each other. We don't construct knowledge in isolation. On another level, again relying on Bakhtin, we negotiate and interrelate multiple voices or I-positions (Akkerman & Meijer, 2011) as we construct and enact our identity in writing or narrative. Engaging in the following conversation, we select certain I-positions to foreground (and concomitantly not select others) in our narrative, which demonstrates multiple voices that make who we are as writers.

***

So, here comes our dialog… *Vira bismillah!* [Godspeed!]

***

***Bedrettin***: When we first started this written conversation, we didn't have a prologue, which is why I started like this: "I know we didn't provide any introduction to this conversation. The reader will find it a little odd to start reading like this but hang in there for a second please." Ufuk, we talked about a potential conversational piece, but we never went for it. I told you about Hossein's invitation to contribute to the *Applied Linguistics Inquiry* journal and we had a chat which again didn't materialize. I just heard from Hossein again and he's wondering if we could write an article for the inaugural issue of the journal. Without telling you (i.e. a Facebook message), I thought "why not start this and get it going as organically as possible." So here I am! Sitting on the balcony of my parents' house in Keşan, waiting for the upcoming meetings with colleagues in half an hour. I am wondering about what question can kindle a conversation between you and me. I know we talked about writing a piece on autoethnography (*critical autoethnography, well, every autoethnography should be critical and political already, right?*), but I am starting this without letting you know. I am aware you're grading your student papers and I'll see you in a few days in person. I think I'll start with a question on how we got interested in autoethnography. You know we have this urge in academic writing, going all the way back to the origins or the beginning of the story. I guess that's what I'm doing here. What do you think? Should we go ahead and write this piece? (*I know all this paragraph will go to waste if you disagree and writing this sentence might be a little nudge or pressure.*). I'll let Hossein (the editor of *Applied Linguistics Inquiry*) accordingly and learn about word count, timeline, and everything. As you write back, I'll keep the conversation going or feel free to do so by asking me questions. I think honest writing (see Casanave, 2017) is a key to connect with our reader.

***Ufuk***: [*Two days later*] Sorry for the late reply, hocam. I have been juggling with the end-of-semester paperwork at Sabancı University (my soon-to-be-ex workplace) and the recruitment documents for Bahçeşehir (University). Reading the conversation starter of yours above… well hocam… I had complicated thoughts; frankly… There is so much to do and so little time to do… I remember we talked about this and we had a rough plan - that was so exciting! Since we both know that we both love to do autoethnography, I said to myself "Hell yeah! Why not?" but a couple of seconds later… I remembered that I did not have time for it! I don't have time for many things these days… Yet, I was so excited… Another meaningful project… I need to make time for it no matter what? [*to be continued*]

[*Three hours later*] I scribbled down the sentences above while waiting for my car to be serviced… Sometimes, I believe we forget that there is life - real life - personal life - family life outside academia… Each time I do something non-academic, my inner voice tells me that I am wasting my time. And yet, my car needs being serviced; my son needs taking out to the ice-cream vendor; my friends to be called; house to be cleaned, and all. And it is a pity that what phd gave me is this scar - the feeling of wasting time if not working on a paper, conference, course and whatsoever. (to be continued - again)

I had to meet the school director to talk about my leave. Now, I actually need to take a rest as today has been so exhausting, yet I could not because… [*to be continued - as always*]

***

[*After two hours*] In the midst of academic smog that surrounds me and penetrates in me as long as I breathe, talking to you hocam is an oasis in the desert. Because you understand what I go through mentally, socially, and emotionally, or to put it in another way; I have "you" to pour out what I have in "me" … uncensored. Also, you experience similar thoughts, beliefs, and relationships as I do. Being at your parents' house… being in Keşan, Edirne, Turkey, Thrace… I have to ask you: Do you feel at home? Is Keşan your home? Where is "home?" Or let me put it more directly: Is academia your home?

For me, academia is sometimes home - and some other times it is what keeps me away from home? Perhaps, we should start by understanding that… to what extent academia is home for us… Us - being transnational individuals who are indebted to academia for their transnational identities in the first place…

***Bedrettin***: [*A few days later*] Wonderful to hear you're interested in writing this piece with me. I just emailed Hossein and let him know that we're working on a manuscript to submit to their inaugural issue. Cc'ed you as well.

Hearing from you, speaking with you, is always refreshing. İyi ki varsın! (*Ayrı yazılan 'ki' bu :) Ah, years of learning Turkish grammar.*)

Two days ago, I just scanned what you wrote. I didn't have the chance to write a response, but your question about home stuck with me. I was working on a different manuscript. Today, I just read your thoughts and reflections and started writing back.

Late reply is totally fine! So glad to hear you're finally transitioning to your new academic job. Best of luck in your transition! You'll do wonderfully at Bahcesehir! They're very lucky to have you. Yet, I can only imagine how much paperwork it requires when both leaving a university job and starting a new one!

Your writing made me think how much and often we forget that we're humans! We're humans first. Yes, we're also teachers, researchers, faculty, academics etc. [insert all the identities/roles here]. So including our stories, personal life in our writing is meaningful. Stories, honest, vulnerable ones humanize us (*thinking about Spry's (2011) "practiced vulnerability" here*). You're right, there's life beyond academia. There's a whole big life which we, for some reason, were trained to hide or allowed to mention in certain contexts or moments to a specific degree. Separation of the personal and the professional! Right? Autoethnography is an effort towards that direction, i.e. getting rid of that artificial, forced separation. So is research on identity in applied linguistics. Well, it depends on the lens you take, but understanding identity at the intersection of all different dimensions of self, all 'I' positions or voices (Bakhtin, 1986; Akkerman & Meijer, 2011), we can open up spaces to humanize pedagogy and research.

As you can imagine, I left your question about 'home' to the end. I said much before I could get to it. Between paragraphs above, I've talked to friends in person and on the phone and they'd ask how I feel visiting Turkey after six years. Honestly, I don't think any place would feel home anymore to me. I knew Kesan wouldn't when I was flying over here. It's the nostalgized home, maybe. Some people, though, feel like home. Being around them, sharing experiences, being vulnerable around … Finding people who are also not fitting in the nation-state border ideologies … I find such people in academia too. So instead of specific places, lands, cultures, I think people, spaces, relationships, conversations might feel home.

You made me think actually and articulate what I think about academia, to some extent. Academia isn't a homogeneous entity, although we tend to frame it that way. Academia this, academia that. We're academia. We make it up and out, in every possible sense. We invent it, construct it, reconstruct it, maintain it, change it. We're in it, why don't we try to change it the way it 'should' be. Well, when there's 'should,' there's ideology involved. We also keep forgetting we researchers have ideologies. We assign value, utility, and power to certain types or ways of being, becoming, doing in academia. Within the imagined borders of 'international academia,' I've found people, spaces, conversations, and relationships which feel home indeed. Also, like you said, academia is the reason we became transnationals and when we realize we are transnationals, with all its benefits and struggles, it's always a challenge to truly 'fit in' a place or call somewhere 'home' again (*well, what does fitting in mean anyways? I gotta stop questioning every possible word or concept, right?*). Maybe we should question the concept of home which has traditionally been attached to a physical place and led to lots of troubles and tribulations across centuries. Oh well, I'll pause here, but I've been meaning to share these two pieces of advice (*in an unsolicited way, though* :) which I stumbled upon when I was writing on autoethnography recently and they resonated with me. Muncey (2010) calls them "two important mantras" and goes on to explain: "[F]irst, 'writers are people who write' – not special people with hidden magical powers but people who take up a pen, or sit down at a keyboard, and write; arguably the rest is practice and having something to say" (p. 57). (*Do I have something to say? I've been practicing for a while and am willing to keep doing so.*) "The second important refrain is 'don't explain and don't complain'. This means setting down the words or drawing the picture without a constant evaluator in your head" (p. 57). That second one is a tough one, indeed! It'll take some time to ignore that evaluator to be able to write freely. We must have constructed that evaluator over the years, through our training and 'feedback' on our writing. I don't mean to say that the training and feedback we received weren't educative, though.

Interesting note here. The first and the only photo you and I took was on this day (June 29) in 2020, exactly two years ago, when I was packing to move from Tuscaloosa, Alabama to San Antonio, Texas. I bumped into that photo while I was writing the paragraph above when it popped up on my Google photos. How serendipitous? (*Am I trying to assign extra value to this writing by bringing in coincidences?*) [*I'll pause here to hear from you.*]

***Ufuk***: [*Next day, while drinking his morning coffee*] Wow! You said so much! So, I think my response will be long as well - accordingly. First of all, let me respond to your "concern" about assigning meaning to every single word we utter. Well, hocam… This conversation requires us to be hypersensitive, right? After all, we are not talking about our holiday plans, or a hobby you took up recently, or a movie we went to last night… We are talking about hardcore stuff… So, I guess, it is OK to overthink, well, as much as it is OK to write freely, right? As long as we enjoy it… I am perfectly fine with assigning additional meaning to "anything." What are "double quotation marks" for, huh?

Another point to consider is why we shy away from giving "unsolicited" advice? Although it is something people may frown upon in the US, it is considered a sign of valuing the person you are talking to in Turkey. In the States, it is like: "Who are you to - or how dare you - teach me a lesson? Are you any better than me? What makes you think that I need your advice?" In Turkey, it is more or less something like: "I had this problem about this and that. I value you so much that I do not want you to go through anything similar to that even if it is not very likely that you will. Yet, if you do, this is the solution - or at least it worked for me. Consider it if you have to." So, hocam, it is OK to give me unsolicited advice. At the end of the day, we share similar cultural blueprints. If you give me a piece of advice, I know that you do so because I am valued not that I am weak.

You know why I embrace my transnational identity so much? Well, hocam, it offers me so many alternatives to choose from. Political correctness… Do I like it? Certainly do. Personal space? I definitely need it from time to time… These are not the concepts I grew up with when I was a child. Yet, here I am. I believe political correctness is the right action to do - well most of the time… especially when I meet someone for the first time. Personal space? Just as I like spending time with people

I love and respect, taking a break and having some private space keeps my mind in my head. And I love it when people respect that!

You see, hocam, I grew up in a rather large family with close relationships with my relatives. So I can say that my community's well-being was more important than whether each member had personal space or not. We did not have to be politically correct, as well because we had these feudal bonds that entitled us to speak directly at each other - well of course keeping our respect for the elders. However, when I first learned about political correctness, it fascinated me - people you have just got to know are being very nice to you. Great!

Now I have new options: I can choose to be and act like an individual or behave like an organic part of the group. Thanks to transnationalism, there are more colors and shades on my pallet. And it is not limited to the Western ways of life…

Ubuntu, for instance. Well, I am fascinated by it. It looks a lot like "*imece*" in Anatolia. *I help you now because I know that you will gladly be helping me when I need it.* None of the words in English explains it… collaboration? Maybe… cohabitation? Well maybe… cooperation? Yeah, why not? But still… *imece* is more than that? It is sacred, communal… it is the "co" itself and yet more than that… Ubuntu? I am sure there is a whole bunch of significance behind the sentiment of the word… although I do not fully understand it, I appreciate ubuntu because I know *imece* and I know how individualism may detriment human relationships. What I mean is that being a transnational, I can cherry-pick the customs that I am exposed to… I can drop some habits that used to feel so natural to me… I can fuse things together… Great, isn't it?

Going back to academia… Your response to my question whether academia was "home" for you, made me ponder even deeper. Initially, I asked you the question because I oftentimes feel estranged from my family members in that they have no idea what I actually do for a living. My mother, after I came back from the US, asked me how my job was any different than a teacher's. Well… It is not, huh? On the other hand, being a teacher and an academician are a totally different line of work. Yet, I was unable to explain the difference since my lexicon did not have any simple words to compare the complexities of both jobs. After I came back from the US, I told my mother that I was now a doctor. The second I uttered the word "doctor," I regretted it. She was petrified as I had never worked at a hospital.

Sometimes, I believe my wife does not really understand my job either. And honestly, I do not think that she is interested in it. Of course, she knows that I am an assistant professor, I work at a private university's ELT department, and I conduct some research. However, she knows nothing about my research interests, the courses I offer, how many papers I have published so far. She is not interested in listening and honestly, I do not feel like telling her about the details. So, when I am with her, I am a totally different person. I downgrade my job to being a "lecturer" only - I am like one of those professors whom she took courses with when she was a student at university. For my mother, I am someone between a teacher and a university rector - as she hears the word "rector" quite often on the news. With my friends from outside of academia, I do not talk about translanguaging, pedagogy of the oppressed, second language socialization, or anything that fascinates me as a scholar. If I feel the urge to tell them something about my job, I have to start from the basics. "Well, there is this guy, Pablo Freire, who wrote "Pedagogy of the Oppressed" (1972). He was an educator with some 'leftist' ideas and 'Catholic' faith. He said…" Now what? If I explain to my friend who Freire is and how his work informs mine… and then what. So what? Not that I do not value my family and friends… I just do not think they are invested in listening to some "boring" details of my intellectual, academic, professional, idealist, critical, problematic, emotional and so on life.

An academician's home is their heart and mind, I guess (!?!?!). If you do not share a home, how can you be a family? That is, hocam, what has been bothering me for a while. In order to be - and keep being - a real family, we have to co-exist in "our" home both mentally and emotionally. Yet, with some of my friends, my family members, and even some of my colleagues, I do not feel that we are sharing a home. The problem is… I oftentimes find myself inhabiting their home - not the other way around. And when I am at their home, I feel comfortable only as much as a guest can get.

When the fact that I have been transforming into someone else, the gap between my be(com)ing a transnational person and my monocultural family members drifts me away from home. For them, defining home is easy. The only country they have lived in so far… To me, home is not a physical space anymore. It is more like a network of relationships, each of which requires me to assume, construct, and enact a different identity. Among the many identities I have, the loneliest one is my academic identity.

***Bedrettin***: [*About a week later*] It'd be only fair to the reader to share (to help them contextualize the writing here) that I'm writing this response from the US. Since the last time I worked on this paper, I had a four-hour drive with you from my hometown to the airport, had Covid in San Antonio, finished summer class, and had a Zoom meeting with you about our ongoing research project. And one of those projects is this current paper or writing. I think both of us agree that we can't name this writing we've been engaged in for the last two months. I don't want Hossein, the founding editor of the journal, to think that this conversation here is less than academic writing. I tend to be the harshest critique of myself and my work. (*Gotta remind myself of Muncey's (2010) mantras again*) However, we want to wrap this conversation up soon so that we can submit the manuscript to the editor which is due in a few days.

Before I continue reflecting on my transnational identity and what it matters to me as a teacher educator, researcher, and university faculty, I'd like to summarize our discussion about this very writing when we met last week. Initially, I thought this title might capture what we're trying to accomplish here: "A snippet of an ongoing conversation" but I'm totally fine if you'd like to go ahead and revise it as you wish. I think you suggested a revision which I couldn't recall right now. Also, we didn't have a traditional introduction paragraph for this paper, nor a conclusion, both of which I believe we should add when we're done talking or when there's a good stopping point for both of us.

When I first suggest we write this conversational piece, I had just finished a manuscript on autoethnography and gotten excited about doing something to nudge the academic writing conventions a little. I think there's something in autoethnography that inspires the reader to write or at least engage in a deep reflection. I can tell my earlier writing at the top of this paper seems to be coming from a deeper reflection than now. I feel like I should've have completed the whole paper when I was in that zone of reflective depth sitting on my parents' balcony watching the mulberry tree branches dancing with the breeze of Mediterranean Turkey, making it difficult for the crows to take another bite from the mulberries.

Why is our conversation supposed to be important to the readership of *Applied Linguistics Inquiry*? Well, I know I don't have to rationalize or justify everything I say. But we need to articulate the purpose of this paper, right? I think one of my writing goals is to exercise and "practice vulnerability" (Spry, 2011) and to unmask the distanced voice of the applied linguistics researcher. I remember Ellis and Bochner's (2006) description of vulnerable author which is "a feeling and vulnerable actor" (p. 441), as a complete opposite of "I" as "a disembodied authorial academic voice that argues and tries to persuade" (p. 441). Perhaps, it's something I expected the applied linguistics researcher to sound like or it's my attempt to establish a relationship with the reader or I'm tired of writing in that 'scientific' third person voice, concealing my true self behind the words which are supposedly mine. Isn't it contradictory? Isn't it unfair to the reader to poke my head a little from the authorial window, but expect them to make sense of what I say?

Additionally, I think I actually wanted to share a snippet from our ongoing conversation on being a transnational individual/researcher in applied linguistics who is trying to do autoethnographic research by swimming against the current at times. We've been talking about such issues for a while as we have conducted research. We've had a lot of collaborative reflective, critical conversations around identity, being, belonging, in a new sociopolitical context. I honestly wonder what conversations my colleagues are having in different contexts across the world. I'd like to hear or read what they say which doesn't get in their published papers or which gets lost within the academic writing conventions. That's also probably why I started to write such a paper.

Plus, what Ahmed et al. (2021) said about the collective "we" voice in research and importance of "transparent, dialogic voices":

> most scholarly publications with multiple authors appear as one voice speaking in triumphal consensus and unfolding in predictable patterns that enhance the validity of the ultimate truths claimed. Yet, it is interesting to reflect on the hidden deliberations and silences that may arise when multiple authors are involved, and when power differentials between senior/junior contributors or gendered and racialized collaborators shape the final work. (p. 540)

Ahmed et al. (2021) are discussing their voice in their trio-ethnography, and I wanted us to try this voice out and practice transparent and dialogic writing with you since we're interested in creating similar autoethnographic voices in our scholarship. Do you think such a goal is something we're accomplishing in this paper? How do you feel about our dialogic writing here?

***Ufuk***: [Several hours later] Hocam, "A snippet of an ongoing narrative" is a perfect title for our dialog here. Maybe, adding a subheading would make it clearer for those who read it. So, how about "A snippet of an ongoing narrative: A non-linear, fragmented, and unorthodox autoethnographic conversation?" This way, we could prepare our readers for a collective "stream of consciousness[1]." You know how much I like using this "literary" technique - writing intermittently in short fragments without following traditional written discourses which have an introduction, development, and a conclusion. I mean… why bother organizing our beliefs, thoughts, and emotions in a logical order when we want to focus on the complexities and chaotic fabric of our thinking/feeling processes? Let us unleash our inner voice as is. After all, this is our life, not a fiction, huh?

Years ago, I read it somewhere, perhaps in a book by a feminist literary critic that Western written academic discourse was very similar to fictional prose in that they were both phallocentric and followed a conventional plot: introduction (exposition); development (rising action, climax, and falling action); and conclusion (resolution). These were highly similar to the five steps of the masculine pleasure: desire (or lust should we say), arousal, ejaculation, relief, and relaxation. This author also noted that while pen was a symbolic representation of penis, women would rely on their lips - their oral skills to tell a

[1] Stream of consciousness is a literary narrative technique by which the writer aims at expressing the character's thoughts through direct quotations of the mind instead of creating meaningful statements using cohesive devices. Similar to inner monologues and the spoken word, this technique may be applied exclusively throughout a whole book or section of a book, or intermittently in short fragments (Bowling, 1950).

story. For women, pleasure would be more fragmented and repetitive yet contextual… So… If we regard this as a conversation rather than a traditional "text," it is safe to take up a fragmented and repetitive yet contextual style despite the fact that we both identify as cismales. Just like women had to educate themselves in male ways for male gaze, we can try something opposite, huh? This paper does not have a linear organization, but it is as it is - we depend on our stream of consciousness, the words slipping through our lips, and sound unorganized and fragmented yet meaningful in context.

For the "we" language, it totally makes sense. I mean, given that you were/are my dissertation advisor, an old(er)timer, and more importantly, my mentor, using one voice would most probably lead to the silencing of my voice. For instance… let's say that I have a more pro-feminist approach to academia than you… please tell me if you disagree… Perhaps, you would never bring forward the "penis vs. lips" metaphor above. But here we are. I can freely express "my" own beliefs/thoughts/emotions in ways that I find suitable. You, on the other hand, hocam, may not agree with me. Nevertheless, we are still "friends, huh? … no hurt feelings? We still maintain our relationships intact as a mentor and a mentee; as a teacher and a student, as an advisor and an advisee; and so on. In this conversation of "ours," I keep my voice fearless of what you will say. We do not have to agree, huh? Or at least, we are able to agree to disagree. I am happy to have you (t)here - this way.

Although some of your scholarship is built on identity theories, you know, I am not a big fan of the word "identity" as you know I find it rather broad, quite post-structuralist, and fairly blurred. Unlike many other advisors, you never questioned why I refrained from the concept of identity in my own work. You always respected that I am more invested in critical approaches and view language learning as a process of socialization. We have never discussed this before… I know you pay extra attention to not step on anybody's toes, yet you had every right to do so since I was your advisee, mentee, and student. I think modesty is part of your "identity" ... LOL :) … And I believe it is this modesty that helps me feel relaxed in this conversation of "ours." I think what keeps our bonds strong even long after I completed my dissertation is the healthy "we" relationship that we have maintained successfully. When we disagree, I respect your "I" and vice versa. When we are together, we serve as complements to each other's beliefs/thoughts/emotions.

At this point, we have to tell our readers that, as autoethnographers, you and I agree that opting for first person voice helps us to diverge "radically from the analytic, third-person spectator voice of traditional social science prose" (Bochner & Ellis, 2016, p. 82). Unlike traditional social scientists who write in third person voice to distance themselves from their readers (Adams et al., 2015), we refrain from assuming a 'God's eye' omniscient view in this conversation and write in first person voice in a dialogic tone.

As for the question why our conversation is supposed to be important to the readership of *Applied Linguistics Inquiry*… Agreeing that "practicing vulnerability" is one important reason for us both, I will also add that I find this joint autoethnography therapeutic in the sense that sharing "the intimacies of my world" (Sparkes, 1996, p. 467) with you in this dialog helps me better articulate my life challenges, which would otherwise be left as unspoken anxieties, hollowness, and ambiguities. When I write to you (or speak to you for that matter), I remember that I am not alone in this. There is at least someone out there who listens, who knows, who appreciates, who supports me and my scholarship. Also, when we write, we want our script to make sense. Writing about unwanted feelings/thoughts/beliefs is, therefore, a means of making sense of them (Ellis et al., 2011). As autoethnographers, we are both aware that practicing vulnerability is not easy, yet without acknowledging and articulating the fragility of our narratives, it would be almost impossible to heal our souls, recover from our wounds, and boost our immunity (Keleş, 2023a). Once we are brave enough to relive and narrate some unpleasant experiences, which cause emotional distress, we can enjoy the positive therapeutic effect of autoethnography.

The therapeutic effect of autoethnography… So, what is it to do with the readers? Well, first of all, writing about a phenomenon, an experience, or an incident means that we articulate it by describing it in detail, defining it using our own experiences, and/or explaining it with examples (Keleş, 2023b). When our readers go over our manuscript, they may be able to make sense of similar experiences they go through by comparing them with ours (Ellis et al., 2011). They may also feel that they are not alone. As a result, reading into our conversation may allow the readers to engage with us on academic, professional, and more importantly humanized and humanizing levels and may help them make sense of their own - or we can at least hope for that.

HERE I WANT TO SHOUT OUTLOUD:

Hey! If you feel that going into academia has distanced you from your family members, you are not alone…

Hey! If you find it difficult to explain what you do for a living to your friends, we are with you…

Hey! If you have sacrificed more than others could ever imagine… only to become a scholar/researcher… Well, welcome to the club!

Hey! If you are confused about where home is, we feel you… You are home with us…

***

Hocam, although there is so much to talk about in this "free style" text, we may think of finishing it for the sake of our readers. What do you say? If you would like to add more, I am more than willing to continue. But… I think we do not have much space left. I would like to remind you that we were planning to add a prologue and an epilogue depending on the journal's readership. So, how do you think we should proceed?

***Bedrettin***: [*The next morning*] I totally agree we should wrap up and let the editor and reviewers take a look at this writing. However, couple of brief notes I feel like I need to make: first, yes, I wouldn't think of the feminist critique of the traditional 'scientific' writing conventions! I've always appreciated your background in gender studies which has been the leading theoretical lens in our earlier collaborations as well. Second, the adjective 'therapeutic' to describe the doing, writing, and reading of autoethnography is a great one. Can't agree more! I think you mentioned that when we were discussing this paper last week. Third, the "we" voice in qualitative research, when not dialogic, (*which I'm still finding myself use with colleagues; not that I've completely transformed my collaborative 'we' voice to a dialogic one in my entire writing endeavors*) has this pressure to find ways to converge everyone's ideas/perspectives/worldviews. This pressure, at times and perhaps in a lot of times, makes it really difficult to collaborate in a research writing experience in a true academic sense in which every team member would be pushing each other's thinking. I'm not saying that all team members should be dissenting from each other at all times, but there tends to be less space for divergence in academic collaborations or (*gotta rephrase a little*) there tends be less space for it to be reflected in the written products or reports of the collaborative academic research. I can keep talking about this matter for a while, but another pressure, i.e., traditional article word count, is reminding me that I'm running out of space.

## Epilogue

Ufuk asked me (Bedrettin) to write the epilogue. I said I'd try my best. What I'll do is share what I feel about this writing at the moment, now that we're closing or wrapping it up. Every time I do similar kind of writing, I have the feelings of gratitude and vulnerability. First of all, I'm grateful to Hossein for giving us the space or the reason to start writing. And I'm grateful to Ufuk for being my sounding board, writing buddy, critical friend, and agreeing to let me engage in this experience, with him, which like he said, has been a therapeutic one. I feel vulnerable which I know we said we're actually doing this writing to practice it, but we're at the same time opening ourselves up for public scrutiny and I wonder how our writing is going to be taken by the reader. Are we going to be able to reach out to the reader as powerfully and intimately as we hoped for? Is what we're discussing here going to resonate with the reader? In an autoethnographic sense, our goal is to strike a chord in the reader. It was like therapy to converse in such a critical fashion with Ufuk, but we don't want this conversation to be our conversation only. We want the reader to respond, chime in, and find something to comment on. (*Well… Yes, yes… For sure… That's an invitation. Yes. That's right!*) We want it to be like therapy for the reader, too. I know explicitly saying that isn't going to guarantee the feeling reaching across the reader, but like we said earlier, we're practicing vulnerability.

## References

Adams, T.E, Holman Jones, S., & Ellis, C. (2015). *Autoethnography*. Oxford University Press.

Ahmed, A.U., Morgan, B., & Maciel, R.F. (2021). Feeling our way: A trioethnography on critical affective literacy for applied linguistics. *Calidoscópio, 19*(4), 538-552.

Akkerman, S.F., & Meijer, P.C. (2011). A dialogical approach to conceptualizing teacher identity. *Teaching and Teacher Education, 27*(2), 308-319.

Bakhtin, M.M. (1986) (V.W. McGee, Trans.). In C. Emerson, & M. Holquist (Eds.), *Speech genres and other late essays*. University of Texas Press.

Bochner, A.P., & Ellis, C. (2016). *Evocative autoethnography: Writing stories and telling lives*. Routledge.

Bowling, L.E. (1950). What is the stream of consciousness technique? *Publications of the Modern Language Association of America, 65*(4), 333-345.

Casanave, C. P. (2017). Representing the self honestly in published research. In J. McKinley & H. Rose (Eds.), *Doing research in applied linguistics* (pp. 235-243). Routledge.

Ellis, C. (2004). *The Ethnographic I: A methodological novel about autoethnography.* AltaMira Press.

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. *Historical social research/Historische sozialforschung*, *36*(4), 273-290.

Freire, P. (1972). *Pedagogy of the oppressed*. 1968. (M. Bergman Ramos, Trans.). Herder.

Keleş, U. (2020). *My language learning, using, and researching stories: critical autoethnography of socialization* (Doctoral dissertation, The University of Alabama).

Keleş, U. (2022a). Autoethnography as a recent methodology in applied linguistics: A methodological review. *The Qualitative Report, (27)*2, 448-474. https://doi.org/10.46743/2160-3715/2022.5131

Keleş, U**.** (2022b). In an effort to write a "good" autoethnography in qualitative educational research: A modest proposal. *The Qualitative Report, 27*(9), 2026-2046. https://doi.org/10.46743/2160-3715/2022.5662

Keleş, U. (2023a). Long-term washback effect of university entrance exam: A critical autoethnography of socialization. In D. Köksal, N. Kavaklı, & S. Arslan (Eds.), *Handbook of research on perspectives in foreign language assessment*. IGI Global.

Keleş, U. (2023b). Exploring my in-betweenness as a growing transnational scholar through poetic autoethnography. In L.J. Pentón Herrera, E. Trinh, & B. Yazan (Eds.), *Doctoral Students' identities and emotional wellbeing in applied linguistics: Autoethnographic accounts*. Routledge. https://doi.org/10.4324/9781003305934-7

Muncey, T. (2010). *Creating autoethnographies*. SAGE.

Sparkes, A.C. (1996). The fatal flaw: A narrative of the fragile body-self. *Qualitative Inquiry*, *2*(4), 463-494. https://doi.org/10.1177/107780049600200405

Spry, T. (2011). Performative autoethnography: Critical embodiments and possibilities. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 497–512). SAGE.

Yazan, B. (2018). TESL teacher educators' professional self-development, identity, and agency. *TESL Canada Journal*, *35*(2), 140-155. https://doi.org/10.18806/tesl.v35i2.1294

Yazan, B. (2019a). Identities and ideologies in a language teacher candidate's autoethnography: Making meaning of storied experience. *TESOL Journal, 10*(4), 1-21. https://doi.org/10.1002/tesj.500

Yazan, B. (2019b). An autoethnography of a language teacher educator. *Teacher Education Quarterly*, *46*(3), 34-56.

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 16-33*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2023.6056.1009

# The Pedagogical Integration of Web 2.0 in K–12 English Teaching: A Systematic Review

Xundan Wang[1]

[1]Faculty of Education, The University of Hong Kong, China

**ARTICLE HISTORY**

Received: 14 August 2022
Revised: 22 October 2022
Accepted: 25 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: u3584698@connect.hku.hk

**ABSTRACT**

Due to the high development of technology and limited research on pedagogical integration of web 2.0 in K-12 English classrooms, the objective of this study is to understand the web 2.0 technology integration in pedagogy in K-12 English classrooms in recent five years, find the gap, and provide insightful suggestions for further improvement. A systematic review with PRISMA 2020 guidance and theme analysis were conducted to achieve the research goal. The findings of the selected paper indicated that current English teachers valued technology-integrated content instruction, preferring to use more relevant and acceptable elements of teaching tools. The ongoing usage of adopting web 2.0 tool not only benefited students' effective language learning but also enriched teachers' teaching methods and improved teaching levels. Additionally, some obstacles were also revealed to web 2.0 tools' successful adoption in education, including the tool, teachers, and external influences. Relevant suggestions are made to improve the pedagogical integration of various tools.

**KEYWORDS:** Digital teaching; K-12 English; Pedagogical integration; Systematic review; Web 2.0

## 1. Introduction

Current schooling has been modified because of technological progress and digitization. It demands new pedagogy in the classroom to fit with contemporary characteristics and enhance the professional development of instructors. Significantly, ICTs, the means for acquiring new-quality information during data conversion (Shugaeva et al., 2021), occur. It has played a growing role in enabling teachers to educate effectively and efficiently and achieve 21st-century education goals (Jimoyiannis et al., 2013). Web 2.0, founded by O'Reilly, is one of the ICT applications that this article focuses on. It acts as a platform for harnessing collective intelligence derived from collective wisdom (O'reilly, 2007). Therefore, instructors can utilize web 2.0 technologies such as Wikis, blogs, websites, and YouTube channels to acquire useful information for their classroom instruction and to share their experience and teaching ideas online. However, even though the concept of web 2.0 technology emerged in 2004, its use in the classroom remains limited (Jimoyiannis et al., 2013). Moreover, governments and educational administrations are progressively encouraging K-12 teachers' technology preparation so that they can successfully apply technology in the classroom (King & South, 2017). However, most teachers are unprepared to use and adapt technology in teaching or learning (Lei, 2009; Ottenbreit-Leftwich et al., 2010).

In the past decade, there has been little systematic research on the pedagogical integration of Web 2.0 into this subject. Meanwhile, the demand for teaching with technology is greater than ever before. Some existing research is obsolete in comparison to the requirements of today. In addition, the increased use and focus on Web 2.0 technology in higher education diminishes the significance of the technology in K-12 education (Norton & Hathaway, 2008). Typically, English

Cite this article: Wang, X. (2023). The pedagogical integration of web 2.0 in K–12 English teaching: A systematic review. *Applied Linguistics Inquiry*, *1*(1), 16-33. doi: 10.22077/ali.2023.6056.1009

teachers are the most inclined to pursue creative measures, such as integrating technology into their classrooms. Researchers should pay more attention to Web 2.0 integration in the K–12 English teaching. Consequently, the aim of this study is to comprehend the amount of the particular technology— web 2.0 integration in K-12 English classroom pedagogy during the past five years, identify the gap, and make suggestions for further improvement. To fulfil the research objectives, the following questions will be addressed:

- What types of web 2.0 technologies are integrated into English instruction?
- What are the benefits of incorporating web 2.0 technologies into English instruction?
- What are the main barriers teachers experience when integrating Web 2.0 into English instruction?

To address the research issues and accomplish the study's objective, the author utilized the PRISMA 2020 declaration of systematic review as a checking and searching guideline to supply a comparatively transparent and exhaustive shifting data procedure. Theme analysis (Braun & Clarke, 2006) was used to classify and analyse the chosen data sample in a systematic way so that the three study goals could be met. The results of this study reveal the actual deployment status of web 2.0 technologies in K-12 English classrooms. Similarly, it can provide more intelligent ideas regarding the pedagogical integration of various tools to ultimately enhance the quality of digital instruction.

## 2. Literature review

### 2.1. Web 2.0 in education

In 1989, Berners-Lee presented the concept of the World Wide Web, a platform where users can read and write on the same device (Carvin, 2005). O'Reilly formally defined Web 2.0 for the first time at a technology-related conference in 2005. He described the considerable changes in web functions and identified various basic competences of web 2.0, particularly with regard to user engagement and writable capabilities (O'Reilly, 2007). Nonetheless, since the notion of Web 2.0 has been introduced and popularized, researchers and academics hold divergent viewpoints on this matter. Abram (2005) argued that Web 2.0 demonstrates interactivity via discussions, personalisation, and interpersonal networking, etc. By stressing the social phenomenon, Barsky (2006) continues to add Web 2.0 traits. Warschauer and Grimes (2007) asserted the actual function of Web 2.0 in society, altering the communicative use of web platforms to introduce fresh energy and version. In addition, the proliferation of Web 2.0 will boost creativity, information exchange, and collaboration (Tu et al., 2008).

Together, these studies emphasize the following basic themes: knowledge creation, information sharing, interaction, and collaboration, with the exception of the various representations of the Web 2.0 definition. Web 2.0 is a read-and-write medium that depends primarily on user engagement and collaboration, as opposed to a read-only tool (Thompson, 2007; Richardson, 2006). Similarly, Downes (2005) supports this claim by explaining how Web 2.0 might encourage users to develop, share, and distribute information. On Web 1.0, a read-only platform, the majority of users passively receive tools without participating actively or interactively. As a result, Web 2.0 has been created to provide a platform for consumer contact and collaboration to compensate for this predicament (Usluel & Mazman, 2009). The usage of Web 2.0 in education represents a dramatic shift from traditional to 21st-century teaching methods. The educational sector's transitional stage is dominated by web 2.0 technologies (Bull & Garofalo, 2006). This remark laid the groundwork for the use of technology aides in the classroom in a mixed capacity. Educators may use Web 2.0 as a platform and tool to revolutionize the teaching and learning process (Alexander & Levine, 2008). Common Web 2.0 technologies include blogs (self-expression), wikis (collaborative content creation), podcasts, and social bookmarking (the social organizing of collective knowledge) (Sykes & Thorne, 2008).

Then, instructors can build a contemporary background and a learning environment in which students can engage and collaborate. Today's language teachers and educators are increasingly incorporating Web 2.0 into foreign language instruction (McLoughlin & Lee, 2007). Simultaneously, Web 2.0 has been recognized as a valuable and conducive application for language teaching and learning. Numerous language instructors and teachers view student participation favourably. Through virtual learning environments and transmissive pedagogy, for example, the incorporation of Web 2.0 into instructional strategies might improve students' performance (Hew & Cheung, 2013). Zeng (2020) stated that web 2.0 might improve language learning in a number of ways, including language input and output, interaction, the language learning environment, and learner autonomy. In addition, Halim and Hashim (2019) argue that prior research demonstrates the clear benefits of web 2.0 technology for ESL learning, such as learning engagement, writing skills, environment, social skills, communication, self-confidence, and peer coaching, among other learners. These findings are roughly compatible with prior research findings about the pedagogical benefits of web 2.0 technology (Alexander, 2006; Brown & Adler, 2008). However, the use of web 2.0 technology in the English classroom will have negative effects on technology relevance, efficacy, adequate teaching resources, and gadgets. It suggests that the majority of educators are optimistic and web-savvy. There is scant mention of the appropriate implementation of web 2.0 technology in schools (Bingimlas, 2017; Pan & Franklin, 2011). Consequently, the systematic evaluation of web 2.0 technology in language schools of the twenty-first century must be ensured and guaranteed.

### 2.2. Web 2.0 technological pedagogical content knowledge (TPACK 2.0)

Insufficient research and analysis have been conducted over the past decade about instructors' adaptation of technology in the teaching context. To increase teachers' pedagogical use of ICT and their knowledge of technology, Mishra, and Koehler (2006) created the TPACK model, which is based on Shulman's (1986) pedagogical content knowledge framework, by adding a new perspective: technology (PCK). The derived model of TPACK consists of seven parts that are both independent and interdependent: (a) content knowledge (requiring educators' knowledge of a particular area), (b) pedagogical knowledge (mastering different teaching strategies and methods), (c) technological knowledge (understanding current widely used technology tools in the educational field), (d) pedagogical content knowledge (knowledge of teaching methods pertinent to subject-matter content), and (e) technological pedagogic content knowledge (knowledge of teaching methods (Teo et al., 2019).

The TPACK concept places greater emphasis on technology and provides a framework for instructors to integrate technology successfully and flexibly into their instruction. The TPACK model can serve as a crucial conceptual and theoretical framework for demonstrating teachers' knowledge of technology in education and guiding teachers to develop their understanding of classroom technology integration (Chai et al., 2010; Mishra & Koehler, 2006; Teo et al., 2019). This model will assist instructors or educators in comprehending the rationale for effective technology adoption in their classrooms and revealing the adaptable TPACK framework (Mishra & Koehler, 2008). Gradually, more teachers will demonstrate their digital savvy and facilitate the use of TPACK in contemporary classroom environments (Borko et al., 2009). Teachers with such expertise will be able to select appropriate teaching resources, access globally connected shared material, and build student-centred activities that correspond with curriculum standards more dynamically and easily (Nelson, 2009).

Thus, teachers will no longer focus on a tool, or the types of activities associated with it, but rather on the actual material that connects with the chosen technology. Input and output of acquired knowledge will be more efficient and effective among students. In a collaborative, authentic, user-friendly, and meaningful environment, students will have more motivation and incentive to acquire knowledge. These will demonstrate the finest practices of teachers and the genuine nature of the pedagogy, content, and technological combination. In the intervening decade, Web 2.0 has been promoted and developed in the education sector due to its consumer-centric, collaborative, and participatory characteristics (Jimoyiannis, 2015). The nature of 21st-century education, the new reformed policy, the digital skills of new generations, and the slow updating of instructors' skills are the four primary components that garner the most growth attention in Web 2.0. (Jimoyiannis, 2015).

In addition, the growth of teachers' professional knowledge of technology indicates a condition of expansion. The whole curriculum design will gradually incorporate internet literacy. Current educators must be able to utilize and integrate Web 2.0 resources to boost student learning. Add 2.0 to the TPACK model to investigate the link between web 2.0 and the TPACK framework. Consequently, Technological Pedagogical Knowledge (TPACK 2.0) is a framework that includes a particular technology-web 2.0-to drive and support instructors' implementation. The original TPACK framework remains unmodified. Web 2.0 will be viewed as a well-defined pedagogical measurement rather than a simple technological tool. For instance, it can direct teachers to employ a variety of teaching methodologies and integrate specific Web 2.0 tools while instructing subjects such as English. It integrates dynamically with the TPACK model to increase the sustainability of technology in education.

In TPACK 2.0, three types of knowledge are required: pedagogical content knowledge (knowledge of specific content), technological content knowledge (knowledge of the application of specific web 2.0 tools for content representation), and technological pedagogical knowledge (knowledge of web 2.0 technology for teaching strategies) (Jimoyiannis et al., 2013). It illustrates the relationship between the technology of web 2.0, content, and pedagogy to integrate the specifical tools into the instructional practice of teachers (Jimoyiannis et al., 2011). These three constituents are interdependent. The overlapping connection can facilitate instructors' strategic educational thinking and the implementation of how to apply their professional knowledge or teaching practices with web 2.0 tools. The TPACK 2.0 framework encourages the incorporation of Web 2.0 in teaching settings (Jimoyiannis, 2015).

## 3. Method

The systematic review was utilized to answer the three specific research questions stated previously: the specific web 2.0 tools and their benefits and drawbacks. It can provide a thorough, evidence-based synthesis of existing knowledge on the issue (Aromataris & Pearson, 2014). The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) 2020 statement was developed as a reporting standard for systematic reviews. In 2009, the PRISMA protocol was published, which provides precise and high-quality reporting requirements for future systematic reviews and meta-analyses (Liberati et al., 2009). The PRISMA 2020, a revised version with 27 checklists and a flowchart of the data collection process, will provide a more complete transparent and exhaustive reporting to combat the inadequately reported systematic review

publications (Page et al., 2021). Identification, screening, eligibility, and inclusion are the four essential processes in gathering data. The following will provide specifics:

### 3.1. Identification phase

The author went through two additional rounds of the identification step to increase the comprehensiveness of the data. In the initial step, the author concentrated on identifying information sources. The author identified four journal publishers, Taylor & Francis, ScienceDirect, Sage, and Wiley, based on the advice and help of professional researchers and significant reference sources for education and social sciences research. The author employed search terms linked to "web 2.0 technology" and "English teaching" (ESL, ELT, EFL, TESOEL, English teaching) to find the research topic. These search keywords were also tweaked and altered to accommodate the search engines used for different Journal articles. The time frame was restricted from 2018 to 2022. The objective was to examine the actual implementation of Web 2.0 over the past few years, particularly following Covid-19. Only open-access, full-text publications were accessible in the original set of 1059 findings. For the second step, the author desired to enrich the data sample and obtain more pertinent papers on the issue by including the ERIC database. The ERIC database was renowned for its comprehensive access to education-specific full-text resources. Even though the ERIC database does not align with the previous three journal publishers, the author modified the selection criteria by selecting only peer-reviewed journal articles and omitting novels that do not correspond to journal articles. The ERIC keyword search terms are shown in table 2. The time frame was also 2018 to 2020, and the complete text was accessible. Twelve hundred and twenty items were chosen from ERIC.

### 3.2. Screening phase

Following the screening process, fifty duplicate articles were eliminated. Inclusion and exclusion criteria were established for selecting more specialized publications. The author examined the titles and abstracts of selected publications to ensure they met the predetermined criteria. English-language articles were included to improve data extraction. For their secondary data, systematic review and meta-analysis were removed from the methodology section. The author wished to collect initial data from selected articles in order to contextualize the setting more thoroughly. Higher education, vocational education, and traditional education were excluded since they did not correspond with the scope of the study. K-12 was covered. Since this study focuses on teachers' viewpoints, student and parent research sets were excluded. To capture instructors' actual implementation levels and attitudes in the classroom, the author included both in-service K-12 English teachers and pre-service English teachers who would teach in the K-12 grades. The subject who met English requirements were included. Ultimately, 78 articles matched the aforementioned criteria.

### 3.3. Eligibility phase

To improve the excellent quality and reliability of this article, the author rechecked each abstract and conduct a full-text review to ensure conformity with the criteria already established (see Table 3). Non-English language, quantitative, non-peer-reviewed, and non-journal publications were excluded. In addition, to improve the comprehensiveness and inclusion of data, references that were highly relevant to the issue and match the inclusion criteria were also included in the selected articles. Fifty-eight articles were omitted following this inspection. These were eliminated from three viewpoints. For instance, the first reason was prioritizing students over teachers. The teaching focus shifted to universities and other forms of higher education rather than K-12. Last but not least, the poor quality of the data source and the vague procedure description was excluded.

### 3.4. Extraction phase

Eventually, the author extracted and included 20 articles which were all based on the previous purification (See Appendix). The whole procedure of the PRSMA from stage 1 to stage was presented in the following (See figure 1).

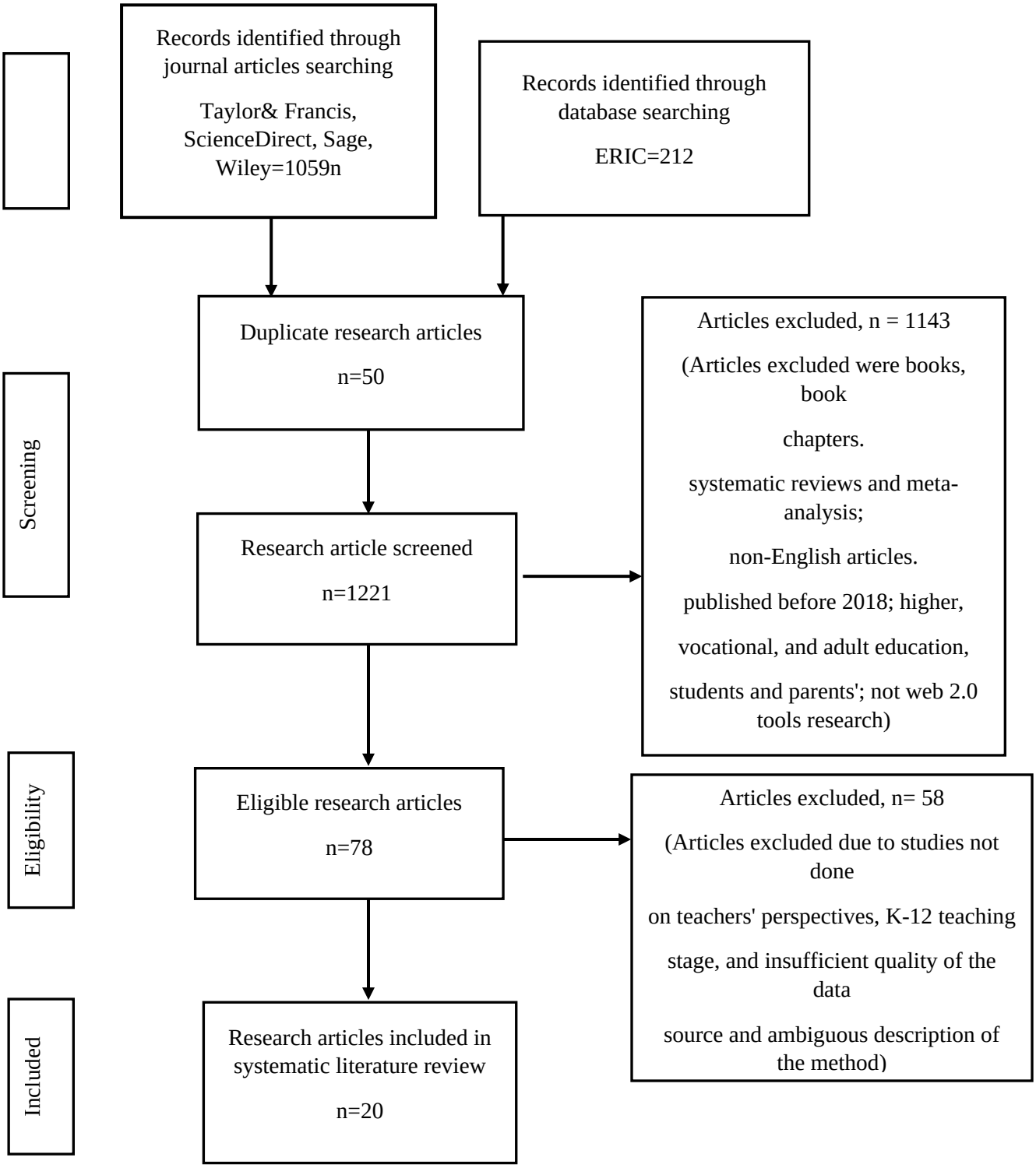



**Figure 1**, The straightforward procedure from stage 1 to stage 4 of the PRISMA.

### 3.5. Data analysis

The selected 20 articles were evaluated and classified by theme analysis (Braun & Clarke, 2006), a qualitative data analysis, after data extraction and full-text reading. In the thematic analysis, an inductive, bottom-up method was utilized to examine the content and theme of the data set. The thematic analysis (TA) is a method for searching and identifying themes systematically in response to specific research topics (DeSantis & Ugarriza, 2000). Primarily, the TA corresponds to the research objective. All detected themes have been subjected to a recursive and repeating procedure in which the author

swings back and forth between the data and coded themes to verify their exhaustiveness, authenticity, and dependability (Braun & Clarke, 2006). As for the analysis procedure, the author went through six processes to do the thematic analysis, including data familiarization, first coding, looking for themes, defining themes, and drafting a report (Braun & Clarke, 2012). In step one, the author engages in "active reading" (Kiteley & Stogdon, 2014) by underlining, taking notes, and making comments. By actively reading, the authors intended to familiarize themselves with the data set and connect it to their complex research concerns. After familiarization, the author carefully began first coding. She categorized the data according to their content significance, and her interpretation was consistent with the research questions. The labeled codes are both descriptive and interpretive. Microsoft Excel was applied to control the coding systems throughout the second step. The spreadsheet has six columns, including articles, aims, methodology, web 2.0 technology, benefits, and challenges about incorporating web 2.0 tools into the classroom. Significantly, the outline of three points: web 2.0 tools, advantages, and problems of integrating web 2.0 in the classroom separately, as opposed to using a single column: significant findings, was intended to more clearly and unambiguously convey salient aspects pertaining to the research questions. When the coding systems were complete, the third step—the search for themes—began. Nonetheless, each must be aligned with the research topics. Then, in step four, reviewing prospective themes, a recursive and iterative procedure was adopted to convey the completeness and reliability of the themes created. By reading articles, the author validated the coding and topics to capture all the data in a relevant manner. After reviewing and double-checking, the author began defining each concept. The designated themes must be informative, succinct, and captivating (Braun & Clarke, 2012). To facilitate comprehension, the author provided subthemes for each theme. The final phase is producing the report, an appropriate process for reporting and analyzing the data in a consistent, logical, and persuasive manner.

## 4. Findings

In this section, the results collected from the twenty papers listed above will be discussed in detail, separately, and methodically to answer the questions. Each topic is consistent with the research questions of the study. Three themes are included: (1) the types of web 2.0 tools integrated into teaching, (2) the advantages of integrating web 2.0 in the English classroom, and (3) the primary hurdles of integrating web 2.0 in the English classroom (See table 1).

Table 1. The emerging subthemes

| Research questions | Themes |
|---|---|
| 1. What are the kinds of technology of web 2.0 integrated in English teaching? | The categories of Web 2.0 tools integrated in teaching:<br>• Multimodal presenting flatforms<br>• Social networking<br>• Interaction<br>• Authentic teaching resources |
| 2. What are the advantages of using web 2.0 technology in English teaching? | Two advantages of integrating web 2.0 in English classroom:<br>• Promote students' effective language learning and positive learning behaviors<br>• Improve teachers' teaching level and enrich teaching methods. |
| 3. What are the main barriers faced by teachers when implementing web 2.0 in English teaching? | Three barriers of integrating web2.0 in English classroom:<br>• The first-order tool barrier<br>• The second-order teacher barriers<br>• The third- order exterior barriers |

### 4.1. Types of web 2.0 tools in ELT

The author viewed 20 articles and ultimately selected 18 articles pertaining to tools. Regarding the categories of specific web 2.0 technologies used in English instruction, the topic is subdivided into four extended topics for a more comprehensive presentation of the subject matter and scope. Multimodal presentations of flatforms, social networking, engagement, and authentic instructional materials comprise the four expanded themes (See Table 2).

Table 2. Types of web 2.0 tools in ELT

| Themes | Sub-themes | Sources |
|---|---|---|
| Multimodal presenting flatforms | YouTube videos<br>Instagram<br>TIKTOK | Eisenlauer, (2020)<br>Fyfield et al., (2020)<br>Karakas & Kartal, (2020)<br>Lee, (2022)<br>Luy, (2022)<br>Mohammad-Salehi et al., (2021) |
| Social networking | Weibo; WeChat<br>Pinterest<br>Blog<br>EnglishForum.com<br>Twitter<br>WhatsApp<br>Facebook<br>Zalo | Bener & Yildiz, (2019)<br>Le et al., (2021)<br>Liu et al., (2020)<br>Mei et al., (2018)<br>Mohammad-Salehi et al., (2021)<br>Santos & Castro, (2021)<br>Say & Yildirim, (2020) |
| Interaction | Kahoot<br>Quizizz<br>Padlet<br>Edmodo<br>Google classroom | Arabaci & Akilli, (2021)<br>Jong & Tan, (2021)<br>Luy, (2022)<br>Santos & Castro, (2021)<br>Say & Yildirim, (2020) |
| Authentic teaching resources | Canva<br>Podcast<br>TED<br>BBC<br>Wiki | Arabaci & Akilli, (2021)<br>Karakas & Kartal, (2020)<br>Özkan & Güler, (2018)<br>Schmidt & Rye, (2020) |

#### 4.1.1.Multimodal presenting flatforms

Multimodality is one of the most common functions utilized in English technological environments. Six research demonstrate the trend. Mohammad-Salehi and his colleagues studied the factors influencing the application of web 2.0 technologies in English instruction by Iranian educators. They collected data on the familiarity of 160 EFL teachers in Iran with adopting web 2.0 technologies and reported numerical results. The most surprising finding is that most teachers are conversant with web 2.0 video sharing platforms such as Instagram and YouTube. Students in English courses typically use the two apps as a

platform for video sharing or presentations. Eisenlauer (2020) similarly concentrated on multimodal meaning-making techniques in English learning situations. YouTube was described as a video-sharing and content-sharing tool that could facilitate access to authentic language settings in numerous ways. Karakas & Kartal (2020) also investigated the Web 2.0 tools utilized by English teachers in the classroom. Therefore, YouTube was one of the popularization tools for visualization that allows students to learn English interactively. Luy (2022) also discovered that YouTube was one of the most widely used video platforms for remote education, especially during the Covid-19 pandemic. YouTube videos could aid English teachers in demonstrating and sharing video information with their pupils. Fyfield and his colleagues focused their investigation on YouTube instructional videos in formal education classrooms in 2020. This study demonstrated the dominance and pervasiveness of YouTube as a video-sharing medium for English teachers to better disseminate instructional content and prepare students for future content. Additionally, Lee (2022) analysed the use of Instagram and TikTok for language acquisition in a separate study. Images, audio, space, emojis, and gestures could convey learning to pupils using the picture or visual presentation tool. The multimodality feature of Instagram and TikTok could vividly convey the precise meaning of the content and facilitate successful language acquisition. According to all research, YouTube is the most popular video-sharing platform in web 2.0 technology utilized by EFL teachers nowadays. Teachers can locate numerous relevant videos on YouTube to share with pupils. Multimodality platforms, including YouTube, Instagram, and TikTok, may dynamically blend text and visuals to convey the understanding and significance of instructional information effectively.

### 4.1.2. Social networking

The second key role of web 2.0 technologies teachers use is social networking, where English teachers engage with one another to share valuable experiences and exchange intelligent teaching-related information. Eight articles discussed various social networking techniques. Blogs are well-known for their social nature and are the most popular tool. Bener and Yildiz (2019) suggested that web 2.0 blogs could be utilized as social networking sites in education. This study highlights the social networking role of Blogs, which provide adequate space for instructors to talk and share their experiences. Mei and his colleagues (2018) discovered that social networking sites like Weibo and WeChat had a reasonably high rate of familiarity and usage compared to other tools. EFL teachers in China primarily utilized Weibo and WeChat as communication platforms where they could communicate with peers or experienced teachers to exchange experiences and severe ideas for improving English teaching standards. Pinterest was employed by Liu and his colleagues (2020) as a research tool to examine social function, professional aims, and educational resources. It indicated that Pinterest could unite teachers worldwide in a socialized knowledge community where they could access and share educational ideas. Subsequently, Mohammad-Salehi and his helpers (2021) investigated teachers' awareness of and use of web 2.0. They discovered that present teachers utilized EnglishFourm sites, Twitter, and blogs as debate and viewpoint exchange platforms. Santos and Castro (2021) also mentioned specific social connection platforms, such as blogs, Facebook, and WhatsApp, allowing education workers to interact with others and receive the updated information. Le and his peers (2021) stressed the importance of social networking platforms for informal professional development. The results demonstrated that Facebook and Zalo were indispensable social networking platforms in casual social settings. Teachers typically used their informal or spare time to seek academic assistance and network. All tools have a crucial feature: social collaboration, in which teachers interact with others, such as peers, educators, and policymakers, to share fresh instructional ideas and connect to worldwide networks.

### 4.1.3. Interaction

Regarding the integration of instruction, teachers typically prioritize interaction medium in the age of technology-assisted instruction in order to facilitate teacher-student collaboration. Through semi-structured interviews and a questionnaire, Luy (2022) investigated the perspectives of fifty primary school teachers on teaching with technology. Teachers cited Kahoot, Quizizz, and Google classroom as interesting and integrative programs. Particularly following the covid-19 epidemic, teachers had increased their use of this technology. Through the incorporation of online resources, the quality of online instruction was enhanced. Jong and Tan (2021) examined Palette, an online learning platform, to evaluate the most effective educational and technological instruments. The results indicated that Padlet was an effective tool for encouraging kids to write, receiving feedback from teachers, and monitoring their writing development. In their 2021 study, Santos and Castro mentioned Edmodo and Google Classroom as two tools for enhancing online learning engagement. Teachers could use these tools to interact with students more effectively, and students could participate more actively in class. Say, and Yildirim (2020) reported that most teachers consider web 2.0 tools as engagement mediums, such as building classroom groups and delivering presentations. Arabaci and Akilli (2021) highlighted the perceptions of English language teachers towards web 2.0 tools in educational environments. Students and teachers could cooperate interactively to produce a more meaningful event, according to most educators' conceptions of Web 2.0 technology. They had a solid propensity for employing web 2.0 tools in their live instruction classroom to improve students' speaking, listening, reading, and writing skills.

### 4.1.4. Authentic teaching resources

Teachers will seek and find several relevant teaching-related resources and materials in instructional contexts. Five articles out of twenty research discussed search tools for practical and appropriate teaching resources. The most well-known resource for accurate instructional materials is Wikipedia. Schmidt and Rye (2020) investigated the free online encyclopedia - Wiki - and discovered that the Wiki, a specialized online resource aimed at educators, could assist them in accessing numerous real-world examples. In addition to Wiki, Ozkan and Güler (2018) investigated the opinions and attitudes of pre-service EFL teachers towards the podcast application. The results demonstrated that podcasts are valuable listening options for educational information. Podcasts could assist teachers in creating a realistic learning environment. In their publications, Arabaci and Akilli (2021) cited Canva, a platform for developing teaching resources typically used by teachers to design more contextualized teaching resources and materials individually or jointly from a student-centered perspective. Karakas & Kartal (2020) also mentioned BBC and TED as search engines for authentic educational resources. Compared to other methods, these two are highly popular for collecting more native language usage and content for students to learn. The previous search and creation tools for natural teaching resources can significantly facilitate and diversify teachers' instruction.

### 4.2. Advantages of integrating web 2.0 in English classroom

According to the study's findings, the benefits of integrating web 2.0 technology into English instruction could be categorized primarily into two groups: students and teachers (See Table 3). A detailed explanation will be provided in the sections that follow.

Table 3. Advantages of integrating web 2.0 in English classroom

| Themes | Sub-themes | Sources |
|---|---|---|
| Promote students' effective language learning and positive learning behaviors | • Language skills: lexical learning; listening; speaking; writing; reading<br>• Learning motivation<br>• Interactive collaboration | Arabaci & Akilli, (2021)<br>Eisenlauer, (2020)<br>Fakhruddin, (2020)<br>Fyfield et al., (2021)<br>Jong & Tan, (2021)<br>Lee, (2022)<br>Luy, (2022)<br>Mei et al., (2018)<br>Özkan & Güler, (2018) |
| Enrich teaching methods and improve teachers' teaching level | • Authentic and diverse language resources<br>• Contextualize and recontextualize learner-oriented content<br>• Professional learning improvement<br>• Positive attitudes<br>• Social collaboration | Arabaci & Akilli, (2021)<br>Ballıdağ & Dikilitaş, (2021)<br>Bener & Yildiz, (2019)<br>Eisenlauer, (2020)<br>Fakhruddin, (2020)<br>Fyfield et al., (2021)<br>Jong & Tan, (2021)<br>Karakas & Kartal, (2020)<br>Le et al., (2021)<br>Lee, (2022)<br>Liu et al., (2020)<br>Luy, (2022)<br>Özkan & Güler, (2018) |

| | |
|---|---|
| | Say & Yildirim, (2020)<br>Schmidt & Rye, (2020)<br>Teo et al., (2019)<br>Tzotzou, (2018) |

### 4.2.1.Promote students' effective language learning and positive learning behaviours

Arabaci and Akilli (2021) stated in their articles that students' reading, writing, listening, and speaking skills would be enhanced by a foreign language education corporate effective technology. Incorporating Web 2.0 into education or instruction had a high likelihood of improving the situation. Once four fundamental abilities have been mastered, pupils' self-consciousness about automatic learning would skyrocket, and their motivation to learn English would increase. zkan and Güler (2018) concurred that using web 2.0 resources, such as Podcasts, was conducive and valuable for enhancing language abilities (listening, writing, reading, etc.). c. In addition to the four skills of listening, writing, reading, and speaking, Eisenlauer (2020) demonstrated in his study that lexical abilities, such as vocabulary size or chunks, idioms, and phrases, can be gradually enhanced with the help of web 2.0 tools, such as multimodal meaning-making tools in instruction. Under such accumulation, Native-like Fluency and accuracy in English are likely to be attained by students.

As for motivation, Fakhruddin's findings in 2020 indicated that pupils tend to demonstrate a strong and positive attitude about learning when teachers create engaging materials using web 2.0 technology. Under the tech-tool assistance, the difficulty of learning a foreign language was lessened, allowing pupils to identify and comprehend words and sentences readily. Jong and Tan (2021) noted that integrating technological tools could boost students' motivation for writing. Students would like the learning process because of the attractive characteristics of technology tools such as a pallet. Positive attitudes would increase, while frustration and resistance to learning would decrease. Students' active participation and collaboration in the classroom could also be encouraged. Mei and his colleagues (2018) have concluded that ubiquitous learning could be generated by students' high motivation for language learning when technology is integrated. Students quickly access diverse or multimodal resources thanks to technology tools in the classroom. Diverse instructional activities boost students' interest and enthusiasm for English study.

For the interaction, Lee (2022) indicated that some characteristics of technology tools can support the interactive possibilities. Due to the capabilities of technological tools, students are more inclined to communicate and exchange ideas with others. Simultaneously, Luy (2022) claimed that introducing digital tools, such as Kahoot and others, increased student engagement in the classroom. The continuity of students in the classroom was enhanced. Web 2.0 adoption in the classroom facilitated the organization of synchronous instructional activities, particularly following the Coronavirus pandemic. Compared to the traditional knowledge delivery technique, classroom participation and collaboration among students have increased.

### 4.2.2.Enrich teaching methods and improve teachers' teaching level

According to the 17 papers, teachers' teaching methods and levels can be enhanced by acquiring authentic and diversified teaching resources, contextualizing learning content, a continuum of professional learning growth, increased social collaboration, and improved attitudes. For example, Eisenhauer (2020) demonstrated that technology support not only assists teachers in locating authentic or diverse language materials, but also encourages teachers to contextualize or recontextualize texts. Students can acquire a language more effectively in a natural and contextualized learning setting. Fakhruddin (2020) emphasized that teachers might utilize tool functions, such as editing, combining, and generating, to contextualize and recontextualize student-centered content. These improved teaching or learning resources, which are more authentic and diversified, could spark motivation and positive learning attitudes among students. Moreover, when teachers edited, their professional competencies in digital literacy could be enhanced. Fyfield and colleagues (2021) outlined how a web 2.0 platform could periodically update teachers' instructional materials. These materials serve as a supplement to present instructional content. Teachers' excitement for teaching and learning would increase, and their views toward teaching using technology would continue to improve. In his essay published in 2022, Lee discussed how many search engines could assist teachers in locating authentic content and designing materials based on the teaching context. The role of "hashtag" could facilitate social collaboration by integrating the same experience or group. zkan and Güler (2018) also demonstrated that a positive effect of technology integration in instruction was that teachers not only provide a means to expose students to authentic contexts with authentic resources and a natural learning environment but also equipped themselves with the 21st-century teaching skill requirement of e-teaching. g. Schmidt and Rye (2020) wanted to analyze the web 2.0 tool Wikipedia

usage in the classroom. They discovered that the tool could assist teachers in locating various materials and developing practical teaching contexts. Using technology could provide teachers with authentic and contextualized resources.

For the professional development and social collaboration, Arabaci and Akilli (2021) examined the impact of web 2.0 tools on the educational environment of English instructors and found that most teachers' professional development is encouraged. Specifically, concerning digital literacies, participants reported that their technological competence, such as organizing, designing, editing, and disseminating information with technology, improved significantly as their technological skills grew. Ballda and Dikilitaş (2021) also emphasized the online professional development of teachers. Their knowledge and fluency with web 2.0 tools would no longer be a barrier to teachers' efficient use of technology in the classroom. Additionally, social collaboration among teachers could be enhanced. Instructors would devote more time collaborating with experts and seasoned teachers to discuss the optimal implementation in English classrooms. Bener and Yildiz (2019) observed that integrating technology into practice teaching would increase the rate of teacher reflection and social collaboration. Reflection could gradually repair errors and enhance teachers' teaching levels. Conversely, when teachers expressed themselves in an online learning community, it could be seen that their social participation promoted mutual learning by getting diverse and numerous perspectives. Le and colleagues (2021) saw technology instruments' use in professional learning and social collaboration as significant. Teachers actively share their teaching approaches and successes online and connect. Their digital skills and English ability could be enhanced through professional learning development. Liu and his partners (2020) have also demonstrated the benefits of social integration through the general use of technology in the classroom. The formed socialized communities could assist teachers in discovering more diversified materials and acquiring new teaching techniques, which would be conducive to their future professional development. Say and Yildirim (2020) investigated that teacher generally identified web 2.0 as the medium for interacting with others and acquiring English teaching-related information. Teachers demonstrated a strong willingness and good attitudes to use them extensively in the classroom.

As for the positive attitudes, Jong and Tan (2021) analyzed educators' utilization of online teaching platforms. They found that most teachers had good opinions about implementing web 2.0 tools in the classroom. Karakas and Kartal (2020) discovered that teachers had a strong propensity to use social media websites where they may freely engage with others and obtain helpful feedback to improve teaching standards. These allowed teachers to engage in a competitive environment and develop a sense of self-improvement. Teachers would insist on changing their instructional methods. Many teachers had good attitudes toward e-teaching—technology integration in the classroom, as described by Luy (2022) because of the emergence and ubiquity of tech tools. With web 2.0 tools, instructors could build numerous engaging, synchronous, and dynamic educational activities, and their digital literacies of professional development have also been supported. Teo and colleagues (2019) evaluated teachers' adoption of web 2.0 tools in the classroom and found that teachers viewed the benefits of web 2.0 use positively. Particularly after consolidating knowledge of TPACK and widespread implementation in instruction, teachers demonstrated firm, positive, and pleasurable attitudes for its continued use in the future. In his survey, Tzotzou (2018) noted that most EFL teachers had positive attitudes toward the continued integration of web 2.0 tools in EFL classroom settings. Overall, teachers can benefit themselves, like gaining authentic resources, contextualizing content, improving professional development and social collaboration, increasing positive attitudes under the pedagogical application of web 2.0 technology.

### 4.3. The barriers of integrating web 2.0 in English classroom

Across the reviewed 20 articles, the investigated barriers to impede appropriate implementation of web 2.0 are divided in to three dimensions: the web 2. Tool itself, teachers, and some exterior perspectives (See Table 4).

Table 7. The barriers of integrating web 2.0 in English classroom

| Themes | Sub-themes | Sources |
|---|---|---|
| The first-order tool barrier | • The lack of systematic arrangement<br>• low reliability of shared content<br>• commercial-dominated platform | Arabaci & Akilli, (2021).<br>Eisenlauer, V. (2020)<br>Fyfield et al., (2021)<br>Le et al., (2021)<br>Liu et al., (2020)<br>Schmidt, K. J., & Rye, E. (2020) |

| | | |
|---|---|---|
| The second-order teacher barriers | • Increased workload<br>• Insufficient knowledge of web 2.0 knowledge and skills Social collaboration | Fakhruddin, (2020)<br>Fyfield et al., (2021)<br>Karakas & Kartal, (2020)<br>Le et al., (2021)<br>Lee, (2022)<br>Luy, (2022)<br>Mei et al., (2018)<br>Mohammad-Salehi et al., (2021)<br>Teo et al., (2019)<br>Tzotzou, (2018) |
| The third- order exterior barriers | • Disparity of digital tool infrastructure<br>• Inadequate guidance and training<br>• Internet connections problems<br>• The lack of parents' support | Arabaci & Akilli, (2021)<br>Jong & Tan, (2021)<br>Karakas & Kartal, (2020)<br>Le et al., (2021)<br>Luy, (2022)<br>Mei et al., (2018)<br>Mohammad-Salehi et al., (2021)<br>Santos & Castro, (2021)<br>Teo et al., (2019)<br>Tzotzou, (2018) |

### 4.3.1. The first-order tool barrier

Even though web 2.0 technology has been widely disseminated and popularized in the education sector, it has not yet been implemented smoothly. There are three slightly noticeable parts of the tool barrier: lack of systematic organization, limited trustworthiness of shared content, and a platform controlled by corporate entities. For instance, Isenlauer (2020) indicated that some tools often displayed a variety of unorganized material types. Especially for new users, this layout and overview of the information would be overwhelming. Teachers and students would lose interest in its continued use due to its time-consuming approach. Arabaci and Akilli (2021) also noted that the unregulated system of the tool was the most evident detrimental effect of web 2.0 technology on instruction. Due to the unexpected content, which often took considerable time to search, it would be difficult for teachers to teach continuously. For content reliability, Schmidt & Rye (2020) noted in their study that although most internet resources are free, the quality of these sources cannot be assured due to their low academic or school curriculum standards. Due to its incompatibility with Curriculum requirements, Liu and other researchers (2020) asserted that the veracity and precision of disseminated content on various digital tools may be relatively uncertain. On the commercial-dominated aspect, Fyfield and other fellows (2021) asserted that since some technologies were commercially dominated, advertising on those tools would maximize teachers' screen time and may misdirect teachers' initial focus. Teachers are prone to hesitating when deciding which tool is ideal for incorporating into the classroom. Le and helpers (2021) noted that some tools were commercially dominated, which might significantly impact the teacher's attention and time spent searching owing to advertisements. It would be difficult for teachers to choose dependable teaching resources. Ultimately, these obstacles result in poor teaching quality and untargeted content for language learners.

### 4.3.2. The second-order teachers' barriers

Increased workload and insufficient web 2.0 knowledge and abilities are the primary obstacles aimed at teachers that restrict technology deployment in education. For the increased workload, Le and helpers (2021) found that teachers' workloads

would increase if they were inundated with information. In his article, Lee (2022) showed that implementing technological interaction could increase teachers' workload. Teachers must consciously select and create more multimodal, interactive, and engaging materials. Teachers would most likely be overburdened by the research and planning processes. Luy (2022) remarked that elementary school teachers had a more challenging workload than secondary school teachers since young students required engaging activities to capture their attention. This situation needed teachers to find more relevant interaction design tools, which would increase their workload and cause them to feel stressed. According to Mei and other academics (2018), the increased workload for teachers was one of the primary issues of using Web 2.0 in the classroom. Teachers must devote more effort and time to pre-class and post-class activities if they want to provide a flawless class with technological integration.

As for the insufficient web 2.0 knowledge and skills, Fyfield and other colleagues (2021) stressed that pre-service or beginner instructors would likely lack sufficient web 2.0 knowledge and skills more than in-service or experienced teachers with more collected experience. In their article, Fakhruddin (2020) and Luy (2022) stated that if teachers wished to dynamically and flexibly incorporate web 2.0 in teaching, they needed a high level of digital literacy. However, most teachers still lack the knowledge and abilities necessary to conduct high-demand digital instruction. Karakas and Kartal (2020) observed that teachers were not adequately integrating technological tools into instruction because they lacked the necessary knowledge and skills. Mohammad-Salehi and colleagues (2021) inferred that most teachers could not guarantee the successful incorporation of web 2.0 technology into English instruction due to a lack of specific technical, pedagogy, and subject understanding. A successful classroom of the twenty-first century should demonstrate the combination of technology, pedagogy, and content insertion. Teo and colleagues (2019) discovered that pre-service teachers had little acquaintance with the necessary technologies deployed in contemporary educational environments. Teachers' competency with Web 2.0 tools falls short of expectations. In this essay, Tzotzou (2018) argued that teachers' lack of pedagogical and technical knowledge and utilization would lead to a lack of confidence to continue teaching with technology.

### 4.3.3. The third-order exterior barriers

The composition of exterior barriers comes from the infrastructure disparity, limited tech training, technical issues, and low support from parents. For equal access to infrastructure, Karakas & Kartal (2020) acknowledged the absence of accessibility issues. There would be a disparity in the distribution of resources among teachers in various geographic regions. Some teachers had a more difficult time gaining access to advanced tools. Luy (2022) stated that the infrastructure in some schools, such as digital gadgets, was incomplete. In addition, many parents demonstrated negative support for teachers' instruction, such as a lack of comprehension of teachers' teaching tempo and a lack of support for learning equipment. Mohammad-Salehi and coworkers (2021) found that the degree to which schools were equipped with technology differs significantly throughout regions according to administrative policies. Some sub-regions had limited digital infrastructure, but others had advanced technological assistance. It was impossible to guarantee the uniformity of web 2.0 implementation in schools. Tzotzou (2018) and Jong and his coworker (2021) showed that certain EFL teachers encountered a lack of digital tool infrastructure availability and accessibility. In addition to receiving insufficient training, teachers were also given fewer digital gadgets. All could contribute to a lack of faith in web 2.0 integration in English instruction. Santos and Castro (2021) have also highlighted the disparity between real-world infrastructure and policy idealization. Schools did not receive sufficient financial aid to construct a stable and somewhat comprehensive infrastructure.

Insufficient and timely training with technology is the second exterior barrier. Mei and other researchers (2018) noted the lack of appropriate training planning and preparation for instructors to become more acquainted with the web 2.0 integration. Typically, inadequate training and police backup occurred. Teo and colleagues (2019) indicated that the shared knowledge and skill of technology integration among pre-service teachers was due to inadequate guidance and training. Teachers have not received sufficient and comprehensive training to support their future teaching careers. According to Le and his colleagues (2021), the lack of support and direction from schools and relevant policy discouraged teachers continued use of web 2.0 tools in the classroom. The lack of policy support from key authorities would result in most teachers receiving inadequate direction and training (Santos & Castro, 2021).

As for the technical issues, Arabaci and Akilli (2021) noted that the internet problem was the most evident detrimental effect of web 2.0 tools on classroom instruction. Due to the unexpected internet outage, which often took considerable time to repair, it would be difficult for teachers to teach continuously. Jong and Tan (2021) also discussed internet connection technical concerns. Many educators would encounter this circumstance and be unsure of how to respond. The quality of instruction would decline. If the government and relevant authorities do not consider these exterior aspects in improvement policies while the pedagogical integration of web 2.0, the reality will not be ideal as expectations.

## 5. Discussion and conclusion

In the study, 20 articles were obtained from four journal publishers and a database in response to the categories of web 2.0 tools, benefits, and difficulties associated with incorporating web 2.0 technology in the classroom. Even though web 2.0 was launched in 2004, the results suggested that its use in the educational field is gaining prominence in many studies. Classroom uses of web 2.0 by English teachers demonstrate that the educational field has developed its use of web 2.0. Remarkably, the multimodal usage allows EFL teachers to upload and download instructional movies from websites. Teachers who employ Web 2.0 technologies observe each tool's functionality and incorporate it into their English lessons. It implies that current English teachers prefer to use more relevant and acceptable elements of teaching tools, recognizing the significance of using technology in topic instruction. It also confirms the hypothesis of TPACK 2.0 (Jimoyiannis et al., 2013) that technology-integrated content instruction can significantly assist teachers in recognizing the three intersections of pedagogy, content, and technology. Following instructional situations and curriculum standards, teachers can select the most effective teaching instruments more dynamically and flexibly.

As for the benefits, previous studies (Hew & Cheung, 2013; Zeng, 2020; Halim & Hassan, 2019) that placed the learning advantages of students as the top priority can be refuted. Teachers, on the other hand, can also be a prioritized analytical target group. Teachers can acquire authentic and diverse teaching resources (Eisenhauer, 2020; Fakhruddin, 2020); Fyfield et al., 2021; zkan & Güler, 2018; Schmidt & Rye; 2020) as well as contextualizing learning content (Eisenhauer, 2020; Fakhruddin, 2020; Schmidt & Rye; 2020), and professional learning development (Arabic & Akili, 2021; Ball u (2018). These advantages among teachers can further illustrate the use of the TPACK 2.0 model in the education field so that instructors can better comprehend the rationale behind advocating the effective implementation of technology in their classrooms. The ongoing usage of web 2.0 technologies in the classroom will demonstrate the technology's viability in education. When schools or police administrations embrace technology, teachers will exhibit less resistance on the field. Similarly, the initiative-driven demand for teacher knowledge, such as technical content, technological pedagogy, and pedagogical content know-how, will positively impact the classroom. With web 2.0 help, teachers may link and apply their professional knowledge and instructional practices more intelligently and strategically. Additionally, the quality of instruction can be improved, and instructors' professional growth will be promoted automatically.

Even though web 2.0 has been developed and promoted for many years, there are still obstacles to its successful adoption in education, including the tool, teachers, and external influences. If teachers want to compensate for their limited knowledge and abilities of web 2.0 technology, well-developed TPACK knowledge and skills are needed. In that case, they will need to commit more time and effort, increasing their burden. In response to professional development training aligned with the TPACK theory, more individual training practices, such as active technology use and meaningful integration of technology, content, and instructional approaches, are required. In the meantime, unsupported school administration policies can eventually result in disparate digital tool infrastructure and inadequate training. There also include some unpredictable components like technical challenges and unsupported attitudes from parents. All the existing issues indicate that changes and improvements must be implemented to promote technology integration in education. To rectify the disparity between reality and ideal, school or police administrations should provide substantial and sufficient financial assistance for infrastructure development, such as acquiring high-tech digital equipment and providing equal access to various tools, etc. Once the essential infrastructure for efficient technology adoption has been established, pre-service and in-service teachers should receive the following professional development courses. In response to the training content, training and implementation should prioritize actual social behaviors and contextualized instruction (Collinson et al., 2009; Corcoran, 1995). To cultivate well-rounded, application-oriented instructors, the content of pre-service teacher training should emphasize the current teaching environment and future requirements.

To contextualize and correct the lack of technology integration in the classroom, schools should arrange in-service teachers' training closely tied to their everyday experiences. Teachers will become increasingly adept and adaptable in dealing with challenging situations. For instance, pre-service education should include technology pedagogical subject understanding and web 2.0 skills. Then, future pre-service teacher applications of acquired knowledge in actual classroom settings are also required. In-service teacher training must also correct and update TPACK 2.0 knowledge and abilities. However, the focus should be on effectively integrating technology into current teachers' classrooms. As for teachers, they should have a strong sense for developing professional learning, such as English proficiency, digital literacy, innovative pedagogy under web 2.0 sustainability, and technology integration-friendly policies. Following the social constructivist learning theory in professional development (Sheffield et al., 2018), teachers should be strongly encouraged to interact with others on collaborative platforms for their high-quality professional development, where they can freely share and communicate their experiences. Expert or experienced teachers must provide more outstanding help for novice teachers and share more critical information and skills via interactive platforms. Under the stimulation of the social community, this can significantly increase social collaboration among instructors and entice teachers' automatic learning for their professional development. Teachers in the classroom must provide students with more straightforward explanations and instructions to reduce student stress and boost their learning motivation. Parents exert an essential function in supporting the incorporation of technology in education. K-12 pupils have comparatively low levels of self-focus and self-discipline compared to higher students. They have a high propensity for internet addiction and immersion. Parents can simultaneously control their

children's behaviour and assist them with learning devices. Technology integration in teaching can eventually be successful, and the health dynamics of teaching and learning with technology can be realized.

This paper proposes a web 2.0 educational framework on pedagogy to improve teachers' professional development and to equip instructors with a comprehensive framework for adopting web 2.0 as an essential and effective learning tool to be integrated into their instructional practices. Educators will be able to select relevant Web 2.0 technology tools for use in the classroom, boost their motivation to continue using Web 2.0 in the classroom, and be aware of potential problems and challenges. It will improve both the teaching procedure and the quality of instruction. However, this paper may have certain drawbacks. Due to the capstone requirement, only the author conducts a comprehensive systematic evaluation of this work, which will likely result in some bias towards particular concerns and research questions. The accuracy and reliability of this work will be inferior to those written by two or three authors. The author scrupulously adheres to the PRISMA 2020 statement on maintaining data transparency and dependability to mitigate personal bias. As for the sample selected for analysis, the author primarily selected journal publishers for the intended sample, reducing the sample sizes necessary for analysis. Significant accuracy will be compromised. In addition, this article focuses mainly on K-12 English classrooms. Other subjects' web 2.0 implementation and higher education level courses are ignored. It will not reflect the entirety of technology integration instruction across all grade levels and classrooms. Therefore, future studies on the pedagogical integration of Web 2.0 should involve a more significant number of researchers and samples to strengthen the data's dependability and precision. In addition to journal papers, cooperative databases can be accessed to ensure data completeness and evidence-based precision when analysing sample size. Additionally, future research should include topics other than English, such as mathematics, physics, and others, to shed light on the application of web 2.0 pedagogy integration in K-12 classrooms. This may reflect the K-12 application and education level as a whole.

## 6. References


Abram, S. (2005). Web 2.0 – huh?! Library 2.0, librarian, 2.0. *Information Outlook*, *9*(12), 44–46.

Alexander, B. (2006). Web 2.0: A new way of innovation for teaching and learning? *EDUCAUSE Review, 41*, 33-44.

Alexander, B., & Levine, A. (2008). Web 2.0 storytelling: Emergence of a new genre. *EDUCAUSE Review*, *43*(6), 40-56.

Alharbi, A. S. (2020). Wikis in language learning: Merits and limitations. *Journal of English Language Teaching and Applied Linguistics*, *2*(4), 79-88.

Aromataris, E., & Pearson, A. (2014). The systematic review: an overview. *AJN The American Journal of Nursing*, *114*(3), 53-58.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10–20.

Barsky, E. (2006). Introducing Web 2.0: Weblogs and podcasting for health librarians. *Journal of Canadian Health Librarian Association*, *27*, 33–34.

Bingimlas, K. A. (2017). Learning and teaching with Web 2.0 applications in Saudi K-12 schools. *Turkish Online Journal of Educational Technology-TOJET*, *16*(3), 100-115.

Borko, H., Whitcomb, J., & Liston, D. (2009). Wicked problems and other thoughts on issues of technology and teacher learning. *Journal of Teacher Education*, *60*(1), 3-7.

Bowman, M. A., Vongkulluksn, V. W., Jiang, Z., & Xie, K. (2022). Teachers' exposure to professional development and the quality of their instructional technology use: The mediating role of teachers' value and ability beliefs. *Journal of Research on Technology in Education*, *54*(2), 188-204.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychosociology. *Qualitative Research in Psychology*, *3*, 77-101.

Braun, V., & Clarke, V. (2012). *Thematic analysis*. American Psychological Association.

Brown, S., & Adler, R. P. (2008). Minds on fire. Open education, the long trail, and learning 2.0. *Educause, 43*, 17-32.

Bull, G., & Garofalo, J. (2006). Commentary: Ubiquitous computing revisited – A new perspective. *Contemporary Issues in Technology and Teacher Education*, *6*(2), 271-274.

Carvin, A. (2005). *Time Berners-Lee: Weaving a Semantic Web*. Retrieved July 24, 2007, from http://www.digitaldivide.net/articles/view.php?ArticleID = 20

Chai, C. S., Koh, J. H., & Tsai, C.-C. (2010). Facilitating preservice teachers' development of technological, pedagogical, and content knowledge. *Educational Technology & Society,* 13(4), 63–73.

Collinson, V., Kozina, E., Kate Lin, Y. H., Ling, L., Matheson, I., Newcombe, L., & Zogla, I. (2009). Professional development for teachers: A world of change. *European journal of Teacher Education*, *32*(1), 3-19.

Corcoran, T. B. (1995). *Helping teachers teach well: Transforming professional development. Consortium for Policy Research in Education RB-16*. New Brunswick, N.J.: Rutgers, The State University of New Jersey.

DeSantis, L., & Ugarriza, D. N. (2000). The concept of theme as used in qualitative nursing research. *Western Journal of Nursing Research, 22*, 351-372.

Downes, S. (2005). E-learning 2.0. *ELearn, 2005*(10), 1.

Drucker, P. F. (1995). *The age of social transformation*. Centre canadien de gestion.

Ganser, T. (2000). An ambitious vision of professional development for teachers. *NASSP Bulletin*, *84*(618), 6-12.

Halim, M. S. A. A., & Hashim, H. (2019). Integrating web 2.0 technology in ESL classroom: A review on the benefits and barriers. *Journal of Counseling and Educational Technology*, *2*(1), 19-26.

Hew, K. F., & Cheung, W. S. (2013). Use of Web 2.0 technologies in K-12 and higher education: The search for evidence-based practice. *Educational Research Review*, *9*, 47-64.

Hunzicker, J. (2011). Effective professional development for teachers: A checklist. *Professional Development in Education*, *37*(2), 177-179.

Hutchison, A., & Woodward, L. (2014b). A planning cycle for integrating technology into literacy instruction. *Reading Teacher, 67*(6), 455–464.

Jimoyiannis, A. (2015). TPACK 2.0: Towards a framework guiding Web 2.0 integration in educational practice. In M. S. Khine (Ed.), *New directions in technological pedagogical content knowledge research multiple perspectives*. (pp. 83-108). Information Age Publishing.

Jimoyiannis, A., Tsiotakis, P., & Roussinos, D. (2011). Pedagogical and instructional design issues towards the integration of Web 2.0 tools in instruction-Implications of teachers' training pilot courses in Greece. *2011 7th International Conference on Next Generation Web Services Practice*, 530-535,

Jimoyiannis, A., Tsiotakis, P., Roussinos, D., & Siorenta, A. (2013). Preparing teachers to integrate Web 2.0 in school practice: Toward a framework for Pedagogy 2.0. *Australasian Journal of Educational Technology*, *29*(2), 248-267.

Kiteley, R., & Stogdon, C. (2014). *Literature reviews in social work*. SAGE.

Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly*, *50*(1), 37–59.

Mcloughlin, C., & Lee, M. (2007). Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era introduction: Social trends and challenges. *Proceedings ascilite Singapore 2007*, 664– 675.

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054.

Mishra, P., & Koehler, M. J. (2008). Introducing technological pedagogical content knowledge. In *annual meeting of the American Educational Research Association*, 3-16.

Mouza, C. (2002). Learning to teach with new technology: Implications for professional development. *Journal of Research on Computing in Education*, *35*(2), 272-289.

Nelson, J., Christopher, A., & Mims, C. (2009). TPACK and web 2.0: Transformation of teaching and learning. *TechTrends*, *53*(5), 80.

O'reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies*, (1), 17.

Ottenbreit-Leftwich, A. T., Glazewski, K. D., Newby, T. J., & Ertmer, P. A. (2010). Teacher value beliefs associated with using technology: Addressing professional and student needs. *Computers & Education*, *55*(3), 1321-1335.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 1-11.

Pan, S. C., & Franklin, T. (2011). In-service teachers' self-efficacy, professional development, and web 2.0 tools for integration. *New Horizons in Education*, *59*(3), 28-40.

Paulus, M. T., Villegas, S. G., & Howze-Owens, J. (2020). Professional learning communities: Bridging the technology integration gap through effective professional development. *Peabody Journal of Education*, *95*(2), 193-202.

Richardson, W. (2006). *Blogs, wikis, podcasts, and other powerful Web tools for classrooms*. Thousand Oaks, CA: Corwin Press.

Sheffield, R., Blackley, S., & Moro, P. (2018). A professional learning model supporting teachers to integrate digital technologies. *Issues in Educational Research, 28*, 487–510.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Sykes, J. M., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *Calico Journal*, *25*(3), 528-546.

Thompson, J. (2007). Is education 1.0 ready for Web 2.0 students? *Innovate: Journal of Online Education*, *3*(4).

Tu, C., Blocher, M., & Ntoruru, J. (2008). Integrate Web 2.0 technology to facilitate online professional community: EMI special editing experiences. *Educational Media International*, 45(4), 335–341.

King, J., & South, J. (2017). Reimagining the role of technology in higher education: A supplement to the national education technology plan. *US Department of Education, Office of Educational Technology, 16-70*.

Usluel, Y. K., & Mazman, S. G. (2009). Adoption of Web 2.0 tools in distance education. *Procedia-Social and Behavioral Sciences*, *1*(1), 818-823.

Warschauer, M., & Grimes, D. (2007). Audience, authorship, and artifact: The emergent semiotics of web 2.0. *Annual Review of Applied Linguistics*, 27, 1-23.

Zeng, S. (2020). The potential of online technology for language learning. *English Language Teaching*, *13*(10), 23-37.

## 7. Appendix

**(The final 20 articles included in the systematic review)**

Arabaci, I. B., & Akilli, C. (2021). English Teachers' Views on the Use of Web 2.0 Tools in Educational Environments. *Asian Journal of Education and Training*, *7*(2), 115-125.

Ballıdağ, S., & Dikilitaş, K. (2021). Preparatory School Teachers' Self-Directed Online Professional Development. *Iranian Journal of Language Teaching Research*, *9*(3 (Special Issue)), 25-38.

Bener, E., & Yildiz, S. (2019). The use of blog activities to promote reflection in an ELT practicum. *Australian Journal of Teacher Education (Online)*, *44*(8), 38-56.

Eisenlauer, V. (2020). The EFL-YouTube remix: Empowering multimodal and computational literacies for EFL purposes. *Journal of Visual Literacy*, *39*(3-4), 149-166.

Fakhruddin, Z. (2020). Designing English Listening Materials through YouTube Video Editing: Training for English Teachers of Islamic Junior High Schools, Parepare, South Sulawesi. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *4*(2), 275-290.

Fyfield, M., Henderson, M., & Phillips, M. (2021). Navigating four billion videos: teacher search strategies and the YouTube algorithm. *Learning, Media and Technology*, *46*(1), 47-59.

Jong, B., & Tan, K. H. (2021). Using Padlet as a Technological Tool for Assessment of Students' Writing Skills in Online Classroom Settings. *International Journal of Education and Practice*, *9*(2), 411-423.

Karakas, A., & Kartal, G. (2020). Pre-Service Language Teachers' Autonomous Language Learning with Web 2.0 Tools and Mobile Applications. *International Journal of Curriculum and Instruction*, *12*(1), 51-79.

Le, V. H., Maor, D., & McConney, A. (2021). The potential of social networking sites for continuing professional learning: investigating the experiences of teachers with limited resources. *Studies in Continuing Education*, 1-17.

Lee, Y. J. (2022). Language learning affordances of Instagram and TikTok. *Innovation in Language Learning and Teaching*, 1-16.

Liu, Y., Torphy, K. T., Hu, S., Tang, J., & Chen, Z. (2020). Examining the virtual diffusion of educational resources across teachers' social networks over time. *Teachers College Record*, *122*(6), 1-34.

Luy, D. T. T. (2022). Remote Teaching amid the Covid-19 Pandemic in Vietnam: Primary School EFL Teachers' Practices and Perceptions. *Dau, TTL (2021). Remote Teaching amid the Covid-19 Pandemic in Vietnam: Primary School EFL Teachers' Practices and Perceptions. AsiaCALL Online Journal*, *13*(1), 1-21.

Mei, B., Brown, G. T., & Teo, T. (2018). Toward an understanding of preservice English as a Foreign Language teachers' acceptance of computer-assisted language learning 2.0 in the People's Republic of China. *Journal of Educational Computing Research*, *56*(1), 74-104.

Mohammad-Salehi, B., Vaez-Dalili, M., & Heidari Tabrizi, H. (2021). Investigating Factors That Influence EFL Teachers' Adoption of Web 2.0 Technologies: Evidence from Applying the UTAUT and TPACK. *TESL-EJ*, *25*(1), 1-21.

Özkan, Y., & Güler, S. (2018). Podcast applications in pre-service language teacher education from a constructivist perspective. *World Journal on Educational Technology: Current Issues*, *10*(3), 131-141.

Santos, J. M., & Castro, R. D. (2021). Technological Pedagogical content knowledge (TPACK) in action: Application of learning in the classroom by pre-service teachers (PST). *Social Sciences & Humanities Open*, *3*(1), 100110.

Say, S., & Yildirim, F. S. (2020). Investigation of Pre-Service Teachers' Web 2.0 Rapid Content Development Self-Efficacy Belief Levels and Their Views on Web 2.0 Tools. *International Journal of Educational Methodology*, *6*(2), 345-354.

Schmidt, K. J., & Rye, E. (2020). Simple English Wikipedia: Free Resources for Beginner to Intermediate Levels. In *English Teaching Forum*, 58(2), 14-21.

Teo, T., Sang, G., Mei, B., & Hoi, C. K. W. (2019). Investigating pre-service teachers' acceptance of Web 2.0 technologies in their future teaching: a Chinese perspective. *Interactive Learning Environments*, *27*(4), 530-546.

Tzotzou, M. (2018). Integrating Web 2.0 technologies into EFL learning in the Greek state-school context: A mixed-method study. *Research Papers in Language Teaching & Learning*, *9*(1).

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 34-49*

Applied Linguistics Inquiry

https://doi.org/10.22077/ALI.2023.6087.1010

# EFL Teachers' Burnout during the Covid-19 Pandemic: Can Teaching Context Make a Difference?

Hossein Navidinia[1]
Fatemeh Zahra Gholizadeh[2]
Fateme Chahkandi[3]

[1]Associate Professor in Applied Linguistics, Department of English Language, University of Birjand, Iran
[2]M.A. Student in Applied Linguistics, Department of English Language, University of Birjand, Iran
[3]Assistant Professor in Applied Linguistics, Department of English Language, University of Birjand, Iran

**ARTICLE HISTORY**

Received: 14 June 2022
Revised: 14 October 2022
Accepted: 21 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: navidinia@birjand.ac.ir

**ABSTRACT**

The present study aimed to examine the differences between Iranian EFL teachers' burnout in online classes during the Covid-19 pandemic in public schools (PSs) and private language institutes (PLIs). It also investigated the causes of teacher burnout in the two contexts. The participants included 268 Iranian EFL teachers (108 teaching at PSs and 160 teaching at PLIs). The study employed a mixed-method design. In the quantitative phase, 268 EFL teachers were asked to answer Maslach et al.'s (1996) Burnout Inventory and in the qualitative phase, semi-structured interviews were conducted with 20 EFL teachers (10 from PSs and 10 from PLIs). The results of the first phase of the study indicated a significant difference between the two groups in the two dimensions of burnout including "Emotional Exhaustion", and "Personal Accomplishment". However, the difference between the "Depersonalization" aspect of burnout was not significant. The results of the second phase of the study indicated that factors such as challenges of online teaching, concerns about teaching effectiveness, lack of collegiality and principal/manager support, financial concerns and low wages, lack of job security, students' misbehavior and lack of motivation, lack of teacher autonomy, heavy workload, as well as school/institute policies were conducive to teacher burnout.

**KEYWORDS:** Covid-19 pandemic; EFL teachers; Online teaching; Teacher burnout

## 1. Introduction

The outbreak of the Covid-19 has created educational systems across the world unprecedented problems. As a result of the pandemic and the concomitant quarantine, many educational institutes were obliged to transition to online delivery mode of education to ensure public health. Despite its advantages, this abrupt shift in education has been accompanied by particular hassles, one of the most significant of which is ensuring the quality of instruction in online classes. As research has frequently demonstrated, the quality of instruction is among the most crucial factors affecting student achievement (Clotfelter et al., 2007; Haycock, 1998; Nye et al., 2004). Accordingly, teachers have critical roles to play both in virtual courses and in in-person courses and influence the success or failure of educational systems.

Teachers experience a plethora of stressors in online teaching including increased workload, use of new technologies, conflict, and role ambiguity, to name a few (Kaya, 2022; Pressley, 2021; Vargas Rubilar & Oros, 2021). In the absence of

**Cite this article:** Navidinia, H., Gholizadeh, F. Z., & Chahkandi, F. (2023). EFL teachers' burnout during the Covid-19 pandemic: Can teaching context make a difference?. *Applied Linguistics Inquiry*, *1*(1), 34-49. doi: 10.22077/ali.2023.6087.1010

efficient coping strategies, these stressors can result in teacher burnout which is referred to as "the state of physical and emotional depletion resulting from conditions of work'' (Freudenberger, 1974, p. 160). Prior research has indicated that burnout can adversely influence the quality of instruction and can result in job dissatisfaction which can culminate in teachers' decision to leave the profession (Huberman & Vandenberghe, 1999). Furthermore, the literature has documented a multitude of factors which can give rise to teacher burnout (Navidinia & Heiran, 2017). These factors can be broadly classified into contextual and personal ones. Contextual factors, among others, may include the type of institution, its curriculum, its guidelines, and instruction modality (Bartosiewicz et al., 2022; Khezerlou, 2017; Kimsesiz, 2019; Panisoara et al., 2020). Personal determinants which are both affected by and affect the contextual variables are emotional intelligence, resilience, self-esteem, and self-efficacy to name a few (Salovey & Mayer, 1990).

Taking into account that online teaching particularly under the conditions of the pandemic is a new phenomenon and that teachers experience a host of stressors which might be different from those specific to face-to-face, it is imperative to examine the causes of teacher burnout in this milieu. Furthermore, in the view of the crucial role of the contextual factors in teachers' experience of burnout (Kokkinos 2007) and the contextual variations that exist between public schools (PSs) and private language institutes (PLIs) in Iran (e.g. differential levels of institutional demand, teachers' view toward teaching, access to resources, and the amount of collegial support available (Moradkhani & Shirazizadeh, 2017), this study aims to look into how EFL teachers' experience of burnout is different with regard to the peculiarity of the two contexts.

As a consequence, this study aims to first, compare the level of Iranian EFL teachers' burnout during the Covid-19 pandemic in PSs and PLIs and second, to investigate the factors causing teacher burnout in two contexts. An investigation of this kind can harness the stakeholders' awareness including teacher educators' and the administrators' of the factors causing burnout in web-based teaching before any reactive measures can be taken to prevent and remedy such burnout. In particular, the study is guided by the following research questions:

1. Is there any significant difference between Iranian EFL teachers' level of burnout in PSs and PLIs?
2. What factors can cause Iranian EFL teachers' burnout in online education during the Covid-19 pandemic in PSs and PLIs?

## 2. Review of literature

### 2.1. Teacher burnout: definitions and contributing factors

Over the years, various definitions have been proposed for the term "burnout". Maslach, as one of the most prominent figures in the literature on burnout, alongside Jackson, defined burnout as "a syndrome of emotional exhaustion and cynicism that occurs frequently among individuals who do people-work of some kind" (Maslach & Jackson, 1981, p. 99). On the basis of this definition, they divided the concept into three key aspects of emotional exhaustion (EE), depersonalization (DP), and personal accomplishment (PA) (Maslach & Jackson, 1981).

Emotional exhaustion which arises as a result of experiencing high levels of stress and the demands of the job, describes a person's emotional resources being depleted forcing them to withdraw psychologically from their jobs (Maslach & Jackson, 1981). Depersonalization which is also related to emotional exhaustion, is manifested through negative attitudes toward the clients (Maslach & Jackson, 1981), which in the case of the teaching profession includes the students. Of course, such negative views are not limited to others considering that individuals suffering from burnout tend to evaluate themselves and their professional achievements negatively as well, thereby feeling dissatisfied with themselves, their clients, and their jobs (Edú-Valsania et al., 2022). As the third aspect of burnout, Personal Accomplishment (PA) refers to the tendency to view one's work and professional achievements negatively (Maslach & Jackson, 1981). In the case of teachers, diminished levels of PA are displayed through feelings of the lack of contribution to students' development, while developing a feeling of fatigue and tiredness and staying physically and emotionally distant are manifestations of emotional exhaustion and depersonalization, respectively (Grayson & Alvarez, 2008; Maslach et al., 1996).

In addition to Maslach and Jackson's (1981) classification of burnout, which is the most inclusive and widely-accepted one, other definitions discuss the various dimensions of the concept. One such definition is provided by Pines and Aronson (1988) which focuses on long-term exposure to the emotional demands of the job and defines burnout as the subjective experience of physical, emotional and mental exhaustion due to demands of the profession. In general, the physical and affective exhaustion as well as the environmental variables seem to play a role in the majority of the definitions of burnout. Furthermore, teachers constitute a popular subject pool for the studies of burnout considering the emotional demands of their profession (Schaufeli & Enzmann, 1998) and a bulk of research has endeavored to document the effect of different factors on teacher burnout.

With regard to the role of demographic variables in teacher burnout, there exists a host of conflicting results. On either side of the spectrum, studies show positive or negative correlations, making it difficult to draw any firm conclusions on the links between such variables and burnout levels (Bayani et al., 2013; Farshi & Omranzadeh, 2014; González-Morales et al., 2010; Green et al., 2014; Mousavy & Nimehchisalem, 2014; Purvanova & Muros, 2010). Gender is a significant indicator in different dimensions of burnout where women have been found to be more inclined toward emotional exhaustion and men toward depersonalization (Grayson & Alvarez, 2008; González-Morales et al., 2010; Purvanova & Muros, 2010).

The role of marital status in the experience of burnout is up for debate considering the conclusions from Mousavy and Nimehchisalem (2014), who identified significantly higher levels of burnout in married teachers as opposed to the findings of Jackson et al. (1986) in this regard. From among the other demographic variables included in the literature, teaching experience seems to account for burnout to some extent, specifically for burnout intensity and emotional exhaustion (Brewer & Shapard, 2004; Capel, 1987). Still, there are studies which fail to link burnout to any of the demographic variables including years of experience (Farshi & Omranzadeh, 2014; Hock, 1988; Yastibas, 2021).

Another category of variables in relation to burnout includes individual factors such as resilience, self-regulation, self-efficacy, and adaptability. It is not surprising that the literature supports a negative association between emotional intelligence, stress, and burnout levels which can be due to more efficient emotional regulation and a stronger sense of self (Gohm et al., 2005; Mayer et al., 2000; Mérida-López & Extremera, 2017; Pishghadam & Sahebjam, 2012; Vaezi & Fallah, 2011). Emotional intelligence helps emotional regulation and building tolerance, thus allowing for more efficient ways to cope with the stress and challenges of the job (Kariou et al., 2021; Mérida-López & Extremera, 2017; Pishghadam & Sahebjam, 2012; Vaezi & Fallah, 2011). As Pishghadam and Sahebjam (2012) reported emotional quotient (EQ) is associated with lower scores in emotional exhaustion and depersonalization and higher scores in personal accomplishment.

Self-efficacy is also related to emotional intelligence since it is a strong positive predictor of personal accomplishment and is linked to lower levels of burnout (Bartosiewicz et al., 2022; Foley & Murphy, 2015; and Motallebzadeh et al., 2014). Other personality traits including, adaptability, ambiguity tolerance, and neuroticism have been also indicated to affect burnout. While ambiguity tolerance and adaptability are both negative predictors of burnout (Azadianbojnordi et al., 2021; Zhaleh et al., 2018), higher scores in emotional exhaustion are best predicted by high scores in neuroticism and low scores in extroversion (Pishghadam & Sahebjam, 2012).

Another line of research is concerned with the relation between teacher burnout and the contextual factors, including type of institution, class size, administrative policies, and the demands of the job. The literature confirms the role of the socioeconomic status of the institute, excessive workload, low financial rewards, inadequate resources and training, as well as poor work conditions in manipulating burnout levels among teachers (Barutçu & Serinkan, 2013; Brissie et al., 1988; Carson III, 2006; Roohani & Dayeri, 2019). In this regard, Grayson and Alvarez (2008) studied the effects of school climate on the core three dimensions of burnout, revealing that teachers' experience of burnout was greatly affected by the presence of strong community and administrative support as well as proper instructional management. Yet in another study, Roohani and Dayeri (2019) contented that Iranian EFL teachers had lower levels of burnout since they were intrinsically motivated in their profession, verifying the role of motivation and self-regulation as predictors of burnout.

The teachers' choice of teaching style can also have a preventive role in burnout experience as demonstrated by Ghanizadeh and Jahedizadeh (2016). This seems to be affected by the type of institution the teacher works in, as the instructional policies and demands may vary in different contexts. Interestingly, in two separate studies aiming to identify context-laden factors affecting burnout by comparing Iranian high schools and language institutes, Rostami et al. (2015) and Heiran and Navidinia (2015) found that teachers in the public sector demonstrate significantly higher levels of burnout compared to their private section counterparts. These results are in sharp contrast with the findings of Soleimani and Bolourchi (2021) and Bahrami and Moradkhani (2019). The choice of teaching approach and the existence of such a choice for the teacher can have a role in their burnout experience. For example, in private language institutes teachers have the freedom to choose a more modern approach and thus can work together with their learners providing a chance to build rapport and flexibility in the class compared to the traditional approaches which are often used at schools (Rostami et al., 2015).

Studies also demonstrate that when the instructional rules and policies are perceived as limiting or incongruent with teachers' beliefs and outlooks, they can lead to a higher risk of burnout (Grayson & Alvarez, 2008; Salahshour & Esmaeili, 2021). These policies may exclude teachers from decision-making at school, and limit school resources and teachers' salaries, thus contributing to the experience of burnout (Betoret, 2009; Foley & Murphy, 2015).

### 2.2. Teacher burnout in online classes

Online education has gained more momentum since the outbreak of the Covid-19 pandemic and some researchers have sought to investigate teacher burnout during this trying time (Bartosiewicz et al., 2022; Bravo et al., 2021; Ferdous & Shifat, 2020; Panisoara et al., 2020). During the first wave of the national lockdowns, most educational institutes opted for fully online classes which proved to be demanding especially for teachers. Even as the teachers returned to school after a period of online classes, they were faced with dramatically different instructional routines and approaches (Pressley, 2021). More teachers are reported to have experienced burnout during the pandemic and the levels of burnout have also been shown to increase significantly (Pressley, 2021; Vargas Rubilar & Oros, 2021).

Studies also searched for the reasons behind such a trend and provided justifications and counteractions for the matter. For example, Izquierdo et al. (2021) investigated the relationship between language instructors' unpreparedness and their burnout level during the pandemic. Their findings point to an increase in work hours and workload as the main culprits in teacher burnout during the pandemic (Izquierdo et al., 2021). Similarly, Pressley (2021) and Kaya (2022) believe that increased job demands such as the need for more detailed lesson plans alongside the lack of efficient support particularly on technology-related skills, emotional as well as instructional support have substantially contributed to the increase in burnout levels. Being forced to acquire technology-related skills and pursue training with regards to technology-assisted language learning independently, despite having access to the internet, was among the factors that increased their workload (Izquierdo et al., 2021).

Ghanbari and Nowroozi (2022) provided a more inclusive categorization of the sources of stress and burnout during the pandemic in the form of the four categories of technological, pedagogical, affective, and administrative challenges. Poor digital literacy, online assessment concerns, and technological availability belong to the first category, while the pedagogical demands of online education including the change in roles and teaching approach belong to the second (Ghanbari & Nowroozi, 2022). Azadianbojnordi et al. (2021) exhibited that an overall positive attitude towards change, in general, can indirectly predict burnout levels during the pandemic, since it affects their acceptance of e-learning and virtual education. Overall, the transition to online education which took place over the past years had negative psychological consequences, especially among language teachers (Escudero-Nahón, 2021).

## 3. Methods

### 3.1. Design of the study

To delve into the differences in burnout levels of Iranian EFL teachers in PSs and PLIs during the pandemic and to examine the factors contributing to burnout, this study adopted a mixed-method design. In the quantitative phase of the study, Maslach et al.'s (1996) Burnout Inventory was distributed among EFL teachers using Google Docs. In the qualitative phase of the study, 20 EFL teachers, 10 from PSs and 10 from PLIs were interviewed.

### 3.2. Participants

In the quantitative part of the study, 268 EFL teachers completed the survey. The participants were 108 (40.3%) teachers teaching at public schools and 160 (59.7%) teachers teaching in private language institutes. Of 268 participants, 152 (56.7%) were female and 116 (43.3%) were male with the age range of 20 to 61 and a mean of 30.95. Regarding their academic degrees, 139 (51.9%) teachers had Bachelor's degrees, 117 (43.7%) had Master's degrees, and 12 (4.5%) were Ph.D. holders and students. Although 292 teachers participated in this study, the participants with missing values were excluded, thereby, leaving us with 268 participants.

From among the participants, 20 teachers (10 from PSs and 10 from PLIs) were selected for the qualitative part of the study to be interviewed using purposeful sampling and employing homogeneous and intensity strategies on the account of the richness of the information they provided (Creswell, 2012). Attempt was made to choose teachers of the same demographic information in pairs from PSs and PLIs to minimize the differential effects of demographic attributes on teachers' experience of burnout. In so doing, a pair of teachers (one from a PS and one from a PLI) was chosen with the same teaching experience, gender, academic degrees, and the age levels they taught. Moreover, care was taken to select pairs with different demographic information to ensure a diversity of factors which may affect teacher burnout. From the participants selected for the interview, 3 had experience teaching both in PSs and PLIs which helped take a comparative look at the issue in both contexts.

### 3.3. Instruments

#### 3.3.1.Maslach et al.'s (1996) burnout inventory

Maslach et al.'s (1996) Burnout Inventory is the most widely-used scale for measuring teachers' level of Burnout. It includes 22 statements, and participants are asked to rate each item on a seven-point Likert scale from Never (0) to Everyday (6). This

scale includes three sub-scales of "emotional exhaustion", "depersonalization", and "personal accomplishment". High scores on the first two components and low scores on the last one are considered as indications of burnout. In this study, the reliability coefficient of the inventory using Cronbach's alpha was 0.76. Moreover, pervious factor analytic studies (e.g. Schaufeli & Van Dierendonck, 1993) have confirmed the three dimensions of burnout as represented in the scale.

#### 3.3.2. Semi-structured interviews

In order to obtain a deeper understanding of the reasons that can cause teacher burnout in virtual teaching, semi-structured interviews were conducted with 20 EFL teachers, 10 from PSs and 10 from PLIs. The interviews were semi-structured allowing the researchers to raise further questions to clarify and probe the issue in case the respondents' answers were not adequate. Following the methodological procedures in other studies on teacher burnout (e.g. EI Helou, Nabhani & Bahous, 2016), the participant teachers were asked to describe their different activities inside and outside the online classes and what they did to cope. In addition, they were encouraged to share their feelings, perceptions, and thoughts regarding online education and the stressors they experienced in their teaching context which provided an understanding of the causes and origins of burnout in the contexts of the online teaching. The interviews were mainly conducted through social media due to the lockdown and in participants' mother tongue (Farsi). The interview data were then transcribed verbatim and analyzed using thematic analysis approach.

### 3.4. Procedure and data analysis

The quantitative data of this study were collected using Google Docs. Having obtained the EFL teachers' consent, the link to the questionnaire was sent to them. The qualitative data were obtained through semi-structured interviews using social media platforms considering the spread of the Covid-19 and the lack of in-person access to the participants. The quantitative data of the study were analyzed using SPSS. Both descriptive and inferential statistics (including Independent Sample T-tests) were used to answer the first research question. Regarding the qualitative data, Braun and Clarke's (2006) thematic analysis was used to analyze the interviews. The recorded voices sent via social media platforms were first transcribed verbatim and then the comments were read by the researchers to familiarize ourselves with the data. Simultaneously, notes of summaries were taken with the aim of obtaining the initial ideas for analysis and coding. Next, the initial codes were reviewed, merged and reduced to get the recurrent main themes and taxonomies. Having finalized the major themes, we did several re-readings to ensure that no pieces of information are left. Further, segments of verbatim were chosen to illustrate the themes emerged from the data. To ensure the trustworthiness of analysis, 30% of the data were analyzed by the second and the third researchers and the disagreements were resolved through discussion and negotiation.

## 4. Result

### 4.1. Result of the quantitative phase

In this phase of the study, Maslach et al.'s (1996) Burnout Inventory was distributed among 268 EFL teachers (108 teaching at public schools and 160 teaching at private language institutes). Table 1 shows the descriptive statistics for the two groups. To examine if the level of difference between the two groups was significant, an Independent Sample T-test was administered. As indicated in Table 2, there is a significant difference between the level of burnout that EFL teachers in PSs and PLIs experience in virtual education (P.<.05).

**Table 1. Descriptive Statistics of EFL Teachers' Burnout in PSs and PLIs.**

| | Context | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Teacher Burnout | Public Schools | 108 | 53.7685 | 15.40033 | 1.48190 |
| | Private Institutes | 160 | 63.9250 | 20.43291 | 1.61536 |

**Table 2. The Results of Independent Sample T-test of the Differences between Teachers' Burnout in PSs and PLIs.**

| | | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
|---|---|---|---|---|---|---|
| Burnout | Equal variances assumed | -4.391 | 266 | .000 | -10.15648 | 2.31303 |

In order to examine the differences between the two groups in the three components of burnout including "emotional exhaustion", "depersonalization", and "personal accomplishment", three more Independent Sample T-tests were administered. Table 3 shows the descriptive statistics for the three components of burnout for both groups. As indicated in the table, the mean difference between the two groups is higher for the two components of "Emotional Exhaustion" and "Personal Accomplishment", and in both of them, the private language institute teachers have a higher mean.

**Table 3. Descriptive Statistics of the Three Components of Burnout for EFL Teachers in PSs and PLIs.**

| | Context | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| Emotional Exhaustion | Public Schools | 108 | 17.6111 | 11.34409 | 1.09159 |
| | Private Institutes | 160 | 23.9813 | 14.28967 | 1.12970 |
| Depersonalization | Public Schools | 108 | 11.1481 | 6.16565 | .59329 |
| | Private Institutes | 160 | 10.1688 | 5.40931 | .42764 |
| Personal Accomplishment | Public Schools | 108 | 25.0093 | 9.27915 | .89289 |
| | Private Institutes | 160 | 29.8750 | 8.83158 | .69820 |

In order to examine if the mean differences between the two groups are significant, three Independent Sample T-tests were administered. As Table 4 portrays, there is a significant difference between the two groups in the two components of "Emotional Exhaustion", and "Personal Accomplishment" (P.<.05). However, the mean difference between the two groups in "Depersonalization" is not significant (P.>.05). It means that the private institute teachers have a significantly higher level of "Emotional Exhaustion", and "Personal Accomplishment" compared to the public school teachers.

**Table 4. The Results of Independent Sample T-test of the Differences between PS and PLI Teachers in the Three Components of Burnout**

| | | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
|---|---|---|---|---|---|---|
| Emotional Exhaustion | Equal variances assumed | -3.880 | 266 | .000 | -6.37014 | 1.64190 |
| Depersonalization | Equal variances assumed | 1.374 | 266 | .171 | .97940 | .71304 |
| Personal Accomplishment | Equal variances assumed | -4.334 | 266 | .000 | -4.86574 | 1.12261 |

## 4.2. Result of the qualitative phase

The analysis of the qualitative data pointed to some reasons for teacher burnout in online teaching contexts including the hassles with online teaching, lack of teaching effectiveness, lack of collegiality and support, financial issues and low wages, lack of job security, student misbehavior and lack of motivation, lack of teacher autonomy, heavy workload, and workplace context which will be discussed below.

### 4.2.1. Online teaching challenges and obstacles

Part of the reasons for teacher burnout in both PSs and PLIs concerned the issues related to the online teaching. Many factors such as teachers' lack of familiarity with online teaching platforms especially at the beginning of the Covid-19 pandemic, poor internet connection and internet coverage, the loss of teacher-student interaction, lack of students' engagement, student evaluation, and health issues because of working with digital tools for long hours created occupational burnout for EFL teachers in both contexts. Even some teachers, especially the older ones, preferred to quit the profession temporarily until in-person classes are resumed. Here are some excerpts of the participants' ideas.

Excerpt 1

> "After so many years of teaching in-person classes, it has been very disparate for me to shift to online teaching especially at the beginning of the pandemic. I simply couldn't get used to it…. The reasons are mainly my lack of familiarity with online teaching and the poor internet connection". (*Interviewee 9, PS*)

Excerpt 2

> "Although I am familiar with teaching online, unfortunately, many students either do not have access to the internet or are not willing to participate in classroom activities. Even for those attending the class, I am not sure if all of them are really following the class". (*Interviewee 3, Ps and PLI*)

Excerpt 3

> "There is not much interaction between the students and the teacher. Many of the students are reluctant to turn on their webcams or mics and be involved in classroom activities. So, I should talk more in the class which is against the norms of a language class". (*Interviewee 6, PLI*)

Excerpt 4

> "The main issue with my online classes is evaluating students. Because of its limitations, I think online testing is not very reliable. Sometimes you see that the answer to the same question is written by many students using the same wording showing that they may be sharing the answers in online platforms". (*Interviewee 1, PLI*)

Excerpt 5

> "Because of teaching online classes for a long time, unfortunately, I have a very bad neck pain. So, I have canceled all of my online classes". (*Interviewee 4, PLI*)

### 4.2.2. Concerns about teaching effectiveness

Another factor causing teacher burnout was the feeling that their teaching was ineffective. This was more evident among public school teachers who believed that due to reasons such as heterogeneous and crowded classes, time limitations, lack of technological facilities, lack of teaching materials, and students' absenteeism from the online classes as a result of the lack of access to the internet or digital devices, their teaching was not very effective. This is in line with the results of the quantitative phase of the study showing that public school teachers had a significantly lower level of "personal accomplishment". Here are some excerpts to illustrate the point:

Excerpt 6

> "The main reason causing burnout for me is the feeling that my teaching is not effective. I consider myself a competent person who can make positive changes in society. But, because of the limitations of online classes, I think my teaching is almost a waste of time". (*Interviewee 2, PS*)

Excerpt 7

> "Because of the class size, heterogeneous students, lack of possibility to involve all the students in productive tasks, and the problems with online teaching platform, my teaching is not as effective as it should be". (*Interviewee 7, PS*)

Excerpt 8

> "Compared with my experience of teaching in private language institutes, the language proficiency level of my students in public schools do not improve due to time limitations and many other factors". (*Interviewee 6, PS and PLI*)

#### 4.2.3.Lack of collegiality and principal/manager support

One of the factors causing stress for EFL teachers in both contexts was the lack of a good relationship among colleagues and principal/manager support. This was specially the case in online classes where teachers were left on their own to find solutions to their problems through trial and error and the opportunities for professional development and growth were few and far between considering the teachers fell physically apart and felt professionally isolated. This point is indicated in the following excerpts:

Excerpt 9

> "The working condition in my institute is very problematic. There is a fierce negative competition among teachers. We don't often share the issues which arise in our profession with each other. If I can find another institute, I'd prefer to leave here". (*Interviewee 6, PS and PLI*)

Excerpt 10

> "The institute manager expects too much from the teachers. He expects us to be available online to answer students' questions all the time and give them feedback which is beyond what I can do. Unfortunately, he prefers punishing teachers to encouraging them". (*Interviewee 10, PLI*)

Excerpt 11

> "The problems that I have with my colleagues started from a classroom observation which is a very routine practice in my institute. Unfortunately, one of them tried to highlight the weaknesses of my practice and ignore my strengths. Since that time, I do not have a good attitude toward classroom observation". (*Interviewee 1, PLI*)

Excerpt 12

> "In online courses, especially the beginning days, we were left on our own to run the classes. As teachers, we were not connected since asking for help meant raising questions on our own competencies. In the cases where teachers were comfortable asking each other their questions, little help received as we were all perplexed with this new online teaching." (*Interviewee 3, PS and PLI*)

#### 4.2.4.Financial concerns and low wages

One of the main problems causing teacher burnout in both contexts was the financial problems and low wages. In fact, for the majority of teachers, this was the main reason leading them to leave the profession, in case they could find another job. Considering the high inflation rate in the country, many teachers believed that they could not afford to pay for their expenses without working extra hours and having another part-time job. The excerpts show this point:

Excerpt 13

"The wages are very low in the language institute that I am teaching. I am still single and live with my parents. I am sure I wouldn't be able to afford family expenses with this salary".(*Interviewee 6, PLI*)

Excerpt 14

"Considering the high inflation and low salary of the teaching profession, if I can find another job with a better salary, for sure I'll prefer to leave the teaching profession, despite my love for it". (*Interviewee 4, PLI*)

Excerpt 15

"In order to afford my expenses, I need to work from the morning till evening. I know that it can harm my health and family relationships, but I have to do that. I know it might cause burnout after a while". (*Interviewee 3, PS and PLI*)

#### 4.2.5.Lack of job security

The issue of job security was a concern for teachers working in PLIs or those working in PSs based on an annual contract. The fact that they may lose their job due to various reasons caused them extra stress as indicated in the following excerpts:

Excerpt 16

"I work in a language institute and I receive a monthly salary, but I know that at any time I may lose my job if they can find another teacher with fewer expectations". (*Interviewee 6, PLI*)

Excerpt 17

"One of the main problems of working in private language institutes is that you usually do not have a fixed salary, insurance, retirement pension, etc. So, if teachers have the chance to find another job, they usually leave the teaching profession". (*Interviewee 5, PLI*)

Excerpt 18

"In public schools, we have a monthly salary and we have job security which is good, though the salary is not enough. But in our school, there are teachers teaching based on an annual contract; I know that their stress and dissatisfaction with the teaching profession is much more than that of us". (*Interviewee 3, PS and PLI*)

#### 4.2.6.Students' misbehavior and lack of motivation

Although, this category is not specific to online classes, teachers complained about how student misbehavior and lack of motivation caused them discomfort in both contexts. Some teachers believed that despite high expectations from parents, many students are not motivated enough to study and learn. In online classes, the problem of students' demotivation was deteriorated since they lost their social networks and were isolated. Moreover, they could make excuses with regard to the internet and their technical failures and evade their duties and coursework more easily. Here are some excerpts of teachers' comments:

Excerpt 19

"My students, who are teenagers, do not have enough motivation to learn. Their parents force them to participate in the classes. Also, some of them are very noisy. So, a significant portion of the class time should be spent on classroom management which is very stressful for me". (*Interviewee 5, PLI*)

Excerpt 20

In online classes, students frequently make excuses. When I ask them to turn on their mics or webcams and share their ideas or responses, they pretend that their mics and cameras don't work. Or, when they are not following the class and I call them out, they use pretexts like, I lost my connection or my device had got frozen I couldn't answer." (Interviewee *9, PS).*

Excerpt 21

> "Generally speaking, my students are not motivated enough to study. This can adversely influence my teaching effectiveness". (*Interviewee 9, PS*)

#### 4.2.7. Lack of teacher autonomy

The lack of teacher autonomy was another factor leading to teacher burnout. Some teachers believed that they needed to follow exactly what the institute and school dictated to them and there was no scope for maneuvering or applying their creativity in teaching pedagogy as indicated in the following excerpts:

Excerpt 22

> "In our language institute, you have to follow a fixed schedule given to you at the beginning of the semester. So, the theoretical issues we studied about creativity in teaching and making decisions based on the contextual variables and individual characteristics cannot come in handy". (*Interviewee 1, PLI*)

Excerpt 23

> "During the online classes, it has happened to us a couple of times that the platform specified for teaching failed. However, since the class archives were checked, we were not allowed to use an alternative platform." (*Interviewee 7, PS*)

#### 4.2.8. Heavy workload and obligations

Some teachers especially those teaching in private language institutes believed that heavy workload caused many problems for them. The burden was reported to be added in the online courses with many teachers having children at home whose academic responsibilities were assigned to their parents. In addition, being obliged to share the devices with their children to get connected to their classes which were often at the same of their parents' classes created further hassles.

Excerpt 24

> "I teach for long hours each day. I have to teach both my students and my kids at home. We often share the same cellphone for online classes and this has created me some problems as my kids classes and mine are often at the same time.". (*Interviewee 7, PS*)

Excerpt 25

> "With the online classes that I have, I need to teach long hours and after the class, I should mark my students' homework on WhatsApp or other platforms. In addition, my students ask questions at any time they want and I am supposed to answer them, so I do not have any time to rest" (*Interviewee 4, PLI*)

#### 4.2.9. School/institute policies

The different policies adopted by PSs and PLIs were often a source of fatigue and boredom for teachers driving them to experience extra burden and stress. One such policy was put forward for PSs in rural areas during the pandemic where students did not have access to the internet and technical devices necessary for online classes. In such areas, the teachers were obliged to commute to the areas to attend the in-person classes despite the outbreak and the associated health risk (*interviewee 2, PS*). Another such policy taken by language institutes was running hybrid classes to ensure maximum enrollment in classes. That is, the class was run both face-to-face for those attending the class in-person and virtually for those students who were online. Managing two groups of students at the same time was also believed to be tiresome and demanding and required the teachers to exert themselves even more (*interviewee 5, PLI*).

## 5. Discussion

The purpose of this study was to compare Iranian EFL teachers' level of burnout during the Covid-19 pandemic in PSs and PLIs and to investigate the factors causing teacher burnout in both contexts. To this aim, quantitative and qualitative data were collected and analyzed using Maslach et al.'s (1996) Burnout Inventory and semi-structured interviews respectively.

With regard to the first research question, the findings in the quantitative phase of the study exhibited a significantly higher level of Emotional Exhaustion and Personal Accomplishment among PLI teachers compared to their PS counterparts. Evers et al. (2004) suggested considering Emotional Exhaustion as the first evidence of experienced burnout. The difference between the experienced burnout in the three dimensions can be justified considering the findings of the qualitative phase. Higher levels of certain aspects of burnout among PLI teachers are partially in line with the findings of Soleimani and Bolourchi (2021) and Hosseini Fatemi and Raoufi (2014) who observed a higher overall burnout experience among private language institute teachers. the findings are, however, in contrast to those of Heriran and Navidinia's (2015) and Ferreira and Martinez's (2012) which indicated a higher burnout level among public school teachers in studies conducted prior to the pandemic.

One significant contribution of our study was revealing the intricate differences in the experience of burnout between public school and private language institute teachers. While previous studies report an overall difference between the levels of burnout among PS and PLI teachers, our findings demonstrated a more complex picture of their burnout experience. Accordingly, PS teachers reported higher levels of Depersonalization whereas PLI teachers stated higher levels of Emotional Exhaustion and Personal Achievement. Previously, Ferreira and Martinez (2012) revealed that PS teachers reveal higher levels of both emotional exhaustion and Depersonalization, while there was no significant difference with regard to personal accomplishment. Although it is partially in line with the findings of this study with regard to the level of depersonalization, it contradicts our findings in the other two subscales of burnout. Furthermore, echoing our findings, Nikolaos et al. (2006) reported higher levels of Emotional Exhaustion among private sector early educators compared to public school teachers.

Prior research (e.g. Ghanbari & Nowroozi, 2022; Hassani, 2021; Malekolkalami, 2020; Shobeiry, 2022; Soleimani & Bolourchi, 2021) has highlighted some challenges facing teachers in remote education including promoting student engagement, lack of suitable material and proper online assessment methods, digital literacy and access, as well as disruptive student behavior. These challenges could result in decreased levels of teaching effectiveness among EFL teachers. For instance, in the context of public schools, running large and heterogeneous classes can cause teachers classroom management challenges which directly affect the quality of teaching, thereby prompting lower personal achievement scores.

Additionally, being positioned in a rigid environment that restricts teachers' autonomy and dictates certain behavior and schedules can foster the development of teacher burnout as it fosters the perceptions of unmet goals and a lack of professional development (Grayson & Alvarez, 2008). Teachers' professional beliefs shape their teaching practice and help them organize their teaching according to what they believe to be appropriate (Salahshour & Esmaeili, 2021). Limiting their creativity and scope of maneuvering through strict and rigid policies and regulations can lead the way to burnout according to a number of studies (Khezerlou, 2017; Salahshour & Esmaeili, 2021; Soleimani & Bolourchi, 2021; Van Droogenbroeck et al., 2021).

On the other hand, exhausted from online classes, the teacher is surrounded by demotivated, burnt-out, and often competitive peers who fail to provide the much-needed emotional and professional support. This lack of support network coupled with demanding and criticizing parents and authorities exacerbates the feeling of low personal achievement and ineffectiveness which might already be present due to the novelty of online teaching experience and lack of familiarity with the new environment and the required skills.

An amicable work climate can help teachers not only to maintain a realistic view of their teaching circumstances but also aid them to manage the stress and challenges of online education. Bahrami and Moradkhani (2019) support this, confirming the significance of support from peers and principals. By the same token, van Emmerik (2002) supports the role of departmental support and peer assistance in reducing emotional exhaustion. It is also in accordance with the findings of Moradkhani and Shirazizadeh (2017) who emphasized the crucial role of collegial support, stating lack of appropriate rapport can be demotivating to teachers. The presence of such collegiality and collaboration not only fosters but also entails the existence of an atmosphere of professionalism, the value of which cannot be undermined in the prevention of burnout. As Seyedjalali et al. (2021) stress professional behavior in administrators necessitates creating an open and accepting environment for instructors (Seyedjalali et al., 2021).

The issue of demotivation is another culprit in both contexts. For example, in PLIs, on the one hand, teachers are demotivated due to the lack of job security and an imbalance between the demands of the job and the payment (Pishghadam & Sahebjam, 2012). In PSs, on the other hand, teachers often enjoy a secure job and payment which can compensate for the negative factors present in this context at least to some extent, though students may lack the necessary intrinsic motivation. In any circumstances, however, having a secure job can be a driver in the prevention of burnout (Bahrami and Moradkhani, 2019).

Understandably, demotivated students can create a host of problems in the class, complicating class management for the teacher. Furthermore, a combination of the aforementioned factors has devastatingly increased teachers' workload leading to higher burnout levels, especially emotional exhaustion. In a similar vein, an increase in workload drives teachers to experience increased mental stress and frustration (Ferdous & Shifat, 2020; Shobeiry, 2022) and deprives them of the time to be spent with the family and the support that can be provided by them. Studies show that those who benefit from family and friends' support also benefit from better mental and emotional health which can prevent burnout (Bahrami & Moradkhani, 2019). Ferdous and Shifat (2020) along with Honarzad (2022), and Van Droogenbroeck et al. (2021) corroborate the effect of

increased workload and time pressure in the development of burnout. Although increased workload cannot independently aggravate burnout, concerns about the fit between working hours and personal affairs may have a stronger effect (Barnett et al., 1999).

## 6. Conclusion

Considering the importance of teachers in the success or failure of educational systems, it is incumbent upon policymakers to pay more attention to the stressors causing teacher burnout. Increasing the technological literacy of teachers, creating a more teacher-friendly environment, decreasing teachers' workload, and increasing their payment can be considered as some ways that can alleviate the problems and prevent burnout. In addition, considering the lack of effectiveness of current teacher professional development programs in the Iranian context (Navidinia, 2021), more need-responsive teacher professional development programs are required to prepare teachers and arm them with the necessary coping strategies to prevent burnout.

This study has some limitations. Apart from the relatively small sample size particularly, in the interview section, the data collected by administering Maslach et al.'s (1996) Burnout Inventory were based on teachers' self-explanatory perceptions which might affect the generalizability of the findings. Despite its limitations, the study has gone some way in improving our knowledge of the status and causes of burnout among Iranian EFL teachers in the milieu of online education which is one the first studies of this kind.

Given the importance of teacher burnout, it is hoped that future lines of research delve more deeply into causes and origins of burnout in other contexts. The interplay of contextual and personal factors causing burnout, for instance, is one of the related area that future researchers can address. Furthermore, identifying some effective and context-sensitive strategies that can help teachers cope with job stressors is another area for further re

## 7. Disclosure statement

The authors have no conflicts of interests to declare.

## 8. References

Azadianbojnordi, M., Bakhtiarpour, S., Hatami, M., Sadimajd, G., Khelghati, A., Bakhtiarpour, S., Bakhtiarpour, M., Bakhtiarpour, Z., & Bakhtiari, M. (2021). The Relationship between Teachers' Attitudes Toward Change and Job Burnout in the period of the Covid-19 Pandemic with the Mediating Role of Virtual Education Acceptance. *Journal of Research in Educational Science*, *15*(55).

Bahrami, M., & Moradkhani, S. (2019). A comparison of public school and private institute EFL teachers' burnout. *Technology of Education Journal (TEJ)*, *13*(3), 626-637. https://doi.org/10.22061/jte.2018.3542.1890

Barnett, R. C., Gareis, K. C., & Brennan, R. T. (1999). Fit as a mediator of the relationship between work hours and burnout. *Journal of Occupational Health Psychology*, *4*(4), 307.

Bartosiewicz, A., Łuszczki, E., Zaręba, L., Kuchciak, M., Bobula, G., Dereń, K., & Król, P. (2022). Assessment of job satisfaction, self-efficacy, and the level of professional burnout of primary and secondary school teachers in Poland during the COVID-19 pandemic. *PeerJ*, *10*, e13349.

Barutçu, E., & Serinkan, C. (2013). Burnout syndrome of teachers: An empirical study in Denizli in Turkey. *Procedia-Social and Behavioral Sciences*, *89*, 318-322.

Bayani, A. A., Bagheri, H., & Bayani, A. (2013). Influence of gender, age, and years of teaching experience on burnout. *Annals of Biological Research*, *4*(4), 239-243.

Betoret, F. D. (2009). Self-efficacy, school resources, job stressors and burnout among Spanish primary and secondary school teachers: a structural equation approach. *Educational Psychology*, *29*(1), 45-68.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Bravo, A., Buenaflor, N., Baloloy, J., Guarte, L., Osinaga, A., Salartin, A., & Tus, J. (2021). Amidst the COVID-19 Pandemic: The Job Burnout and Job Satisfaction of Public School Teachers in the Philippines. *International Journal Of Advance Research And Innovative Ideas In Education*.

Brewer, E. W., & Shapard, L. (2004). Employee Burnout: A Meta-Analysis of the Relationship Between Age or Years of Experience. *Human Resource Development Review*, *3*(2), 102-123. https://doi.org/10.1177/1534484304263335

Brissie, J. S., Hoover-Dempsey, K. V., & Bassler, O. C. (1988). Individual, situational contributors to teacher burnout. *The Journal of Educational Research*, *82*(2), 106-112.

Capel, S. A. (1987). The incidence of and influences on stress and burnout in secondary school teachers. *British Journal of Educational Psychology*, *57*(3), 279-288.

Carson III, R. L. (2006). *Exploring the episodic nature of teachers' emotions as it relates to teacher burnout*. Purdue University.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of education review*, *26*(6), 673-682.

Creswell, John, W. (2012). *Educational Research: Planning, conducting, and evaluating*

*quantitative and qualitative research (4th ed)*. Boston: Pearson Education, Inc.

Edú-Valsania, S., Laguía, A., & Moriano, J. A. (2022). Burnout: A Review of Theory and Measurement. *Int J Environ Res Public Health*, *19*(3). https://doi.org/10.3390/ijerph19031780

El Helou, M., Nabhani, M., & Bahous, R. (2016). Teachers' views on causes leading to their

burnout. *School leadership & management*, *36*(5), 551-567.

Escudero-Nahón, A. (2021). A meta-synthesis regarding the educational narrative during the COVID-19 Pandemic. *Diálogos sobre educación. Temas actuales en investigación educativa*, *12*(22).

Evers, W. J., Tomic, W., & Brouwers, A. (2004). Burnout among teachers: Students' and teachers' perceptions compared. *School Psychology International*, *25*(2), 131-148.

Farshi, S. S., & Omranzadeh, F. (2014). The effect of gender, education level, and marital status on Iranian EFL teachers' burnout level. *International Journal of Applied Linguistics and English Literature*, *3*(5), 128-133.

Ferdous, A. U., & Shifat, N. F. (2020). Dealing with mental health in online learning: A retrospect on ELT teachers and EFL learners during COVID-19 pandemic. *REiLA: Journal of Research and Innovation in Language*, *2*(3), 101-107.

Ferreira, A. I., & Martinez, L. F. (2012). Presenteeism and burnout among teachers in public and private Portuguese elementary schools. *The International Journal of Human Resource Management*, *23*(20), 4380-4390. https://doi.org/10.1080/09585192.2012.667435

Foley, C., & Murphy, M. (2015). Burnout in Irish teachers: Investigating the role of individual differences, work environment and coping factors. *Teaching and Teacher Education*, *50*, 46-55.

Freudenberger, H. J. (1974). Staff Burn-Out. *Journal of Social Issues*, *30*(1), 159-165. https://doi.org/https://doi.org/10.1111/j.1540-4560.1974.tb00706.x

Ghanbari, N., & Nowroozi, S. (2022). Iranian EFL Teachers' Challenges and Coping Strategies During COVID-19 Pandemic: A Case Study. *The Qualitative Report*, *27*(3), 605-625.

Ghanizadeh, A., & Jahedizadeh, S. (2016). EFL teachers' teaching style, creativity, and burnout: A path analysis approach. *Cogent Education*, *3*(1), 1151997.

Gohm, C. L., Corser, G. C., & Dalsky, D. J. (2005). Emotional intelligence under stress: Useful, unnecessary, or irrelevant? *Personality and individual differences*, *39*(6), 1017-1028.

González-Morales, M. G., Rodríguez, I., & Peiró, J. M. (2010). A longitudinal study of coping and gender in a female-dominated occupation: Predicting teachers' burnout. *Journal of Occupational Health Psychology*, *15*(1), 29.

Grayson, J. L., & Alvarez, H. K. (2008). School climate factors relating to teacher burnout: A mediator model. *Teaching and Teacher Education*, *24*(5), 1349-1363.

Green, A. E., Albanese, B. J., Shapiro, N. M., & Aarons, G. A. (2014). The roles of individual and organizational factors in burnout among community-based mental health service providers. *Psychological services*, *11*(1), 41.

Hassani, V. (2021). The Impacts of COVID-19 Pandemic on English Language Teacher Education in Iran: Challenges and Opportunities. *Journal of Teaching Language Skills*, *40*(3), 83-116.

Haycock, K. (1998). Good teaching matters... a lot. *OAH Magazine of History*, *13*(1), 61-63.

Heiran, A., & Navidinia, H. (2015). Private and public EFL teachers' level of burnout and its relationship with their emotional intelligence: A comparative study. *International Journal of English Language & Translation Studies*, *3*(3), 1-10.

Hock, R. R. (1988). Professional burnout among public school teachers. *Public personnel management*, *17*(2), 167-189.

Honarzad, R. (2022). An assessment of online education during the COVID-19 pandemic: A survey of instructors in Iran. *AJELP: Asian Journal of English Language and Pedagogy*, *10*(1), 79-93.

Hosseini Fatemi, A., & Raoufi, R. (2014). Burnout and teaching style among Iranian English language educators in public schools and private institutes: A cross-comparison study. *International Journal of Research Studies in Language Learning*, *17*.

Huberman, A. M., & Vandenberghe, R. (1999). Burnout and the teaching profession.

Izquierdo, J., Sandoval Caraveo, M. d. C., De la Cruz Villegas, V., & Zapata Díaz, R. (2021). University Language Instructors' Preparedness for TechnologyMediated Instruction and Burnout During the COVID-19 Pandemic. *Íkala, Revista de Lenguaje y Cultura*, *26*(3), 661-695.

Jackson, S. E., Schwab, R. L., & Schuler, R. S. (1986). Toward an understanding of the burnout phenomenon. *Journal of applied psychology*, *71*(4), 630.

Kariou, A., Koutsimani, P., Montgomery, A., & Lainidi, O. (2021). Emotional Labor and Burnout among Teachers: A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(23), 12760. https://www.mdpi.com/1660-4601/18/23/12760

Kaya, N. (2022). EFL teacher's anxiety towards educational technologies in relation to teacher burnout during Covid-19 pandemic.

Khezerlou, E. (2017). Professional Self-Esteem as a Predictor of Teacher Burnout across Iranian and Turkish EFL Teachers. *Iranian Journal of Language Teaching Research*, *5*(1), 113-130.

Kimsesiz, F. (2019). The effect of school type on EFL teachers' burnout: The case in Turkey. *Journal of Language and Linguistic Studies*, *15*(4), 1413-1425.

Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers.

*British journal of educational psychology*, *77*(1), 229-243.

Malekolkalami, M. (2020). The Perception of Iranian Teachers on Online Teaching Using Digital Carrier During the COVID-19 Pandemic. *International Journal of Digital Content Management*, *1*(1), 109-126.

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, *2*(2), 99-113. https://doi.org/https://doi.org/10.1002/job.4030020205

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). The maslach burnout inventory-test manual. *Palo Alto*.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sternberg, R. J. (2000). Models of emotional intelligence. *JD Mayer*.

Mérida-López, S., & Extremera, N. (2017). Emotional intelligence and teacher burnout: A systematic review. *International Journal of Educational Research*, *85*, 121-130. https://doi.org/https://doi.org/10.1016/j.ijer.2017.07.006

Moradkhani, S., & Shirazizadeh, M. (2017). Context-based variations in EFL teachers' reflection: The case of public schools versus private institutes in Iran. *Reflective Practice*, *18*(2), 206-218.

Motallebzadeh, K., Ashraf, H., & Yazdi, M. T. (2014). On the relationship between Iranian EFL teachers' burnout and self-efficacy. *Procedia-Social and Behavioral Sciences*, *98*, 1255-1262.

Mousavy, S., & Nimehchisalem, V. (2014). Contribution of gender, marital status, and age to English language teachers' burnout. *Advances in language and literary studies*, *5*(6), 39-47.

Navidinia, H. (2021). EFL teachers as lifelong learners: Rethinking the link between teacher evaluation and in-service professional development. *International Journal of Language Studies*, *15*(3), 107-126.

Navidinia, H. & Heiran (2017). Exploring the demographic and contextual factors influencing English language teachers' burnout: Implications for teacher education programs. *Asean Journal of Teaching and Learning in Higher Education, 9 (1), 1-19.*

Nikolaos, T., Evridiki, Z., & Vasilios, G. (2006). Job satisfaction and burnout among Greek early educators: A comparison between public and private sector employees. *Educational Research and Reviews*, *1*(8), 256-261.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational evaluation and policy analysis*, *26*(3), 237-257.

Panisoara, I. O., Lazar, I., Panisoara, G., Chirca, R., & Ursu, A. S. (2020). Motivation and continuance intention towards online instruction among teachers during the COVID-19 pandemic: The mediating effect of burnout and technostress. *International Journal of Environmental Research and Public Health*, *17*(21), 8002.

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cures*. Free press.

Pishghadam, R., & Sahebjam, S. (2012). Personality and Emotional Intelligence in Teacher Burnout. *The Spanish journal of psychology*, *15*(1), 227-236. https://doi.org/10.5209/rev_SJOP.2012.v15.n1.37314

Pressley, T. (2021). Factors Contributing to Teacher Burnout During COVID-19. *Educational Researcher*, *50*(5), 325-327. https://doi.org/10.3102/0013189x211004138

Purvanova, R. K., & Muros, J. P. (2010). Gender differences in burnout: A meta-analysis. *Journal of Vocational Behavior*, *77*(2), 168-185. https://doi.org/https://doi.org/10.1016/j.jvb.2010.04.006

Roohani, A., & Dayeri, K. (2019). On the Relationship between Iranian EFL Teachers' Burnout and Motivation: A Mixed Methods Study. *Iranian Journal of Language Teaching Research*, *7*(1), 77-99.

Rostami, S., Ghanizadeh, A., & Ghapanchi, Z. (2015). A study of contextual precursors of burnout among EFL teachers. *International Journal of Research Studies in Psychology*, *4*(1), 13-24.

Salahshour, F., & Esmaeili, F. (2021). The causes of burnout among Iranian EFL Academics: A case study. *Iranian Journal of Language Teaching Research*, *9*(1), 123-137.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, cognition and personality*, *9*(3), 185-211.

Schaufeli, W., & Enzmann, D. (1998). *The burnout companion to study and practice: A critical analysis*. CRC press.

Seyedjalali, K., Daneshvar, Z., Khadivi, A., & Talebi, B. (2021). Structural analysis of teachers' burnout based on professional development and participatory management in school. *Journal of School Administration*, *9*(2), 72-51.

Schaufeli, W.B. & Van Dierendonck, D. (1993). the construct validity of two burnpur measures. *Journal of Organizational Behvaior*, *14*, 631-647.

Shobeiry, M. (2022). Covid-19 and analysis of needs of the primary school teachers in online-schooling in Iran. *Revista on line de Política e Gestão Educacional*, e022050-e022050.

Soleimani, M., & Bolourchi, A. (2021). EFL Teachers' Job Satisfaction and Burnout in Iran. *PASAA: Journal of Language Teaching and Learning in Thailand*, *61*, 229-252.

Vaezi, S., & Fallah, N. (2011). The Relationship between Emotional Intelligence and Burnout among Iranian EFL Teachers. *Journal of Language Teaching & Research*, *2*(5).

Van Droogenbroeck, F., Spruyt, B., Quittre, V., & Lafontaine, D. (2021). Does the School Context Really Matter for Teacher Burnout? Review of Existing Multilevel Teacher Burnout Research and Results From the Teaching and Learning International Survey 2018 in the Flemish- and French-Speaking Communities of Belgium. *Educational Researcher*, *50*(5), 290-305. https://doi.org/10.3102/0013189X21992361

van Emmerik, I. H. (2002). Gender differences in the effects of coping assistance on the reduction of burnout in academic staff. *Work & Stress*, *16*(3), 251-263.

Vargas Rubilar, N., & Oros, L. B. (2021). Stress and burnout in teachers during times of pandemic. *Frontiers in Psychology*, *12*, 756007.

Yastibas, A. E. (2021). Comparing the Burnout-Levels of Turkish EFL Instructors at State and Private Universities Depending on Several Variables. *i-Manager's Journal on Educational Psychology*, *14*(4), 34.

Zhaleh, K., Ghonsooly, B., & Pishghadam, R. (2018). Effects of conceptions of intelligence and ambiguity tolerance on teacher burnout: A case of Iranian EFL teachers. *Journal of Research in Applied Linguistics*, *9*(2), 118-140.

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 50-66*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2022.5258.1001

# A Survey on Viewing Preferences and Habits of Iranian Audience of Audiovisual Translation

Saeed Ameri[1]
Masood Khoshsaligheh[2]

[1]Ph.D. in Translation Studies, Department of English, Ferdowsi University of Mashhad, Iran
[2]Full Professor in Translation Studies, Department of English, Ferdowsi University of Mashhad, Iran

**ARTICLE HISTORY**

Received: 14 June 2022
Revised: 14 October 2022
Accepted: 21 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: s.ameri@mail.um.ac.ir

**ABSTRACT**

The remarkable popularity and success of foreign feature films and television series as a source of entertainment cannot be disputed. The emergence of amateur subtitling communities has led to dramatic changes in individuals' viewing habits and styles as fansubbers offer an ideal opportunity for people to access international audiovisual materials, especially in dubbing countries like Iran. However, empirical research on audience viewing preferences and habits in terms of audiovisual translation is scarce. To bridge the gap, over 1200 Iranian viewers filled in an online questionnaire. The results suggest the emergence of new habits and preferences as proportionately more Iranians nowadays access international films and TV series with Persian subtitles, at least among young adults. Also, a small number of the respondents prefer to watch foreign films and programs dubbed. Additionally, there was a weak interplay between viewers' English proficiency and their tendency towards watching dubbed or subtitled programs, and national TV channels do not appear to be the primary entertainment source when it comes to international cinematic products. Understanding viewers' preferences and viewing styles forges possible future paths for the translation industry to cater for the needs of individuals with different viewing styles and needs.

**KEYWORDS:** Audience; Audiovisual translation; Dubbing; Foreign and domestic productions; Subtitling; Viewing habits and preferences

## 1. Introduction

Dubbing is deep-rooted in Iran and its history can be traced back to many decades ago. While the current Iranian governmental policies still resist the introduction of professional subtitling on Television, considerable innovation in recent years in terms of audiovisual material consumption has put forward new solutions to remedy this problem, mirroring the active role of the audience in creating their own materials. In the past, the audience had to wait until their favorite program being broadcast on Television or screens in cinemas. But now online translators work around the clock to offer the quickest release of the translation for foreign cinematic products (Khoshsaligheh et al. 2020) even though this activity has given rise to ethical issues, such as the breach of copyright and film piracy (Orrego-Carmona, 2018). The foreign programs, especially the fictional genres are pirated and widely shared and distributed across the Internet (Dwyer, 2017, p. 2019). Illegal distribution of foreign cinematic production has been very frequent in Iran due to the lack of copyright laws there (Zeydabadi-Nejad, 2016). This comfortable and inexpensive access to original materials has brought about a change in people's behavior towards cinematic productions; from being a passive receiver to an active producer of such products (Pérez-González, 2014; Orrego-Carmona & Lee, 2017).

**Cite this article:** Ameri, S., & Khoshsaligheh, M. (2023). A survey on viewing preferences and habits of Iranian Audience of Audiovisual Translation. *Applied Linguistics Inquiry*, *1*(1), 50-66. doi: 10.22077/ali.2022.5258.1001

As far as the translation of foreign programs in Iran is concerned, before the introduction of Iranian online video-on-demand (VOD) services, such as *https://www.filimo.com*, there was always much delay in broadcasting and screening the dubbing of foreign films and TV series. The Iranian broadcasting center, IRIB, is virtually lagging behind, on account of reluctance to dub new releases. Some programs are dubbed and shown with a several-year delay. Some people, especially film buffs, are unlikely to put up with a grueling dubbing or subtitling release schedule where the delayed screen of a foreign product is expected as is the case with the Iranian broadcasting center, IRIB (Ameri & Khoshsaligheh 2018). We are therefore witnessing the emergence of translations outside the professional and official circles in such a way that the border of the translation practice has been widened to embrace new forms, and consequently, this has opened up many opportunities for viewers (Khoshsaligheh et al., 2020).

According to Díaz Cintas, "the way in which viewers can consume audiovisual productions has been revolutionized" because viewers "who are now in the driving seat" have the chance to decide on how to watch their favorite audiovisual programs (Bogucki & Díaz Cintas 2020, p. 15). Therefore, one can claim there have been some shifts towards subtitling in dubbing countries, especially among the younger users. Chaume (2019) considers it a new way of choosing a given form of audiovisual translation (AVT) because, unlike what was common in dubbing countries in the past, now the audience is not necessarily obliged to consume only dubbed products and they have more access to subtitled programs. This generally means that AVT choice is no longer nationally defined but should be understood from an individual perspective, stressing individual variations among viewers (Pedersen, 2018). Empirically speaking, a few researchers have reported that people, in particular the youth, in dubbing countries have begun to watch subtitled programs (Matamala, Perego, & Bottiroli, 2017; Matamala & Ortiz-Boix, 2018, Orrego-Carmona, 2014; Enríquez-Aranda & García Luque, 2018). This is likely because there is a relationship between audience age and their language competency, as the youngsters are more competent in languages; therefore, they are more interested in subtitled programs than dubbed ones (MECD, 2012).

This should be noted that some scholars argue that it is rather difficult for individuals accustomed to a particular AVT modality, say dubbing, to change their viewing preference and habit (Antonini & Chiaro, 2009) just because individuals "are creatures of habit" (Ivarsson as cited in Antonini & Chiaro 2009, p. 97). For Díaz-Cintas, tradition is not the only decisive factor as many other factors, such as age and language mastery, have a role to play (Bogucki & Díaz Cintas, 2020, p. 15). The discussion above therefore takes us to believe that the research objectives should be complemented with such variables as language proficiency to figure out if habits and preferences may also be driven by this variable. Especially significant is how viewers access cinematic programs, which can provide insights into the popularity of old and new multimedia service providers.

Against this backdrop, this paper queries Iranian viewers' AVT habits and preferences, drawing on an online survey. The empirical evidence of this investigation will have implications for modifying the current norms and polices of the Iranian AVT industry, which has been long known to ignore audience needs and habits. The questions, which guide this study are:

- To what extent do Iranian viewers watch domestic and foreign productions?
- To what extent do Iranian viewers watch dubbed and subtitled productions?
- Is there an association between viewers' command of English and their tendency to watch dubbed and subtitled productions?
- How do Iranian viewers prefer to access foreign productions?
- What are Iranian viewers' preferences concerning a film's country of origin?

## 2. Theoretical framework

### 2.1. The changing landscape of audiovisual translation

The introduction of talkies or sound films in the European society of the 1930s resulted in the emergence of two crucial translation types, which facilitated the circulation of foreign films around the world. These translation types were called dubbing and subtitling, where the former replaces the original dialogue track with a new one that is understandable to target viewers and the latter presents the translation in the form of captions at the bottom of screen, leaving the original soundtrack untouched (Pérez-González, 2009; Gambier, 2013; O'Sullivan & Cornu 2019).

The introduction of sound films was, however, more complicated than it was assumed and thought, especially in Europe as it crippled the local film industry in countries whose budget was small and their film infrastructure was not sophisticated. Home productions in these countries were consequently suppressed, leading to an increased import of foreign programs (O'Connell, 2007; see also O'Sullivan & Cornu 2019). Small countries, such as Belgium, Portugal, Denmark and

Scandinavian countries chose subtitling to drive the expenses down. Such big countries as France, Italy, Germany, and Spain, however, selected dubbing because it became "an assertion of the supremacy of the national language" (Danan, 1991, p. 612) and it was their government goal to treasure their national language, and to shield it from "the onslaught of anglicisms" (Chiaro, 2013, p. 3) or because some countries, such as Iran, were suffering from a high degree of illiteracy (Naficy, 2011). This division of countries into dubbing, subtitling and voice-over, however, does not work anymore. This is because not only are these classifications obsolete but also they had been devised on the basis of clichés (Chaume, 2012, 2013). In other words, "dubbing/subtitling boundaries regarding consumer choice remain fuzzy and fraught with bias" (Chiaro, 2019, p. 171). Notwithstanding this, it is still reasonable to think that some countries still have one primary AVT modality, along with other (emerging) modalities.

Using the concept of "digitalization", which is "the actual process of change", Chaume (2019, 108) contends that AVT practices across the globe have been under the influence of "digital production, digital distribution, digital consumption and digital manipulation". This in turn has led to the belief that new habits of audience and AVT types have begun to rise to the surface. Digital technology, including downloading and streaming, as Chaume (2016, p. 69) puts it, has empowered audience to have more choices on how and what to watch and it has also boosted the diversity of translation types and practices. This digitalization has resulted in shaping audience viewing styles, notably in dubbing countries, as people now use more and more subtitled programs (Chaume, 2019, p. 108). In line with this reality, "the instantaneity and global reach of audiovisual content flows", as Pérez-González (2019, p. 2) notes, have certainly questioned the way we used to define and understand the audience. Evidently, national audiences are no longer homogenous and it is difficult to imagine a coherent AVT market around them, highlighting the fact that "the dominance of specific AVT modalities within individual countries" has started to collapse (Pérez-González, 2019, p. 2). In addition, the widespread availability of large unauthorized online archives of films and TV shows—again as a result of digitalization—has made it easier for audience to access more foreign content (Orrego-Carmona, 2018). Be that as it may, the AVT industry in many countries does not have the capability to cater for the needs of viewers who may wish to watch their favorite program shortly after its original release. This is where audience, especially those in dubbing regions, rely most heavily on amateur subtitling, which is both cheap and swift and does not suffer from content manipulation (He, 2017; Massidda & Casarini, 2017).

The above discussion clearly reveals that watching cinematic programs has experienced a significant shift with the emergence of digitalization and participatory practices like amateur subtitling. Notwithstanding this, such anecdotal evidence has remained unsubstantiated and should be complemented with empirical investigations on audience. Empirical studies, Díaz Cintas (2020, p. 221) points out, "towards evaluating and appraising the prescriptive conventions applied in the profession" are extremely rare. Accordingly, this article aims to bring to light how Iranians watch foreign cinematic programs nowadays and what constitutes their viewing preferences and habits, when it comes to AVT.

### 2.2. Empirical reception studies

Audience or receivers, according to Díaz Cintas and Szarkowska (2020), constitute an essential part of the communicative act and cannot be ignored in process-oriented research. Interest has emerged out of the curiosity of researchers for understanding how viewers perceive, receive and understand translated programs. Borrowing Chaume's (2018) methodological turns, reception studies could be positioned at both social studies and cognitive studies. Exploring viewers' viewing habits, for example, is part of the social turn in AVT where relevant data are accumulated through questionnaires, interviews, and focus groups, allowing for both or a mash-up of quantitative and qualitative approaches. The cognitive turn, in contrast, offers an analysis of what is happening in the mind of the spectator, "bringing feelings and body response to the surface" (Chaume, 2018, p. 54). When conducting a reception study, Gambier (2018) maintains, several variables could be taken into consideration:

- Sociological variables, which include demographic information of the viewers, such as age, gender, language proficiency and so forth
- Audiovisual variables, which embrace, among others, genre, broadcasting time and the relationship between images and dialogues.

Research has it, for example, age and language proficiency could impact the reception of subtitling and dubbing (for example see Orrego-Carmona, 2016; Perego, 2018). The audiovisual variable, such as the complexity of the program could also affect the way viewers apprehend and appreciate dubbed and subtitled programs (Perego, Missier, & Stragà, 2018).

To the best of our knowledge, there have been a few large-scale AVT studies on audience viewing habits and preferences. For instance, Kizeweter (2015) used internet forums as a source to gather data on AVT habits among the Polish. An important question was already posted in the blog which was "voice-over, dubbing or subtitles[?]" (171). The overall analysis showed that 14.8% of the users preferred dubbing, but the majority (62.96%) argued against dubbing, and 22.22% did not provide a flat answer for the question. Along the same line, Szarkowska and Laskowska (2015) reported the result of an online survey on the AVT preferences of Polish viewers (both with and without hearing impairments). In this large-scale study, 815 respondents took part, yet only 427 participants answered all questions. The results showed that the majority of the

participants (82.67%) preferred watching films with subtitles. Such a result was expected as nearly half of the respondents were deaf or hard of hearing. However, other results showed that 77.25% of hearing participants preferred subtitling, while only 6.88% favored the voice-over modality. The findings of a quantitative survey by Orrego-Carmona (2014) with 332 university students in Spain suggest that Spaniards mainly use the Internet to watch their favorite program and a good—but not large—number of the respondents reported watching subtitled programs. Dubbing was the favorite choice for more than half of the respondents.

In a Spanish context and with 72 viewers, Enríquez-Aranda and García Luque (2018) reported that most of the audience prefer to watch foreign films (than Spanish films) and with Spanish subtitles (than with Spanish dubbing) and they prefer to watch foreign programs in the cinemas. Also, English-language and French-language films were audience's preferences. Though not as viewing habits investigations, Matamala, Perego, and Bottiroli (2017), Matamala and Ortiz-Boix (2018) and Ameri and Ghodrati (2019) found that the participants of their studies are mainly subtitling users even though they had grown up with dubbed content.

This survey of the small body of literature that has so far examined audience habits and preferences brings to the fore the need for further research in new contexts. Taken all together, the principal underlying purpose for doing this study is to survey the current pattern in terms of Iranian audience habits and preferences when watching foreign films and TV shows. Also, the study investigates if there is a correlation between the language mastery and viewing preferences of the audience of audiovisual translation. Broadly, it offers insights into Iranian viewers' engagement with AVT and explores the situation of subtitling and dubbing in Iran.

## 3. Method

### 3.1. Research design

This study enjoys an Internet survey. A survey is a quantitative research instrument which collects large-scale data like attitudes, opinions, habits, etc. from an enormous number of audience and it aims to discover possible patterns and trends among a specific population, say viewers here (Cohen, Manion, & Morrison 2018; Mellinger & Hanson 2017). The present web-based survey was administered online and respondents were requested to fill in the questionnaire, that was designed and developed by an Iranian online survey tool. More specifically, "in web surveys the questionnaire is accessed and answered by respondents using a web browser" (Vehovar & Manfreda 2017, p. 146), as is the case with the present survey. Both open-ended and close-ended items collect data concerning individuals' demographic information, attitudes and opinions for this study. Additionally, the close-ended items were not based on a Likert scale and were measured at the categorical level.

### 3.2. Participants

A total of 1261 Iranians participated in the study. The final sample, however, comprised 1137 Iranians who completely filled in the questionnaire, including 357 females (31.4%) and 780 males (68.6%). Owing to the online data collection procedure, the participants were from across Iran. Convenience sampling—a non-probability sampling technique—was employed as the participants "included in the sample cannot be determined" in advance and "it is left up to each individual to choose to participate in the survey" (Fricker Jr, 2017, p. 166). Here, it suffices to mention that the survey link was posted on Telegram channels, where it was left up to users visiting posts of the channels to choose to click on the link and attend the survey.

The age range of the respondents was between 12 and 59, with the mean age of 25.19 ($SD$ = 7.96). As for the educational background, the participants were students or graduates of high school ($N$ = 333, 29.3%), B.A. ($N$ = 507, 44.6%), M.A. ($N$ = 176, 15.5%) and Ph.D. ($N$ = 70, 6.2%) programs, and others ($N$ = 51, 4.5%). With reference to the European Framework of Reference for Languages, the participants self-reported their level of English as follows; not familiar ($N$ = 42, 3.7%), basic ($N$ = 291, 25.6%), intermediate ($N$ = 419, 36.9%), upper-intermediate ($N$ = 266, 23.4%), and advanced ($N$ = 119, 10.5%), which more and less resembles the report of EPI (2020) where Iran holds a low level of proficiency in English. This should be added that self-reported language proficiency, Kaushanskaya, Blumenfeld, and Marian (2020) maintain, is quite in line with the objective assessment of individuals' competency.

### 3.3. Instrumentation

Due to the lack of any relevant questionnaire for gauging individuals' viewing habits and preferences in AVT, the questionnaire was built by the researcher. The questionnaire was designed and developed in consultation with translation studies scholars and a review of the relevant literature so that the items are comprehensive and representative, and act in accordance with the research purposes. The questionnaire had 14 items; four measured demographic information and the other ten items explored individuals' preferences and habits. In other words, the constructs of the questionnaire include:

- demographic information: age, gender, educational background and English language mastery
- film viewing habits and preferences: weekly hours of and preferences for watching domestic and foreign cinematic products, favorite genres, favorite types of cinematic programs, favorite international cinemas and favorite ways of accessing foreign productions
- AVT habits and preferences: dubbing and subtitling viewing habits

To ensure the validity of the questionnaire, it was sent to four Iranian translation studies researchers to comment on the accuracy and wording of the items. After minor modifications were made following the expert reviews, the questionnaire was piloted and refined using a sample of ten individuals who read the items one by one and stated their understanding of each item. This helped us mark problems with the understanding of the items. The scale was in Persian, which matches the respondents' mother tongue, and the reliability was not possible to be reported as the items were not based on a Likert scale and were measured at the categorical level.

### 3.4. Data collection and analysis

The present Internet-based survey was operated on an Iranian website that hosted the questionnaire (http://cafepardazesh.ir). Given that anyone could be a viewer of audiovisual programs, an "unrestricted sample" was favored, indicating that anyone can complete the questionnaire (Cohen, Manion, & Morrison, 2018). In addition, there is a bias with the Internet surveys, because, at times, those who are not tech-savvy or have no Internet connection or cannot click the links to unknown or unsafe websites do not attend the survey. Also, the sampling procedure is non-probability and mainly attracts volunteers (Cohen, Manion, and Morrison 2018). These defects may decrease the generalizability of the findings. To recruit the respondents, attempts were made to use some specialized channels on Telegram. Telegram is dominant in Iran and this helped us collect considerably more answers and include a wide spectrum of viewers from entire Iran, and as a result, the study is not confined to a select group of respondents. A couple of Telegram channels on films were chosen and the first author contacted admins of these channels and they were asked to share the online survey link on the channels. The collected data were imported into Statistical Package for Social Sciences Program (SPSS, version 22) and the relevant variables were defined.

## 4. Results

### 4.1. Data screening

Prior to the data analysis, the data were inspected for normality and missing values. This is because the data should be clean and reliable for any further statistical analysis. Given that the questionnaire used for this research lacks any Likert-based items and all the items are categorical, running a normality test sounds pointless. When the respondents deliberately or accidentally left some items unanswered, the researchers are faced with data called missing data or values. As to the missing data treatment, the present literature has come up with three techniques. The first of which is the "deletion method" where the entire data of any case (i.e. the respondent) with one or more missing data are entirely removed from the analysis or the researcher can only remove only those cases with a particular missing item and variable. The former technique is called "exclude cases listwise" and the other is named "exclude cases pairwise" (Cohen et al., 2018, pp. 752-753). The removal of missing data should be done cautiously because this can result in losing numerous cases. The other alternative method for dealing with missing data is "imputation method" wherein the missing values and variables are replaced with educated guesses and it is for situations where the missing data pattern is pretty systematic (Cohen et al., 2018, 752-753). Because a considerable amount of data was collected and the scale had only 14 questions—which means that all items are more and less equally important for the data analysis—, the first technique "exclude cases listwise" was chosen. This exclusion of missing data resulted in losing roughly ten percent of the data (124 cases) and having 1137 cases who filled in the questionnaire completely.

### 4.2. Film viewing habits

The first research question asked how much the respondents watch domestic and foreign fictional programs. The descriptive data, including percentages and frequencies can be found in Figures 1 and 2. As the figures illustrate, while more than half of Iranians watch domestic programs less than one hour a week, roughly 20 percent of them watch more than nine hours of foreign programs per week as the likelihood is much higher.

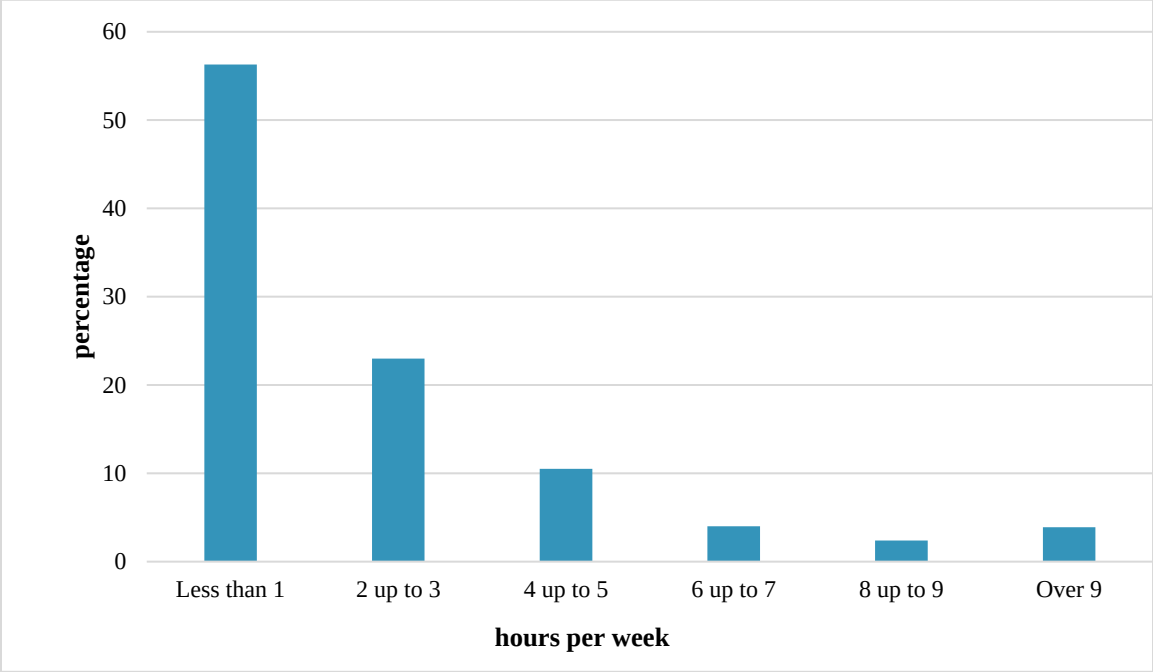



**Fig. 1** Film Viewing Habits: Iranian Productions

What can be implied from these numbers and figures is that Iranians are more inclined toward foreign programs. In other words, foreign programs account for a fair amount of their viewing time. This clearly highlights the central position of foreign programs among Iranians.

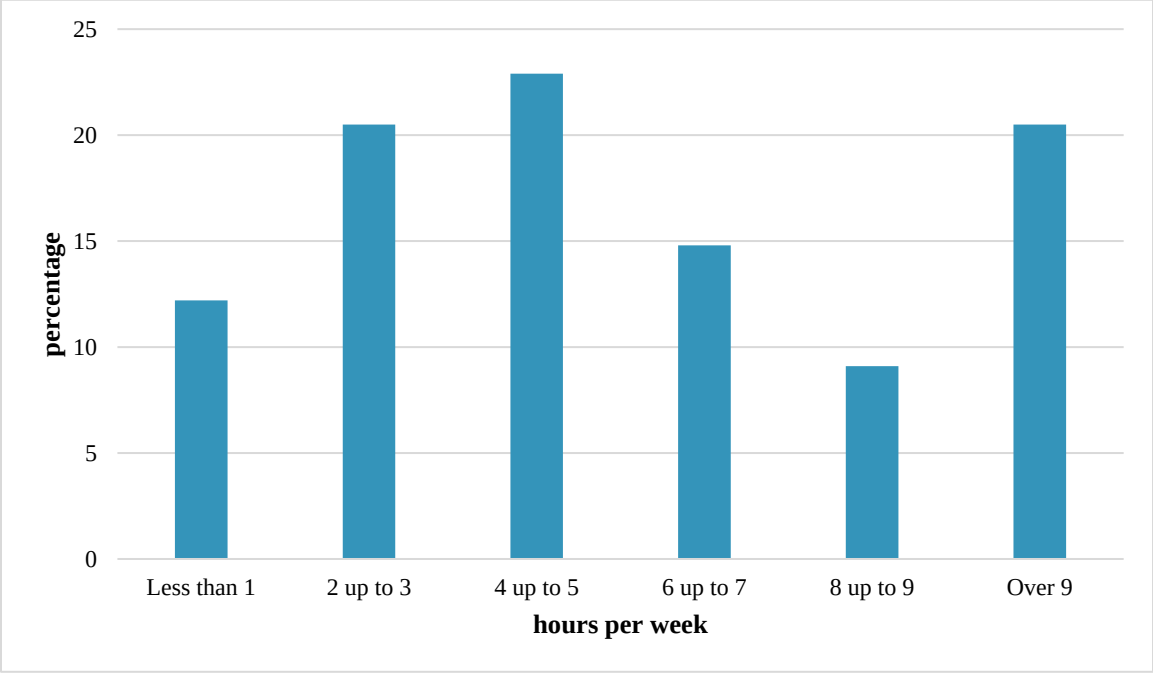



**Fig. 2** Film Viewing Habits: Foreign Productions

### 4.3. Audiovisual translation viewing preferences and habits

The second research question asked how Iranians prefer to watch foreign fictional programs. As Figure 3 illustrates, more than half of the participants choose Persian subtitling, and only 20 percent of them would like to watch programs dubbed. Besides,

16.9 percent of them tend to watch their programs with the original language subtitles, and a handful of them (4.9%) watch foreign programs without the aid of dubbing and subtitling.

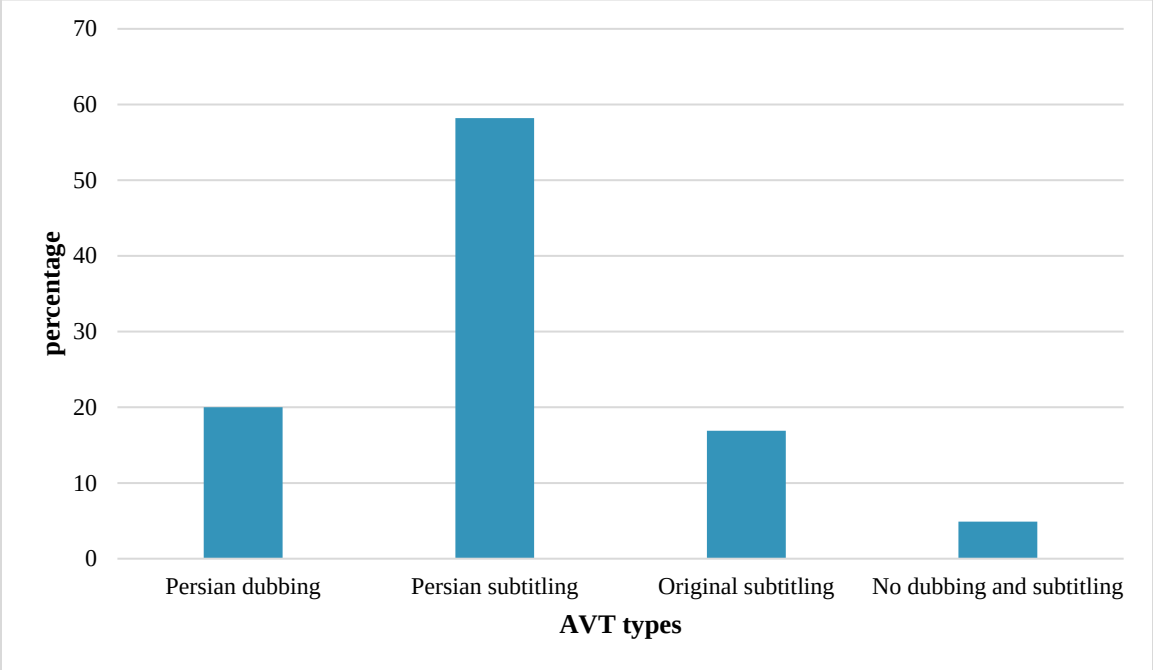



**Fig. 3** AVT Viewing Habits

To deeply understand the viewers' AVT viewing habits and offer more detailed answers to the second research question, the respondents were asked to report to what extent they watch dubbed and subtitled programs, and its results are offered in Figures 4 and 5. What is learned from Figure 4 is that the personal judgment of 66.8 percent of the viewers consisted mainly of *always* and *usually* for subtitling. An opposite trend can be seen about dubbing since the reports of 27.9 percent of the viewers consisted mainly of *always* and *usually*, together with 47.2 percentage for *barely* and *never*.

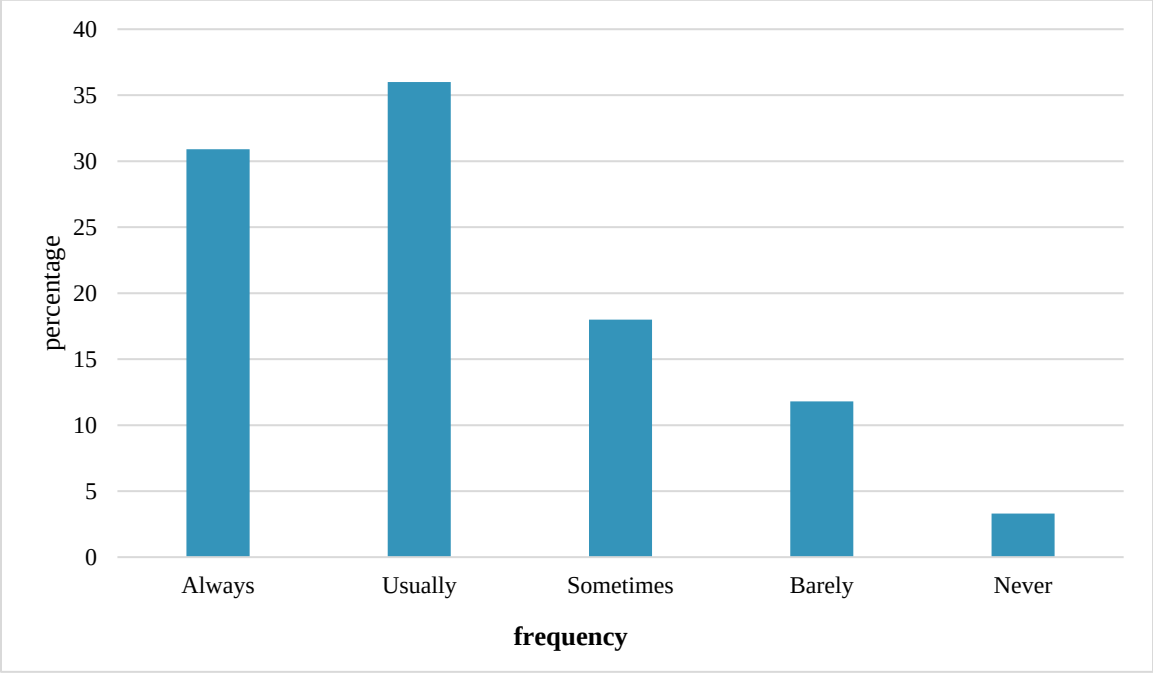



**Fig. 4** Subtitling Viewing Habits

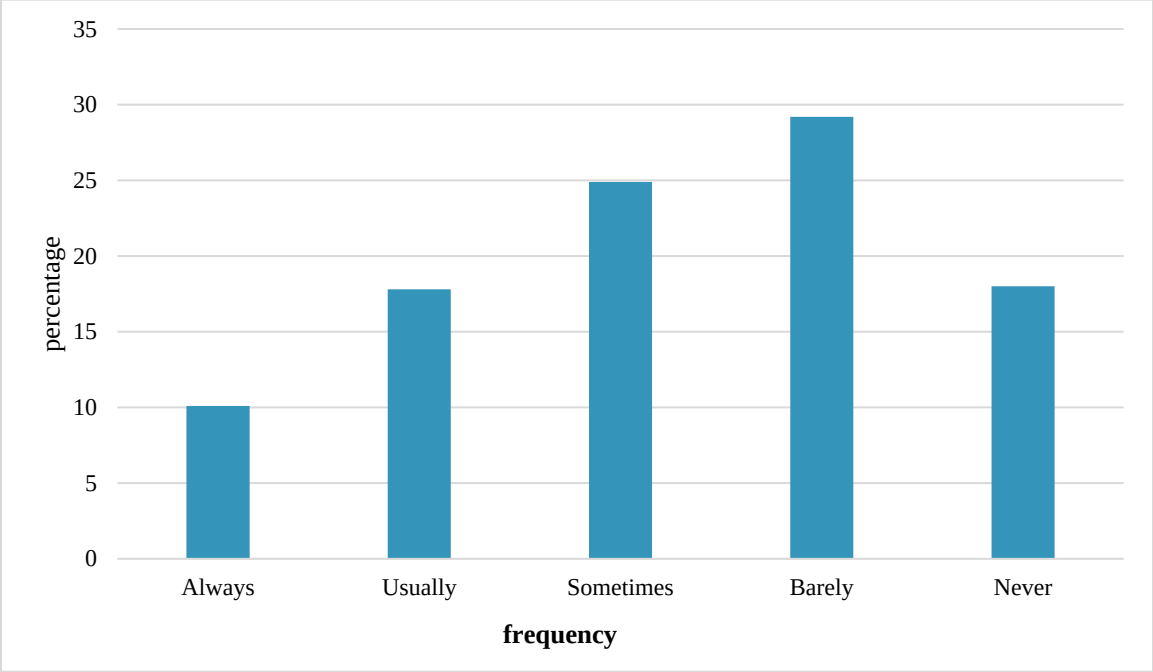



**Fig. 5** Dubbing Viewing Habits

### 4.4. Correlates of audiovisual translation preferences and English language proficiency

In order to answer the third research question concerning the relationships between individuals' AVT viewing habits and their English competency, a chi-squared test was used. A significant association between the two variables emerged from the analysis ($\chi(1) = 240.75$, $p = 0.00$). Nonetheless, the strength of the relationship between the two variables, according to the Cramer's *V* test, is weak. This statistical analysis, in sum, suggests that the two variables are not independent of each other. Given this statistically significant chi-squared test, a post-hock test was employed to discover the significance of each cell. Any statistics beyond 1.96 for the adjusted residuals are significant. However, some types of corrections should be applied here, which is a Bonferroni correction. The p-value of 0.05 should be divided into 20, which results in a Bonferroni corrected p-value of 0.00024 (see Beasley & Schumacker, 1995). The result of computing variables to find the significance of the adjusted values is reported in Appendix A. To sum up, several cells were significantly different from the expectation of the null hypothesis. As far as the interpretation of the findings is concerned (see Appendix B), those with an advanced level of English are more inclined to watch fictional programs with English subtitles or without the help of dubbing and subtitling. The opposite trend is observable in those who are not familiar with English as they prefer to watch dubbed programs. The other telling results indicate that with the increase of one's mastery over English, the tendency towards watching the original program with Persian subtitles increases excepting the advanced group who mainly goes for English subtitling (i.e., intralingual subtitles).

### 4.5. Different ways to access foreign productions

The fourth research question addressed how the Iranian audience access foreign audiovisual programs and the results are shown in Figure 6. The interesting point is that most people download them from the Internet and the number of individuals who watch TV channels is roughly ten percent which is less than those who watch satellite channels. Other ways that were mentioned by the respondents embrace personal archives, borrowing programs from relatives and friends, using the phone application Rubika, and Telegram channels[1].

[1] The reason for excluding cinemas is that Iranian cinemas only screen Iranian productions.

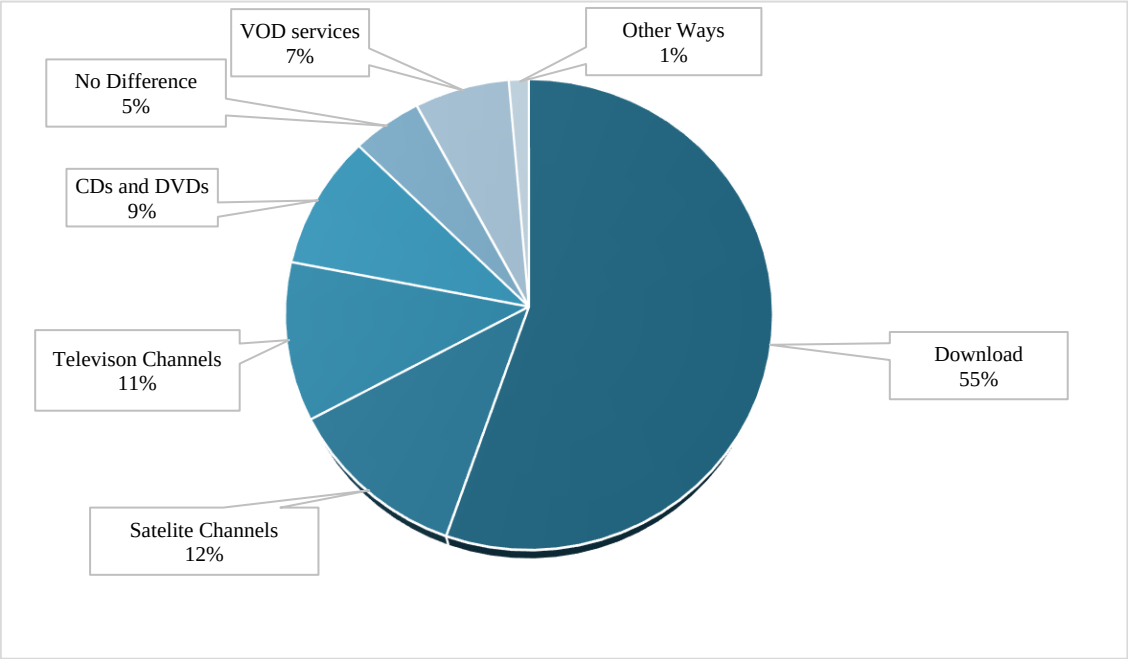



**Fig. 6** Different Ways to Access Foreign Productions

### 4.6. Preferences concerning a film's country of origin

The last research question investigated participants' preferences with regard to a film's country of origin. As could be expected, and presented in Figure 7, more than 60 percent of the respondents stated Hollywood cinema among their favorite foreign cinema. To this, we should add Korean, Turkish and Bollywood productions (9%, 8%, and 7%, respectively)[2]. Other international cinemas mentioned by the respondents include China, Japan, France, England, Russia, Germany, Italy, Australia, and Poland.

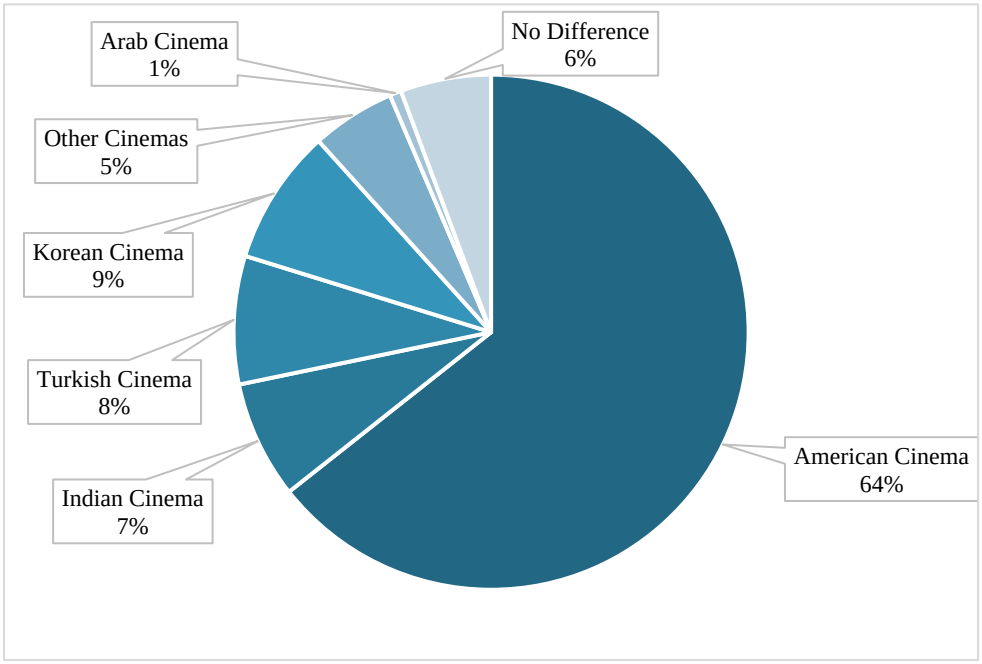



**Fig. 7** Preferences on Film's Country of Origin

[2] Numbers may not total 100 percent as participants were permitted to choose multiple alternatives.

## 5. Discussion and conclusion

This article has explored Iranian viewers' viewing habits and preferences in terms of AVT. The initial results showed that the respondents spend more time watching foreign programs, accounting for a large percentage when compared with domestic productions. A possible explanation for this finding is that the present time has enjoyed an unprecedented boom in the production of English audiovisual programs. Besides, the number and volume of foreign films and TV shows are much larger than those of Iranian programs; therefore, people have a wider range of choices. Foreign cinematic programs have therefore found their way into Iranian society. From a cultural perspective, Zeydabadi-Nejad (2016, p. 110) is of the view that foreign film viewing has become a big part of Iranian culture and is "part of the culture of resistance to regime impositions". This finding tallies with that of Enríquez-Aranda and García Luque (2018) in Spain, where people are using more audiovisual products with a language different from Spanish.

Despite being called a dubbing country on paper and to some extent in practice, Iran has a smaller number of dubbed productions, compared to amateur subtitling productions and given the omnipresence of multimedia products in our world, non-professional subtitling of cinematic products has become a daily practice in Iran (Ameri & Khoshsaligheh, 2019; Khoshsaligheh et al. 2020). The results revealed that dubbing is now way past its prime, which leaves just a small handful of individuals, at least not the younger groups, who still use dubbing. Just 20 percent in total choose dubbing as their preferred modality, while a considerable percentage of the respondents were subtitling users (nearly 60 percent). Needless to say, this discussion does not suggest that such changes and shifts have happened overnight because subtitling, albeit in its amateur form, has existed in Iran since the late 1990s. Therefore, continual exposure of Iranians to subtitling over the past two decades has altered their viewing habits, and watching subtitled programs is now part of people's everyday life. To quote Chaume (2020, p. 321), "habits and tastes are not easily changed in a short space of time".

However, taking into account the issue of habits, Mereu Keating (2019) claims that people may find it hard to change their old habits, and viewers in dubbing countries are scarcely accustomed to reading subtitles, and consequently they "would stay away from screenings of foreign films whose dialogue exchanges were conveyed through fast-disappearing" (2019, p. 72). The role of digitization should not, however, be overlooked here, which appears to play an important role in "moulding the habits of local audiences" (Chaume, 2019, p. 114). VOD platforms, along with the Internet and new technologies, being constantly on the rise in Iran, have allowed audience "to watch what they want, when they want and in the quantities that they want" (Díaz Cintas & Nikolić 2018, p. 3); therefore, they are not bound to a dictated translation modality. In the words of Chaume (2016), "the days of decisions taken by just a few agents, used to dictating what audiences like and dislike, are progressively coming to an end" (p. 72).

The noticeable tendency for watching foreign cinematic programs with subtitling in Iran has already been reported in Ameri and Ghodrati (2019) with a select sample of university students who used subtitled programs as a viable tool to hone their English skills; an issue which has already been well-studied in both applied linguistics and translation studies (e.g. Incalcaterra McLoughlin 2019, Rupérez Micola et al. 2019). Likewise, the study also showed how language competency of viewers could be an essential factor in choosing dubbing or subtitling. As shown, it is more than likely that people with a higher mastery over English would rather watch subtitled programs than dubbed ones. A new piece of research has it that people of subtitling countries have a higher level of English competency as far as the TOEFL exam scores are concerned; thus, a far-reaching effect of exposure to subtitled multimedia content is noticeable (Rupérez Micola et al., 2019).

Besides, the increasing popularity of and preference for subtitling, particularly in the younger groups in other dubbing countries (Matamala & Ortiz-Boix 2018; Matamala, Perego & Bottiroli, 2017, Perego et al., 2016, Orrego-Carmona, 2014, Enríquez-Aranda & García Luque, 2018) or voice-over countries (Szarkowska & Laskowska, 2015) are not just a coincidence. There is a reason why such shifts began to shape; subtitling is associated mainly with "general openness and curiosity for different languages and cultures" (Matamala, Perego, & Bottiroli, 2017, p. 431). Additionally, the unpopularity of dubbing among the respondents may result from the shortcomings and defects observed in the Persian dubbing industry. Ameri and Khoshsaligheh (2020), for example, report that the Iranian dubbing industry, according to the Iranian viewers, is no longer capable of offering high-quality products due to poor translations, lousy voice-acting, wrong policies, such as censorship, and unexpected setbacks in dubbing blockbusters. On a relevant note and in the case of Italy, Casarini (2012) maintains that despite the fact that this country has enjoyed a strong and stable tradition in terms of dubbing, many audiences, who turned out to be tech-savvy, are increasingly voicing their discontent "with the lower quality" and "rushed dubbing" of TV shows. Therefore, it is common to witness that they are willing to "to resort to fansubs – subtitles voluntarily created by language-skilled fans and available for free shortly after each US broadcast" (Casarini, 2012, p. 64).

The other impressive result of this research lends support to the idea that film piracy is very prevalent in Iran. First, only a handful of people watch national or satellite channels and less than this number tends to access video entertainment through VOD services, while roughly half of the viewers download their programs from the Internet, which hosts a growing volume of film piracy websites. Therefore, film piracy, as suggested by Orrego-Carmona (2018), has offered unlimited access to new releases and it costs viewers almost nothing. It is noteworthy that according to Zeydabadi-Nejad (2016), the case of Iran

is far from piracy because Iranians have not been permitted by the authorities to have access to international cinematic productions and more importantly, the copyrights of original works produced outside the Iran borders are barely protected by the Iranian government; therefore, by downloading foreign films and TV shows, "Iranians are not breaking the law so far as copyright is concerned" (p. 100). Overall, people are not watching as much TV channels anymore, and the public service broadcasting is now being replaced by new rivals.

Although American productions still constitute a large part of Iranian audience preference—which makes sense when considering the fact that major film studios feeding international markets are located in the United States—, non-English language films have garnered some popularity in Iran, thanks to national TV and satellite channels which have given rise to Turkish and Korean TV shows in the last decade. This result acts as evidence supporting the idea of "the destabilisation of English as the language par excellence in media entertainment and the renaissance of other original languages" (Bogucki & Díaz Cintas, 2020, p. 30).

As indicated by Díaz-Cintas, reception studies have the potential of establishing the missing link between the research and industry; "a cooperation that in turn holds promise for the development and provision of better products for end users" (Bogucki & Díaz Cintas, 2020, p. 28). Therefore, in light of the research findings, the policymakers in Iran are advised to take into consideration the changes that have happened in individuals' AVT preferences. The findings are indicative of changes but the Iranian national TV channels still resist the introduction of subtitling as an AVT alternative. The time has come for new changes to be made in the Iranian AVT system so as to meet the new needs and habits of Iranians. Therefore, the audience should be granted more extensive choices of AVT in Iran. It should not be taken to mean that nothing has been made to cater to audience needs in terms of subtitling, as Iranian VOD platforms have been forerunners in introducing subtitling to the Iranian society since 2017 or thereabouts (Khoshsaligheh et al., 2020). Notwithstanding these preliminary efforts, considerable steps should still be taken on TV channels. These measures could prevent individuals from the illegal downloading of cinematic products, and professional subtitling on national TV channels can be a possible way to fight piracy and cater to the audience needs. The results could be carried over into film translation classrooms to raise the awareness of translation students regarding the audience current viewing habits and styles and how audience and their reception may shape translation decisions and choices.

The paper has some limitations, which could be promising avenues for future studies. To avoid self-report biases and elicited data and to collect rich data through online observation, netnography (Kozinets, 2019) can be used. Netnography is a reliable method to reach qualitative understanding of, say, film community, blogs or any relevant forums and social media where many people voice and express their ideas and opinions about translation. Journal entries and diaries, to the best of our knowledge, have barely been used in the context of reception studies and AVT preferences. Indeed, research can benefit from these research tools by inviting people to record their daily exposure to subtitled and dubbed content to gain invaluable insights into individuals' AVT consumption. Additionally, the data were gathered through an Internet survey, introducing some biases as the majority of the respondents turned to be young. It is therefore recommended that future studies make use of face-to-face surveys to incorporate more respondents from the elderly group. Finally, at the time of collecting data for this study, the Iranian movie streaming services were few and they were at their early stage of development, but now are mushrooming, especially with the shutdown of many Iranian movie piracy websites in October 2019. In addition to their subtitling activities, the VOD platforms have been very active in dubbing new releases. A final recommendation is therefore the replication of this study to assess the impact of these services on dubbing reception.

## 6. Disclosure statement

The authors have no conflicts of interests to declare.

## 7. References

Ameri, S., & Ghodrati, M. (2019). جایگاه محصولات سینمایی خارجی در یادگیری زبان انگلیسی خارج از کلاس درس [Watching foreign audiovisual programs and out-of-class language learning]. *Journal of Language and Translation Studies, 52*(2), 127-156.

Ameri, S., & Khoshsaligheh, M. (2018). Exploring the attitudes and expectations of Iranian audiences in terms of professional dubbing into Persian. *Hermes, 57*, 175-193. doi:10.7146/hjlcb.v0i57.106206

Ameri, S., & Khoshsaligheh, M. (2019). Iranian amateur subtitling apparatus: A qualitative investigation. *Mutatis Mutandis: Revista Latinoamericana de Traducción*, *12*(2), 431-451. doi:10.27533/udea.mut.v12n2a08

Ameri, S., & Khoshsaligheh, M. (2020). Dubbing viewers in cyberspaces: A netnographic investigation of the attitudes of a Persian-language online community. *KOME, 8*(1), 23-43. doi:10.17646/KOME.75672.45

Antonini, R., & Chiaro, D. (2009). The perception of dubbing by Italian audiences. In J. Díaz-Cintas & G. Anderman (Eds.), *Audiovisual translation: Language transfer on screen* (pp. 97-104). Basingstoke, England: Palgrave MacMillan.

Beasley, T. M., & Schumacker, R. E. (1995). Multiple regression approach to analyzing cntingency tables: Post hoc and planned comparison procedures. *The Journal of Experimental Education, 64*(1), 79-93. doi:10.1080/00220973.1995.9943797

Bogucki, Ł., & Díaz Cintas, J. (2020). An excursus on audiovisual translation. In Ł. Bogucki & M. Deckert (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility* (pp. 11-32). Cham, Switzerland: Palgrave Macmillan.

Casarini, A. (2012). XOXO: Gossip Girl and dubbing in the age of 'Net lingo. In S. Bruti & E. Di Giovanni (Eds.), *Audiovisual translation across Europe: An ever-changing landscape* (pp. 59-78). Oxford, England: Peter Lang.

Chaume, F. (2012). *Audiovisual translation: Dubbing*. Manchester, England: St. Jerome.

Chaume, F. (2013). The turn of audiovisual translation: New audiences and new technologies. *Translation Spaces, 2*(1), 105-123. doi:10.1075/ts.2.06cha

Chaume, F. (2016). Audiovisual translation trends: Growing diversity, Choice, and enhanced localization. In A. Esser, M. Á. Bernal-Merino, & I. R. Smith (Eds.), *Media across borders: Localizing tv, film, and video games* (pp. 68-84). London, England: Routledge.

Chaume, F. (2018). An overview of audiovisual translation: Four methodological turns in a mature discipline. *Journal of Audiovisual Translation, 1*(1), 40-63.

Chaume, F. (2019). Audiovisual translation in the age of digital transformation: Industrial and social implications. In I. Ranzato & S. Zanotti (Eds.), *Reassessing dubbing: Historical approaches and current trends* (pp. 104–124). Amsterdam, Netherlands: John Benjamins.

Chaume, F. (2020). Localizing media contents. Technological shifts, global and social differences and activism in audiovisual translation. In S. Shimpach (Ed.), *The Routledge companion to global television* (pp. 320-331). New York, NY Routledge.

Chiaro, D. (2013). Audiovisual translation. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Oxford, England: Wiley-Blackwell.

Chiaro, D. (2019). To Europe with love: Woody Allen's liquid society. In I. Ranzato & S. Zanotti (Eds.), *Reassessing dubbing: Historical approaches and current trends* (pp. 169-190). Amsterdam, Netherlands: John Benjamins.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). London, England: Routledge.

Danan, M. (1991). Dubbing as an expression of nationalism. *Meta, 36*(4), 606-614. doi:10.7202/002446ar

Díaz Cintas, J. (2020). Audiovisual translation In E. Angelone, M. Ehrensberger-Dow, & G. Massey (Eds.), *The Bloomsbury companion to language industry studies* (pp. 209-230). London, England: Bloomsbury.

Díaz Cintas, J., & Nikolić, K. (2018). Introduction: Fast-Forwarding with Audiovisual Translation. In K. Nikolić & J. Díaz Cintas (Eds.), *Fast-Forwarding with Audiovisual Translation* (pp. 1-13). Bristol, England: Multilingual Matters.

Díaz Cintas, J., & Szarkowska, A. (2020). Experimental Research in Audiovisual Translation – Cognition, Reception, Production. *JoSTrans, 33*, 3-16.

Dwyer, T. (2017). *Speaking in Subtitles: Revaluing Screen Translation*. Edinburgh: Edinburgh University Press.

Dwyer, T. (2019). Audiovisual translation and fandom. In L. Pérez-González (Ed.), *The Routledge Handbook of Audiovisual Translation* (pp. 436-452). London, England: Routledge.

Enríquez-Aranda, M., & García Luque, F. (2018). The reception of subtitled films from a sociological perspective: An empirical case study. *64*(3), 464-489. doi:10.1075/babel.00041.enr

EPI, E. (2020). *The world's largest ranking of countries and regions by English skills: The case of Iran*. Retrieved from https://www.ef.com/wwen/epi/regions/middle-east/iran:

Fricker Jr, R. D. (2017). Sampling Methods for Online Surveys. In N. G. Fielding, R. M. Lee, & G. Blank (Eds.), *The SAGE Handbook of Online Research Methods* (pp. 162-183). London, England: SAGE.

Gambier, Y. (2013). The Position of Audiovisual Translation Studies. In C. Millán & F. Bartrina (Eds.), *The Routledge handbook of translation studies* (pp. 45-59). Abingdon, England: Routledge.

Gambier, Y. (2018). Translation studies, audiovisual translation and reception. In E. Di Giovanni & Y. Gambier (Eds.), *Reception studies and audiovisual translation* (pp. 43-66). Amsterdam, The Netherlands: John Benjamins.

He, Z. (2017). Chinese fansubbers' pursuit of translation accuracy in the absence of state censorship. In D. Orrego-Carmona & Y. Lee (Eds.), *Non-professional subtitling*. Newcastle, London: Cambridge Scholars Publishing.

Incalcaterra McLoughlin, L. (2019). Audiovisual translation in language teaching and learning. In L. Pérez-González (Ed.), *The routledge handbook of audiovisual translation* (pp. 483-497). London, England: Routledge.

Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2020). The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition, 23*(5), 945-950. doi:10.1017/S1366728919000038

Khoshsaligheh, M., Ameri, S., Shokoohmand, F., & Mehdizadkhani, M. (2020). Subtitling in the Iranian Mediascape: Towards a Culture-Specific Typology. *International Journal of Society, Culture & Language, 8*(2), 55-74.

Kizeweter, M. (2015). Voices about Polish voices in foreign films: Using an Internet forum as a source of information about the opinions of Polish viewers on dubbing as a mode of AVT. In Ł. Bogucki & M. Deckert (Eds.), *Accessing audiovisual translation* (pp. 163-178). Frankfurt, Germany: Peter Lang.

Kozinets, R. V. (2019). *Netnography: The Essential Guide to Qualitative Social Media Research* (3rd ed.). Thousand Oaks, CA: Sage.

Massidda, S., & Casarini, A. (2017). Sub me do: The development of fansubbing in traditional dubbing countries – the case of Italy. In D. Orrego-Carmona & Y. Lee (Eds.), *Non-Professional Subtitling* (pp. 63-86). Cambridge: Cambridge Scholars Publishing.

Matamala, A., & Ortiz-Boix, C. (2018). Text-to-speech voice-over? A study on user preferences in the voicing of wildlife documentaries. *SKASE Journal of Translation and Interpretation, 11*(1), 24-38.

Matamala, A., Perego, E., & Bottiroli, S. (2017). Dubbing versus subtitling yet again? An empirical study on user comprehension and preferences in Spain. *Babel, 63*(3), 423–441. doi:10.1075/babel.63.3.07mat

MECD. (2012). *European Survey on Language Competences (ESLC): Expert Analysis.* Retrieved from http://www.mecd.gob.es/dctm/inee/internacional/eeclvol2ingleslinea.pdf?documentId=0901e72b81443480:

Mellinger, C. D., & Hanson, T. A. (2017). *Quantitative research methods in translation and interpreting studies*. Abingdon, England: Routledge.

Mereu Keating, C. (2019). Locked into dubbing. In I. Ranzato & S. Zanotti (Eds.), *Reassessing dubbing: Historical approaches and current trends* (pp. 64-78). Amsterdam, Netherlands: John Benjamins.

Naficy, H. (2011). *A social history of Iranian cinema: The artisanal era, 1897–1941* (Vol. 1). London, England: Duke University Press.

O'Connell, E. (2007). Screen translation. In P. Kuhiwczak & K. Littau (Eds.), *A companion to translation studies* (pp. 120-133). Clevedon, England: Multilingual Matters.

O'Sullivan, C., & Cornu, J.-F. (2019). History of audiovisual translation. In L. Pérez-González (Ed.), *The routledge handbook of audiovisual translation* (pp. 15-30). Abingdon: Routledge.

Orrego-Carmona, D. (2014). Subtitling, video consumption and viewers: The impact of the young audience. *Translation Spaces, 3*, 51-70. doi:10.1075/ts.3.03orr

Orrego-Carmona, D. (2016). A reception study on non-professional subtitling: Do audiences notice any difference? *Across Languages and Cultures, 17*(2), 163–181. doi:10.1556/084.2016.17.2.2

Orrego-Carmona, D. (2018). New audiences, international distribution, and translation. In E. D. Giovanni & Y. Gambier (Eds.), *Reception studies and audiovisual translation* (pp. 321-342). Amsterdam, Netherlands: John Benjamins.

Orrego-Carmona, D., & Lee, Y. (Eds.). (2017). *Non-professional subtitling*. Newcastle, England: Cambridge Scholars Publishing.

Pedersen, J. (2018). From old tricks to Netflix: How local are interlingual subtitling norms for streamed television? . *Journal of Audiovisual Translation, 1*(1), 81-100.

Perego, E. (2018). Watching translated audiovisuals: Does age really matter? In J. Díaz Cintas & K. Nikolić (Eds.), *Fast-forwarding with audiovisual translation* (pp. 80-92). Bristol, London: Multilingual Matters.

Perego, E., Laskowska, M., Matamala, A., Remael, A., Robert, I. S., Szarkowska, A., . . . Bottiroli, S. (2016). Is subtitling equally effective everywhere? A first cross-national study on the reception of interlingually subtitled messages. *Across Languages and Cultures, 16*(2), 205-229. doi:10.1556/084.2016.17.2.4

Perego, E., Missier, F. D., & Stragà, M. (2018). Dubbing vs. subtitling: Complexity matters. *Target, 30*(1), 137-157. doi:10.1075/target.16083.per

Pérez-González, L. (2009). Audiovisual translation. In M. Baker & G. Saldanha (Eds.), *The routledge encyclopedia of translation studies* (2nd ed., pp. 13-20). Abingdon, England: Routledge.

Pérez-González, L. (2014). *Audiovisual translation: Theories, methods and issues*. London, England: Routledge.

Pérez-González, L. (2019). Rewiring the circuitry of audiovisual translation: Introduction. In L. Pérez-González (Ed.), *The Routledge handbook of audiovisual translation* (pp. 1-12). London, England: Routledge.

Rupérez Micola, A., Aparicio Fenoll, A., Banal-Estañol, A., & Bris, A. (2019). TV or not TV? The impact of subtitling on English skills. *Journal of Economic Behavior & Organization, 158*, 487-499. doi:https://doi.org/10.1016/j.jebo.2018.12.019

Szarkowska, A., & Laskowska, M. (2015). Poland – a voice-over country no more? A report on an online survey on subtitling preferences among Polish hearing and hearing-impaired viewers. In Ł. Bogucki & M. Deckert (Eds.), *Accessing audiovisual translation* (pp. 179-197). Frankfurt, Germany: Peter Lang.

Vehovar, V., & Manfreda, K. L. (2017). Overview: Online Surveys. In N. G. Fielding, R. M. Lee, & G. Blank (Eds.), *The Sage Handbook of Online Research Methods* (pp. 143-161). London, England: SAGE.

Zeydabadi-Nejad, S. (2016). Watching the forbidden: Reception of banned films in Iran. In M. Hagener, V. Hediger, & A. Strohmaier (Eds.), *The state of post-cinema: Tracing the moving image in the age of digital dissemination* (pp. 99-113). London, England: Palgrave Macmillan.

## 8. Appendix

**Appendix A. Bonferroni corrected p-value**

**Bonferroni corrected p-value**

| Adjusted squares | chi-squares | Bonferroni corrected p-value |
|---|---|---|
| 3.39 | 11.49 | .00070* |
| -1.42 | 2.02 | .15561 |
| -.78 | .61 | .43539 |
| -1.30 | 1.69 | .19360 |
| 6.27 | 39.31 | .00000 |
| -1.85 | 3.42 | .06431 |
| -3.56 | 12.67 | .00037 |
| -2.20 | 4.84 | .02781 |
| 1.13 | 1.28 | .25848 |
| 3.12 | 9.73 | .00181 |
| -3.59 | 12.89 | .00033* |
| -3.24 | 10.50 | .00120 |
| -5.80 | 33.64 | .00000* |
| 3.00 | 9.00 | .00270 |
| .61 | .37 | .54186 |
| 1.89 | 3.57 | .05876 |
| -4.79 | 22.94 | .00000* |

| | | |
|---|---|---|
| -5.56 | 30.91 | .00000* |
| 10.36 | 107.33 | .00000* |
| 6.44 | 41.47 | .00000* |

## Appendix B. Crosstabulation of English Mastery and AVT Viewing Habits

### Crosstabulation of English Mastery and AVT Viewing Habits

| | | | AVT Viewing Habits | | | | Total |
|---|---|---|---|---|---|---|---|
| | | | Dubbing | Persian SUB | Original Language | English SUB | |
| English Mastery | Not Familiar | Count | 17 | 20 | 1 | 4 | 42 |
| | | Expected Count | 8.4 | 24.5 | 2.1 | 7.1 | 42.0 |
| | | **% within English Mastery** | **40.5%** | **47.6%** | **2.4%** | **9.5%** | **100.0%** |
| | | % within AVT Viewing Habits | 7.5% | 3.0% | 1.8% | 2.1% | 3.7% |
| | | % of Total | 1.5% | 1.8% | 0.1% | 0.4% | 3.7% |
| | | Adjusted Residual | **3.4** | **-1.4** | **-.8** | **-1.3** | |
| | Basic | Count | 95 | 156 | 3 | 37 | 291 |
| | | Expected Count | 58.1 | 169.4 | 14.3 | 49.1 | 291.0 |
| | | **% within English Mastery** | **32.6%** | **53.6%** | **1.0%** | **12.7%** | **100.0%** |
| | | % within AVT Viewing Habits | 41.9% | 23.6% | 5.4% | 19.3% | 25.6% |
| | | % of Total | 8.4% | 13.7% | 0.3% | 3.3% | 25.6% |
| | | Adjusted Residual | **6.3** | **-1.9** | **-3.6** | **-2.2** | |
| | Average | Count | 91 | 269 | 8 | 51 | 419 |
| | | Expected Count | 83.7 | 244.0 | 20.6 | 70.8 | 419.0 |
| | | **% within English Mastery** | **21.7%** | **64.2%** | **1.9%** | **12.2%** | **100.0%** |
| | | % within AVT Viewing Habits | 40.1% | 40.6% | 14.3% | 26.6% | 36.9% |
| | | % of Total | 8.0% | 23.7% | 0.7% | 4.5% | 36.9% |
| | | Adjusted Residual | **1.1** | **3.1** | **-3.6** | **-3.2** | |
| | | Count | 20 | 176 | 15 | 55 | 266 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Expected Count | 53.1 | 154.9 | 13.1 | 44.9 | 266.0 |
| | Above Average | **% within English Mastery** | **7.5%** | **66.2%** | **5.6%** | **20.7%** | **100.0%** |
| | | % within AVT Viewing Habits | 8.8% | 26.6% | 26.8% | 28.6% | 23.4% |
| | | % of Total | 1.8% | 15.5% | 1.3% | 4.8% | 23.4% |
| | | Adjusted Residual | **-5.8** | **3.0** | **.6** | **1.9** | |
| | | Count | 4 | 41 | 29 | 45 | 119 |
| | | Expected Count | 23.8 | 69.3 | 5.9 | 20.1 | 119.0 |
| | Advanced | **% within English Mastery** | **3.4%** | **34.5%** | **24.4%** | **37.8%** | **100.0%** |
| | | % within AVT Viewing Habits | 1.8% | 6.2% | 51.8% | 23.4% | 10.5% |
| | | % of Total | 0.4% | 3.6% | 2.6% | 4.0% | 10.5% |
| | | Adjusted Residual | **-4.8** | **-5.6** | **10.4** | **6.4** | |
| Total | | Count | 227 | 662 | 56 | 192 | 1137 |
| | | Expected Count | 227.0 | 662.0 | 56.0 | 192.0 | 1137.0 |
| | | % within English Mastery | 20.0% | 58.2% | 4.9% | 16.9% | 100.0% |
| | | % within AVT Viewing Habits | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |
| | | % of Total | 20.0% | 58.2% | 4.9% | 16.9% | 100.0% |

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 67-73*

Applied Linguistics Inquiry

https://doi.org/10.22077/ALI.2022.5574.1004

# English Teachers' Motivations for Research Engagement

Mehdi Mehranirad[1]

[1]Associate Professor of Applied Linguistics, English Department, Faculty of Letters and Humanities, University of Neyshabur, Iran

**ARTICLE HISTORY**

Received: 28 July 2022
Revised: 28 October 2022
Accepted: 27 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: Mehrani@neyshabur.ac.ir

**ABSTRACT**

Within the last few decades, the conception of research engagement has been widely considered vital in teachers' professional development. The literature is replete with numerous arguments both about the benefits of doing and reading research and the reasons for doing so. Within these discussions, however, the opinions of teachers are mostly ignored or reflected only circumstantially. The present study was conducted to investigate the reasons for which Iranian English teachers engage in research. First, a provisional survey questionnaire was designed, using experts' opinion and a comprehensive review of the related literature. The instrument was then validated through conducting exploratory factor analysis on teachers' responses to the survey instrument. Analysis of the results showed that teachers' reasons and motivations for research engagement can best be categorized in four groups: benefits for professional development, instrumental and personal motivations, organizational expectations, and pedagogical benefits. Subsequent analyses also revealed that research engagement among Iranian English teachers is mainly shaped for personal and professional reasons. The findings point to the importance of developing a holistic perspective toward educational research in order to promote research engagement, making it a sustainable path to professional excellence for language teachers.

**KEYWORDS:** Factor analysis; Language teachers; Motivations; Questionnaire; Research engagement

## 1. Introduction

Within the last decade, the desire to foster the relationship between Iranian English language teaching (ELT) research and practice has been made apparent by several initiatives and sustained strands of inquiry intended to encourage teacher-conducted research. The annual convention of Teaching English Language and Literature Society of Iran (TELLSI) in 2012 seemed to provide the first serious impetus to explore the role of teachers in ELT research. The lectures delivered by keynote speakers on the inauguration ceremony focused, for the most part, on the criticism that much of current research in academia is poor in terms of lack of relevance to and genuine impact on practice. The most explicit criticism was made in a symposium by Samar and Mehrani (2012).

> The concurrence of several syndromes in our research accounts for why our academic research is dysfunctional in addressing practical aspects of language education. Our researchers very often address topics in which teachers have little, if any, interest; their studies are often conducted in academic settings and on university students, subjects who are better understood as language users rather than language learners; they often use fairly sophisticated statistical procedures that are unfamiliar to many teachers; they write in a technical language that is too complex for teachers; and they often publish their studies in journals that teachers "have never heard" of.

Following this wave of criticism, several initiatives were made to make ELT research more applicable and accessible to practitioners. For instance, *Roshd Foreign Language Teaching Journal* commenced to publish a special column in each issue

**Cite this article:** Mehranirad, M. (2023). English teachers' motivation for research engagement. *Applied Linguistics Inquiry*, *1*(1), 67-73. doi: 10.22077/ali.2022.5574.1004S

where teachers report on action-research studies they carry out in their own classroom settings. A further section, entitled "my contribution" was also devoted to teachers for sharing their lesson plans and procedures for class activities. More recently, private language institutes such as 'Iran Language Institute' have launched several professional periodicals where academic research findings are translated into pedagogically practical ideas for language teachers. Even the scope of some Iranian academic journals shifted from a theoretical orientation toward a more practical endeavor. For example, in an attempt to better reflect its current focus, the formerly known *Journal of Applied Linguistics* experienced an appellation change to *Journal of English Language Pedagogy and Practice*.

Along with these alternations of policy, several areas of inquiries have recently emerged. Characterized by a drive to promote "teacher-research movement", some researchers, for instance, have appraised ELT research for focusing on too trivial issues (Mehrani & Khodi, 2014; Mehrani, 2014; Samar et al., 2012) and yielding inconclusive findings (Pieters & de Vries, 2007). Others have challenged the traditional conceptions of research dominantly held by Iranian researchers and policy makers (e.g. Rahimi, Madani & Rahimi, 2016) and have demanded for a reform in educational policies (e.g. Rahimi & Askari Bigdeli, 2016). In addition, several researchers have explored the problems that teachers experience in engagement with research (e.g. Dehghan & Sahragard, 2015). A critical review of these studies, however, shows that attempts to bring research and practice into a closer harmony are marked by disappointments. For example, Mehrani and Behzadnia (2013) vehemently contend that reducing the gap requires radical political actions and "given the current socio-political and economic conditions of Iran, it seems that at this time no such action is likely" (p. 28). This argument seems justified because systematic exchange structures between the two communities of research and practice are missing. That is, there is no institution, community, agent or even individuals with a responsibility of turning educational insights into pedagogical impact.

Although the establishment of institutionalized pathways between academics and practitioners can be conducive, lack of such mediating channels in the Iranian ELT profession ought not to be perceived as a fiasco for research to influence practice. In fact, discussions in educational science advocate various venues for educational research to provide insights for practice. One such route is to encourage teachers to do systematic inquiries in their own classroom settings. Within this relatively new paradigm, variously known as "exploratory practice", "teacher-researcher movement", "practitioner research", teachers are particularly expected to engage in research projects and address their pedagogical concerns through conducting classroom-scale investigations (Dehghan & Sahragard, 2015).

A fundamental argument underpinning the promotion of research engagement among teachers has to do with the inadequacy of traditional frameworks of educational research. In fact, recent critical reports have "vented serious doubts about the quality and relevance of educational research", claiming that traditional paradigms of research often fail to provide a systematic, refined basis on which educational professionals can build their work (Biesta, 2007, p. 2). Only rarely do educational research studies offer conclusive and practical results (Broekkamp & van Hout-Wolters, 2007) and are most often biased and socio-politically loaded (Pring, 2000).

Although engagement in research seems to be a challenging endeavor for Iranian language teachers, as our educational system has historically considered research luxurious, non-compelling, and voluntary (Mehrmohammadi, 1997), the literature does offer numerous benefits for teachers to embark on research. One such benefit is that research engagement can inform teachers' pedagogical decisions with sound research evidence, and this will have beneficial effects on their professional development. In addition, it can reduce teachers' feelings of frustration and isolation (Roberts, 1993) and push practitioners to move out of their submissive position and take a much more innovative role in education (Gurney, 1989). Mehrani (2017) maintains that engagement in research broadens teachers' understanding of language education, provides them with a framework for reflecting on their practice, empowers them to take leadership in educational changes and heightens their awareness of students' needs. Through engagement with research, teachers can generate and improve their local knowledge of teaching while also continuously addressing their pedagogical problems (Darling-Hammond, 2006). This, in turn, makes teachers less vulnerable to and dependent on academic researchers (Hammersley, 2004), as they develop their capacity for autonomous professional judgments (Lankshear & Knobel, 2004).

The exhaustive list of benefits reviewed above provides quite an impetus for promoting research engagement among teachers; nevertheless, it does not necessarily reflect the perspective of teachers for getting involved in research. In other words, the literature on the advantages of research engagement is characterized by scholars' hypothetical speculations. The challenge is, thus, that these alleged benefits remain unapproved and ought to be inspected through empirical investigations. In particular, studies are required to examine how engagement in research can influence teachers' profession. In doing so, it is important to take teachers' opinions on board, so as to provide an insider account of the incentives and advantages of research engagement.

A further confounding issue in the literature is that the presumed benefits of teachers' research engagement are each articulated for a particular local context, yet they are largely assumed to be generalizable to various educational situations. This is not justified because research-practice divide is a consequence of the interaction of an array of inter-dependent factors that lie deeply textured in the educational system. Thus, investigations into such a multi-layered phenomenon need to be ecologically valid so as to capture the complexities involved in each particular setting.

A problem in examining the benefits of research engagement is the absence of any instrument for reflecting teachers' opinions. That is, to the best of my knowledge, there is not any published piece of research in applied linguistics (particularly in the Iranian context) to empirically report teachers' perceived benefits of research engagement. The purpose of the present study is, therefore, twofold. First, the study intends to design an instrument to allow for the quantification of the benefits of research engagement. Secondly the study aims to empirically examine the extent to which Iranian English teachers engage in research and report on their perceived benefits of research engagement. Empirical inquiries of this type are essentially required for efforts made to diminish the gap between research and practice, because they can inform policymakers and stakeholders of what teachers really think about doing and using research.

## 2. Method

### 2.1. The development of a survey instrument

Following the guidelines suggested by Dornyei (2003) and Brace (2004), the researcher designed, developed and validated a survey research instrument. In doing so, first a general framework with three independent sections was designed. While, the first section intended to collect participants' demographic information, the second section included four open-ended prompts for examining teachers' level of engagement in research. In designing the third section, which was devoted to teachers' perceived benefits of research engagement, a standard procedure of instrument development was carefully followed. Initially, a comprehensive review of the related literature was conducted to shortlist any benefits presumably associated with research engagement. In doing so, the advantages reported in previous exploratory studies and researchers' speculations were all jotted down and resulted in a provisional item pool of over 40 benefits. Through a tripartite cycle of development, arrangement, and categorization, the items with overlapping contents were then removed, and this led to the development of a provisional data driven model with 19 items in four overall categories of benefits: developing teachers' professional knowledge, improving educational instruction, meeting the requirements of educational institution and fulfilling teachers' personal and instrumental motivations.

In the next stage, each item was formulated to fit into a five-point Likert rating scale through which respondents were supposed to indicate the extent to which they agree or disagree with each item. Subsequently, seven experts with a background in language teacher education were asked to review and revise the questionnaire items. Their comments, resulting in alternations in the wordings of some items, were taken into account, and in order to detect any ambiguity in the items, the final version was piloted by administering it to a group of 24 EFL teachers. The Cronbach's Alpha reliability coefficient of the instrument was calculated and yielded 0.91. Having ensured the precision and clarity of the items, the questionnaire was then administered to the participants, as described below.

### 2.2. Participants

The validation of the instrument proceeded by disseminating over 530 copies of the questionnaire among English teachers in different cities covering central, eastern, and northern parts of Iran. Both hard copies and email attachments were used for instrument distribution. A total of 407 copies were completed by the respondents and returned to the researcher. Upon initial inspection, however, the researcher had to remove 84 questionnaires because they either were not carefully completed or were uncompleted. Through a second round of inspection, 108 of the completed questionnaires were also discarded because in response to the questions in the second section of the instrument participants reported a low or lack of research engagement. The screening process at this stage was to exclude teachers who reported that they were never involved in any research projects, nor did they read research studies regularly. This resulted in 215 completed instruments by teachers who reported a moderate-to-high level of research engagement.

Analysis of respondents' demographic information showed that teachers' experience ranged from 1 to 32 years, though the majority had less than 10 years of experience. About 79% of the sample had Bachelors' degree, 15% had postgraduate qualifications, 2.5% had professional upper diploma, and 3.5% either had diploma or did not specify their qualifications. Over 40% of the respondents were affiliated with educational institutes in the private sector, and 59% were engaged in public schools.

### 2.3. Validation procedure

As was explained, the survey developed in this study for investigating the benefits of teachers' research engagement was substantiated by the domain knowledge and experts' opinions. However, since judgments were made a priori as to the number and nature of benefits, this conceptual model had to be validated so that it can be employed as a valid instrument for the present and future studies. Thus, teachers' responses given to each of the items were subjected to a factor analysis statistical procedure. The analysis was conducted in a two-step sequential principal axis factor analysis with varimax rotation in order to assess the underlying structure for the 19 items.

A preliminary analysis was conducted to determine how many factors can be optimally extracted in the main analysis. This was done through a scree plot and yielded in a break after the first four components. In other words, the scree plot indicated that four is the ideal number of categories of factors to account for the benefits of teachers' research engagement. This was in line with the hypothetical framework developed through the literature review. Therefore, four factors were requested in conducting the main analysis. The analysis resulted in four internally consistent factors. The first factor accounted for 14.53% of the variance, the second factor accounted for 14.49%, the third factor accounted for 11.72%, and the fourth factor accounted for 10.24%. Table 1 displays the items and factor loadings for the rotated factors, with loadings less than .30 omitted to improve clarity.

Table 1. Factor loading for the rotated factors of motivations for research engagement

| Item | Factor loading | | | | Communality |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Item 1 • | .801 | | | | .62 |
| Item 2 • | .691 | | | | .63 |
| Item 3 • | .535 | | | | .41 |
| Item 4 • | .531 | | | | .47 |
| Item 5 • | .433 | | | | .40 |
| Item 6 • | .379 | | | | .30 |
| Item 7 • | | .634 | | | .48 |
| Item 8 • | | .620 | | | .49 |
| Item 9 • | | .608 | .384 | | .65 |
| Item 10 • | .321 | .596 | | | .52 |
| Item 11 • | | .582 | | | .43 |
| Item 12 • | | .575 | | | .44 |
| Item 13 • | .378 | .390 | | | .43 |
| Item 14 • | | | .871 | | .70 |
| Item 15 • | | | .763 | | .63 |
| Item 16 • | | .431 | .600 | | .65 |
| Item 17 • | .344 | | | .817 | .68 |
| Item 18 • | .409 | | | .673 | .63 |
| Item 19 • | | .312 | | .639 | .61 |
| Eigen values • | 2.76 | 2.75 | 2.23 | 1.95 | |
| % of variance • | 14.53 | 14.49 | 11.72 | 10.24 | |

Note: Loadings < .30 are omitted.

As Table 1 indicates, the first factor, which indexes variables related to teachers' professional development, loads most strongly on the first six items. The second factor, which accounts for personal and instrumental incentives mostly associated with teachers' personal desires and financial issues, is composed of the seven items with loadings in the second column of the table. As Table 1 shows, the third, fourth and seventh items have their highest loadings on the second factor, but have moderate cross-loadings over .3 on the third and first columns too. The third factor, which indexes the items relevant to educational policies, comprises three items. The third item in this column, however, seems to be moderately associated with the second factor too. Lastly, the fourth factor indexes issues related to classroom procedure and pedagogical items. The first two items in this column show moderate loadings on the first factor, and the last item shows moderate loading on the second factor.

Based on the results, therefore, teachers' reasons and motivations for reading and doing research can be categorized in four groups: benefits for professional development, instrumental and personal motivations, organizational expectations, and pedagogical benefits.

## 3. Results

In order to gain insights into the importance of each item in rousing teachers' drive to engage in research, ratings given to the items were analyzed through calculating a weight score for each item. To this end, a proportional weight score was first obtained for each teacher's response to each item. The proportional weight scores were obtained by assigning a score of +2 to every "strongly agree" response and a score of +1 to every "agree" response. Similarly, "strongly disagree" responses were assigned a score of -2 and "disagree" responses were given a score of -1. Teachers' "undecided" responses received no score, and therefore, the range of proportional scores was between -2 and +2. For each item, the mean of the proportional scores was calculated to derive a weight score. Within this formulation, the magnitude of each weight score suggested the significance of the reason.

Respondents' data revealed that the most important reasons for teachers' research engagement were because they believed that reading and doing research help them develop their professional knowledge and skills, find better techniques and strategies for language teaching, and be aware of the recent developments in language teaching. These reasons clearly have a strong pedagogical focus. In contrast, more instrumental motives such as better income, and fulfilling authorities' expectations were less prominent in teachers' responses. The results of teachers' responses are shown in Table 2Table 2.

Table 2. Weight scores for motivations for research engagement

| Questionnaire items | No. of respondents[1] | Weight scores | SD |
|---|---|---|---|
| • مطالعه و انجام پژوهش به رشد و توسعه شغلی من کمک میکند. | 215 | 1.32 | .707 |
| • با هدف یافتن روشها و تکنیکهای بهتر در تدریس زبان به مطالعه و انجام پژوهش میپردازم. | 215 | 1.20 | .781 |
| • جهت آگاه بودن از مطالب جدید در رشته آموزش زبان مطالعه میکنم. | 213 | 1.17 | .760 |
| • با هدف افزایش دانش و آگاهیهای فراشناختی در فرآیند تدریس به مطالعه و انجام پژوهش میپردازم. | 215 | 1.11 | .824 |
| • از مطالعه و انجام کارهای پژوهشی لذت میبرم. | 215 | 1.02 | .823 |
| • برای برطرف کردن مشکلات و موانع تدریس زبان به مطالعه و انجام پژوهش میپردازم. | 215 | .94 | .884 |
| • برای تقویت رزومه کاری ام به مطالعه و انجام پژوهش میپردازم. | 214 | .82 | .962 |
| • برای بهتر کردن شرایط آموزشی مدرسه (مؤسسه) مطالعه و پژوهش انجام میدهم. | 215 | .78 | .914 |
| • با هدف پذیرفته شدن در دوره تحصیلی بالاتر به مطالعه و انجام پژوهش میپردازم. | 213 | .71 | .976 |
| • برای شرکت در جلسات و سمینارها نیاز به مطالعه و انجام پژوهش دارم. | 213 | .62 | 1.091 |
| • برای گذراندن دوره آموزشی که اکنون در حال طی کردن آن هستم مطالعه و پژوهش میکنم. | 213 | .58 | .936 |
| • برای بحث و تبادل نظر با همکاران و معلمان دیگر به مطالعه و انجام پژوهش میپردازم. | 211 | .52 | .917 |
| • برای کسب ارتقاء شغلی و ترفیع سازمانی نیاز به مطالعه و انجام پژوهش دارم. | 213 | .38 | 1.229 |
| • برای کمک کردن به همکاران و معلمان دیگر به مطالعه و انجام پژوهش میپردازم. | 213 | .35 | .923 |
| • با هدف یافتن شغل بهتر به مطالعه و انجام پژوهش میپردازم. | 213 | .25 | 1.170 |
| • با هدف چاپ و انتشار مطالب پژوهشی خودم به مطالعه و انجام پژوهش میپردازم. | 211 | .22 | 1.079 |
| • سیستم آموزشی از من انتظار دارد که به مطالعه و انجام پژوهش بپردازم. | 213 | -.08 | 1.171 |
| • مدیر مدرسه (مؤسسه) از من انتظار دارد که به مطالعه و انجام پژوهش بپردازم. | 214 | -.12 | 1.181 |
| • با هدف کسب درآمد بیشتر به مطالعه و انجام پژوهش میپردازم. | 211 | -.23 | 1.141 |

## 4. Discussion

The present study described the development and validation of an instrument for examining teachers' perspectives on their research engagement. For this purpose, a conceptual framework was first designed based on a thorough review of the literature, and then was tested on a sample of over 200 Iranian English teachers using factor analysis. The analysis corroborated the initial conceptual model, suggesting that from teachers' perspective the benefits of engagement in research can best be explained in terms of four overall categories: professional motivations such as improving the quality of education and professional development; instrumental and personal motivations including improving occupational resume, finding a better job, getting admission to higher educational programs, publishing manuscripts, obtaining a better income; institutional motives – both a priori motivations to get involved in research and also the posterior incentives experienced after research projects – such as fulfilling manager's expectations, educational system's requirements, and institutional promotion; and finally pedagogical concerns, including finding better techniques and strategies for teaching, developing knowledge and meta-cognitive understanding, and solving educational problems. Analysis of teachers' ratings to each item also revealed the significance of each of these reasons in the Iranian ELT context.

[1] Throughout this section, where totals in tables do not add up to 215, this is due to missing data.

One of the most interesting findings was that, consistent with the literature, the most important outcomes of research projects were not seen just about developing research skills, nor even about an increased understanding of the subject matter of the research. Although these were obviously very important for many teachers in this study, research was clearly seen as a vehicle for professional development that unlike any other form of professional development remains open to teachers throughout their carrier.

In the Iranian context, teachers' professional development programs normally consist of short term programs, often conducted by a group of "imported experts" (Mehrani & Behzadpoor, 2022). These programs are ideally assumed to provide teachers with a chance to meet colleagues and discuss their professional problems, and be exposed to stimulating new ideas. However, investigations show that the knowledge offered is generally far removed from the contexts of the teachers, and the situational factors affecting their classroom practices are not often taken into account (e.g. Khanjani, et al., 2017). As a consequence, the aim of increasing teachers' professional development is rarely achieved. Engagement in research, on the other hand, as an alternative form of professional development could be considered as customizable, in that teachers decide what kind of knowledge to receive how and when. Therefore, in pre-service teacher education programs, research engagement can be promoted as an on-going opportunity for developing teachers' professional skills based upon their own individual needs.

The results of the study also showed that for some individuals the motivations for research engagement are primarily instrumental. That is, for some teachers reading or doing research is not an end in itself, but a path which paves the way for obtaining other achievements such as a better job, salary raise, university admission for higher education, etc. This, by implication, points to the actual unimportance of some of these factors in our educational system. As a fact of the matter, the analysis of the weight scores reflects that many teachers indicated that they do not receive any financial raise in their salaries for doing or otherwise engaging in research. These findings suggest that instrumental motivations such as financial and intellectual reinforcements can function as efficient tools for encouraging teachers to get involved in educational research. Such instruments, thus, can potentially be employed in making the Iranian language education an evidence-based profession.

The examination of the weight scores also indicated that within our educational system organizational requirements, and principals' expectations do not currently have the potential to rouse teachers' drives for keeping up with research. Given the roles that such institutional motives could play in promoting teachers' research engagement it seems necessary that policy makers implement a set of promotive educational policies to push teachers further toward research.

Unlike institutional motives, teachers' pedagogical concerns appeared to be very influential in increasing the level of their research engagement. For example, teachers believed that research findings can provide them with an extra level of assurance about what works in practice, and help them develop a meta-cognitive knowledge about the process of teaching. Despite a wide degree of skepticism about researchers and research findings, the findings in this study suggest that many Iranian teachers still hold positive views toward research. However, this attitude is not likely to be sustained and profitably used, unless material conditions in schools, and more broadly in our educational system are altered, and unless teachers and researchers come together to learn about, conduct and discuss research.

## 5. Conclusion

The present study provided an empirical analysis of the research engagement among Iranian English teachers. Methodologically, the study was robust and resulted in a research instrument that can be used in other contexts. In terms of the findings, the main contribution here is the first-hand data in support of the argument that teacher engagement is by no means a simple, straight forward issue. Although, the participants involved in the study were exclusively limited to those teachers who were moderately-to-highly engaged in research, there was little evidence in favor of the many hypothetical claims made in the literature about the benefits of research engagement for language teachers. As a matter of fact, the results showed that research engagement among Iranian language teachers is mainly shaped for personal and professional reasons. Teachers do not often feel obliged by their institutions to engage in research, nor do they read and do research for receiving financial and intellectual supports. These findings suggest that initiatives made to promote research engagement would be likely to confound the status-quo, unless a holistic and thorough investigation is made in each educational context. In other words, arguing against the simplistic conceptualization of research engagement, the findings point to the fact that in promoting the notion of teacher as a researcher, the idiosyncrasy of each context must be taken into account.

## 6. References

Biesta, G. (2007). Bridging the gap between educational research and educational practice: The need for critical distance. *Educational Research and Evaluation , 13* (3), 295-301.

Brace, I. (2004). *Questionnaire design: How to plan, structure and write survey material for effective market research*. Kogan Page.

Broekkamp, H., & van Hout-Wolters, B. (2007). *The gap between educational research and practice: A literature review, symposium, and questionnaire , 13* (3), 203-220.

Darling-Hammond, L. (2006). *Powerful Teacher Education: Lessons from Exemplary Programs*. John Wiley and Sons, Inc.

Dehghan, F., & Sahragard, R. (2015). Iranian EFL teachers' views on action research and its application in their classrooms: A case study. *Journal of Teacher Education and Educators, 4*, 39-52.

Dornyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Lawrence Erlbaum Associates, Publishers.

Gurney, M. (1989). Implementer or innovator: A teacher's challenge to the restrictive paradigm of traditional research, in P. Lomax, (Eds.), *The Management of Change*. Multilingual Matters.

Hammersley, M. (2004). Action Research: a contradiction in terms? *Oxford Review of Education, 30* (2), 165–181.

Khanjani, A., Vahdany, F., & Jafarigohar, M. (2017). EFL teacher education in Iran: Does it promote trainees' pedagogical content knowledge? *Journal of Research in Applied Linguistics, 8*(2), 159-186.

Lankshear, C. &. Knobel, M (2004). *A handbook for teacher research: From design to implementation*. Open University Press.

Mehrmohammadi, M. (1997). Research in the Ministry of Education: Necessities, hopes. *Quarterly Journal of Education 50*, 45-47.

Mehrani, M. B. (2017). A narrative study of Iranian EFL teachers' experiences of doing action research. *Iranian Journal of Language Teaching Research 5(*1), 93-112.

Mehrani, M. B. (2014). Bridging the gap between research and practice: Voice of mediators. *Pan Pacific Association of Applied Linguistics, 18* (2). 21-38.

Mehrani, M. B. & Behzadnia, A. (2013). English teachers' research engagement: Current barriers and future strategies. *International Journal of Foreign Language Teaching and Research. 1* (4), 17-32.

Mehrani, M. B. & Khodi, A. (2014). An appraisal of the Iranian academic research on English language teaching. *International Journal of Language Learning and Applied Linguistics World, 6* (3) 89-97.

Mehranirad, M. & Behzadpoor, F. (2022). A survey of EFL teachers' research engagement. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances, 10* (1), 79-98.

Pieters, J. & de Vries, B. (2007). Preface to the special issue. *Educational Research and Evaluation , 13* (3), 199-202.

Pring, R. (2000) 'False dualisms: quantitative and qualitative research'. *Journal of Philosophy of Education, 34* (2), 247-260.

Rahimi, A. & Askari Bigdeli, R. (2016). Challenges of action research: Insights from language institutes. *Journal of Research in Applied Linguistics, 7* (2), 3-15.

Rahimi, A., Madani, A. & Rahimi, H. (2016). University teachers' action research: An adventure into uncharted waters? *International Journal of Learning and Teaching, 8* (4), 262-272.

Roberts, J. R. (1993). Evaluating the impacts of teacher research. *System 21* (1), 1–19.

Samar, R. G., Mehrani, M. B., (2012, October 16-19). *Does the supply meet the demands? The mismatch between researchers' interests and practitioners' needs*. [Paper presentation]. The 10th international TELLSI conference, Shahid Beheshti University, Tehran, Iran.

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 74-82*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2022.5755.1006

# Brown and Levinson's Politeness Theory Revisited: Rejection or Acceptance?

Seyyed-Foad Behzadpoor[1]

[1]Assistant Professor of Applied Linguistics, ELT Department, Azarbaijan Shahid Madani University, Iran

**ARTICLE HISTORY**

Received: 05 July 2022

Revised: 05 October 2022

Accepted: 27 December 2022

Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: fouad.behzadpour@gmail.com

**ABSTRACT**

In the present study, I have briefly revisited Brown and Levinson's (1987) universal theory which has attracted the attention of many linguists to the notion of politeness as an essential feature of communication. Although in studies on politeness Brown and Levinson's model is more or less considered the standard model, it seems that there is not an agreed upon consensus that the complex concept of politeness can be simply captured through this linear and static model. In addition, there is a paucity of investigations into the applicability of this model in various cultures. Therefore, I analyzed and identified the strategies found in Iranian English speakers' requests and apologies following this theory. To this end, adopting a qualitative research approach, data was collected through a multiple-choice discourse completion test and think aloud protocols. The results of the study revealed that Brown and Levinson's model can account, to a large extent, for people's choice of politeness strategies in making both requests and apologies. In particular, the findings pointed to the ubiquitous presence of three influential factors in the degree of politeness; that is, power relations, social distance and rank of imposition in the participants' choice of politeness strategies. However, the findings of the study indicated that there are a few shortcomings associated with Brown and Levinson's model. The study suggests that the weight of politeness cannot be simply measured based on a linear, static basis. The findings supported that for the individuals who participated in this study, politeness was a heavily context-bound and highly dynamic concept. On this basis, I hypothesize that a systematic model of politeness can better explain the variations of individuals' choice of politeness strategies.

**KEYWORDS:** Brown and Levinson's politeness model; Politeness models; Request strategies; Apology strategies; Iranian English speakers

## 1. Introduction

Politeness is an essential feature of communication, and it basically represents a speaker's social concerns about how to interact with others appropriately according to their personal status and social norms (Brown, 2001). People employ various politeness strategies to enhance the possibility of getting their messages across without damaging their social relationships with their interlocutors (Eshghinejad & Moini, 2016). The concept of politeness has been extensively investigated in different contexts for a number of good reasons and, consequently, various theories have been put forward to capture the complexity of politeness (e.g., Cruse, 2000; Ellen, 2001; Fraser, 1990; Scollon & Scollon, 2001). One central theory that has attracted many linguists' attention is Brown and Levinson's universal theory of politeness (1987).

According to Brown and Levinson (1987), in studies of politeness, two concepts play an important role: the concept of face and politeness strategies. In their formulation of politeness theory, face is defined as a "self-image" that involves

**Cite this article:** Behzadpoor, S-F. (2023). Brown and Levinson's politeness theory revisited: Rejection or acceptance?. *Applied Linguistics Inquiry*, *1*(1), 74-82. doi: 10.22077/ali.2022.5755.1006

people's emotions. In any normal communicative situation, attempts are made to maintain "self-image." This is often done through recognizing the desires of the interlocutor and understanding their wants. Failure to appreciate such wants and desires results in "self-image" to be lost. Brown and Levinson's (1987) politeness theory further explains which speech acts can threaten and/or protect face. It is also claimed to be capable of providing detailed explanations about when people are expected to protect face, and when they are more vulnerable to face threatening acts.

Brown and Levinson distinguish negative face from positive face. While negative face is considered to be the desire not to be imposed upon, positive face is believed to be the desire to have the approval or agreement of others. Moreover, they speculate that the distinction between positive and negative face is a universal phenomenon, but it can be the subject of cultural elaboration in any speech community (Brown & Levinson, 1987). In this regard, Mir-Fernandez (1994) maintains that the concept of face involves personal decisions about social values and norms for effective communication. On this basis, face is a reflection of cultural norms and can vary from context to context.

Based on Brown and Levinson, there are three important situational variables that influence people' choice of politeness level: the social distance between speaker and hearer, the relative power of speaker and hearer, and the absolute ranking of impositions within a particular culture. Thus, in their tripartite model, the weight of any face threatening speech act can be simply calculated as the sum of these variables (Vinagre, 2008).

Although in studies on politeness Brown and Levinson's model is more or less considered the standard model, it seems that there is not an agreed upon consensus that the complex concept of politeness can be simply captured through this linear and static model. In addition, there is a paucity of investigations into the applicability of this model in various cultures. Thus, this paper intends to apply Brown and Levinson's model for the analysis of politeness strategies employed by Iranian English speakers. In doing so, I follow Brown and Levinson's (1987) tripartite model of politeness variables in order to analyze and identify the strategies found in Iranian English speakers' requests and apologies.

## 2. Review of literature

Brown and Levinson's model of politeness has been subject to empirical investigations, and consequently various criticisms have been leveled against this theory. The first strand of criticism concerns the model's failure to account for cross-cultural variations. Despite the fact that Brown and Levinson claimed their theory is a universal model of politeness, many researchers have cast doubts on the universality of this theory. Central to arguments against Brown and Levinson's model is the claim that it only reflects the norms of highly individualistic societies (Wierzbicka, 1991). For instance, studies on the conception of face, conducted in Asian contexts such as China, resulted in significantly different results (Mao, 1994). In addition, Mao (1994) and Matsumoto (1988) found out that in the Japanese culture the distinction between positive and negative face with equal weight does not hold. Moreover, they reported that in the East the social self is more highly valued than the individual self.

In a different study, Leech (2007) discussed the limitations of Brown and Levinson's theory of politeness. He proposed a super constraint for the study of politeness known as "grand strategy of politeness." Leech (2007) argued that in order to be polite, speakers (especially in more socially stratified contexts) may express or imply meanings that associate a high value to the hearers or meanings that associate a low value to themselves (speakers). The grand strategy of politeness is largely comparable to Brown and Levinson's positive and negative politeness strategies. Leech (2007) also hypothesized that these variables are differently manipulated in different cultures and can reflect the social norms and particular concepts of face in a given society.

Another line of criticism of Brown and Levinson's model is made on the fact that the concept of politeness basically cannot be quantified. Some researchers have particularly challenged Brown and Levinson's attempts to measure the weight of politeness as the sum of particular linguistic tokens. Arguing for politeness as appropriateness, Locher and Watts (2005) and Watts (2003), for example, set the aim of politeness studies as the folk interpretation of politeness and claimed that it is the only logical means of developing a social theory of politeness. From the perspective of these researchers, determining appropriateness is extremely dependent on the local contexts and thus should be only examined by considering the entire context.

Acknowledging the existence of contextually-sensitive and socially-loaded conception of politeness, some linguists, however, focus on the linguistic aspects of politeness as the object of their studies and, therefore, argue that the quantitative studies of politeness should also be possible. Such studies, if successfully designed and conducted, could serve as empirical tests of various models of politeness.

A further wave of challenge toward Brown and Levinson's model has concerned the overriding emphasis on face as a main motivation for politeness (Fraser, 1990; Locher & Watts, 2005). On this score, some linguists such as Locher and Watts argue that Brown and Levinson's theory is indeed a framework for the study of face not a theory of politeness. These authors suggest that Brown and Levinson's model of politeness could be integrated into a larger theory of relational work, where accounts of face, politeness, and political behavior are all taken into account.

Undoubtedly, most interactions can best be understood by taking into account the entire contextual situations. However, in any theoretical study, a certain level of abstraction is unavoidable for the goals of the theory-based studies of language. It might be well the case that some contextual features of language are lost in an abstract formulation of language. Therefore, whether or not the concept of politeness formulated based on a linguistic examination is valid is an empirical question, one that this study intends to explore.

## 3. Method

### 3.1. Participants

The participants of the study were 20 male (7) and female (13) university students studying English at Ershad University, Tehran, Iran. The participants were all native speakers of Persian and had not resided in any English-speaking country prior to their participation in this study.

### 3.2. Design

In order to identify and analyze the decisive factors in determining the participants' choice of politeness strategies, a qualitative approach was adopted in which concurrent think aloud protocols were employed to gain deeper insights into the latent and unobservable learning processes that occurred in participants' minds as they engaged in the designated research task.

### 3.3. Instruments and data collection procedure

Since there was not any appropriate research instrument for the purpose of this study, the researcher developed a multiple-choice discourse completion test/questionnaire. The first part of the questionnaire included demographic information along with the necessary directions and instructions to be followed by the participants to do the required tasks. Following the first part, the second part was devoted to five scenarios each of which involved a situation which required a request. Each scenario was followed by three possible request options one of which was supposed to be chosen by the participants. The third section included five scenarios each including a situation in which an apology was demanded; again, each situation was followed by three apology choices to be selected by the participants based on their preferences (see Appendix). The participants were individually invited to participate in a series of concurrent think aloud protocols as they were selecting the items of the questionnaire. Indeed, the participants were asked to expand on their responses. In particular, they were asked to explain why they chose particular politeness strategies. The think aloud protocols were audio recorded and transcribed in full. The analysis of the transcriptions involved mapping the participants' comments onto the section of the questionnaire they related to. The data provided broad categories for analyzing concerns they had in their strategy choices. The transcripts were then coded in relation to the broad categories formulated in Brown and Levinson's theory (1987).

## 4. Results

The purpose of this study was to explore the main factors that influence the politeness strategies that Iranian English language learners employ when they apologize and make requests. The summaries that follow represent the participants' explanations in the think aloud protocols. Attempts have been made to provide relevant quotes from the participants' responses to convey key themes and what they typically said about their choice of strategies. It should be mentioned, however, that although the participants who participated in this study were heterogeneous in many respects, the obtained qualitative results cannot be generalized to other cultures. Therefore, the findings in this study provide provisional information about the decisive factors in determining politeness strategies.

### 4.1. Analysis of requests

The first socio-cultural theme that emerged from the analysis of the think aloud protocols was the power distance that the participants felt between themselves and the people they were assumed to be interacting with. This was particularly evident in scenarios where the participants were supposed to request something from a person with whom they had an asymmetrical social relationship (e.g., scenarios 1 and 4). The following assertions made by some of the participants capture this concern:

> It does not seem appropriate to ask a teacher in this way…
>
> Because, I should ask my professor very politely to explain the issue for me again.
>
> Basically, when I talk to my teachers, I use formal and respectful wordings.

When you speak so boldly and directly to your instructor, he will definitely get offended.

These assertions imply that from the participants' perspectives, the type of power relationship that they have with the interlocutors is, perhaps, an important factor in determining the degree of their politeness.

Another noteworthy, yet interrelated, theme in the participants' explanations was that some of the learners acknowledged that they sporadically use less elaborate positive strategies or choose to use positive rather than negative politeness when speaking with friends, classmates or family members. They could clearly distinguish kin or friend from people with whom they were of the same social status, but who were still separated by social distance. For instance, in their elaborations, some of the participants expressed that their close relationships with their classmates and family members (as reflected in scenarios 2 and 5) were a driving force behind their choice of politeness strategies:

I choose to say, "You'd better move away or sit down," because I am talking to my sister and my tone of language must be both informal and intimate.

I would simply ask him [my classmate] to give me a piece of paper because I don't think I have to say very formally "I wonder if you…". It is also more common and typical.

You know, here I just want to talk to my classmate not my supervisor. I would try to observe my social bonds and respect our intimacy. When you use a formal language to talk to a classmate, it seems like s/he owes you something.

Here you are in your home, talking with your sister over a football game. The whole context is rather informal, and considering your relationship you are not supposed to speak formally and for a long time to ask for a simple thing… It is also possible to ask her just by a simple gesture or a facial expression.

Consistent with the widespread assumption that the social distance between the parties involved in interpersonal communications plays a key role in determining the quality and choice of politeness strategies, the above quotes suggest that this assumption holds true for the participants of this research as well. Although these quotes do not clearly show how this assumption can be formulated with regard to the concrete activities happening in particular situations, they do emphasize its presence in the eyes of Iranian English language learners.

The review of the literature showed that the absolute ranking of the threat of the face-threatening act is also among the decisive factors in determining the choice of politeness strategies. For the learners involved in this study, weight of imposition was very often a key factor in the development of their interlanguage pragmatics.

It is a simple request I am trying to make…I have not done anything wrong, I just want her to show me the shirt…Just an "excuse me" would suffice.

I prefer the third option because a salesperson's job is to show his stuff and to sell them anyway, and that is exactly what I am asking him to do for me.

I have never seen anybody saying please do me a favor in situations like this.

Although in the scenarios tested in the present study the participants were not asked various questions where they could show the full potential of their pragmatic knowledge, they were, as the above assertions show, quite aware of the fact that the mere magnitude of the request can, to some extent, fluctuate the politeness level of their verbal communications.

## 4.2. Analysis of apologies

Participants' responses to the request items revealed the significance of some of the determining factors in the degree of their politeness strategies. Throughout the follow-up apology scenarios, they were further asked to explain and elaborate on their choice of strategies. In line with the theoretical frameworks in the literature, the findings of the previous section revealed that to determine what is pragmatically appropriate, Iranian English learners evaluated many of the contextual factors surrounding the scenarios. Some specific examples of how these factors affected their strategy choices were mentioned. The three most common factors affecting the politeness of an utterance were shown to be social distance, power relations and level of imposition. In what follows, attempts are made to discuss these results with direct quotes from the participants used to illustrate and exemplify the points.

As mentioned before, social distance refers to the relationship between the interlocutors. If two people are very close, they will have a low degree of social distance. Two strangers would typically have a high degree of social distance. In some of

the scenarios used in this experiment, a high degree of social distance was assumed to exist between the parties. The participants' choices of strategies were consequently rather formal.

> I explain to convince him that it was my mistake.
>
> I would definitely choose C, because it is more formal and more polite.

Basically, we can imagine three types of power relationships between interlocutors. In the first case, one would have equal power with the person he is talking to (e.g., a friend or colleague). In the other two, one would either have more power (e.g., as a boss, teacher) or less power (e.g., employee, student) than the person he is talking to. As the following assertions show, in traditional contexts like Iran, more formal and indirect language is typically used in situations where one is engaged in an interaction with the person who has more power (Rudy & Grusec, 2006).

> I would choose A because I cannot do anything else… option B seems to be too rude, and C is too formal, classmates do not usually talk in this way.
>
> When I apologize and explain for my teachers the reason why I fell asleep…
>
> I have to respect him because the teacher may get angry… so I have to be polite.

Additionally, rank of imposition seemed to have been considered seriously by the learners. As a matter of fact, the explanations that the participants provided seemed to have augmented the reasons that they previously mentioned in the case of request making.

> Because I have hurt him… and this is the least I can do.
>
> I need to buy a new book and make up for what I have done.

### 4.3. Problems associated with Brown and Levinson's model

The results presented above are mainly in line with Brown and Levinson's theory of politeness (1987). However, an in-depth analysis of the results revealed some further complexities in the participants' choice of politeness strategies. Particularly, the findings suggest that the factors that influence people's choice of politeness strategies do not function on a linear basis; rather, they seem to work best within a more dynamic system. In what follows I present parts of the participants' assertions that could not be properly explained in light of Brown and Levinson's theory.

> When I'm watching soccer, or any other program on TV, and somebody blocks my view, I would get angry and shout… I don't care who they are, they should not do that… I would say get out of my face…
>
> It depends on what has made me ask that question. It might be the case that you have a positive relationship with your hearer, so you are normally expected to show respect, but if there is a negative background, you might be discouraged to be very polite in your request.
>
> It's not important what you're asking. It is important who you are talking to. When you talk to a professor, you are expected to show respect…

These quotations seem to imply that the level of politeness in communication is, prior to any choice of politeness strategies, determined at a higher level where the speakers make a *grand* decision to the effect whether or not they intend to behave politely. In other words, the data suggest that the influential factors in regulating speakers' choice of politeness strategies work within a broader system. Therefore, people's choice of level of politeness might be governed at a more holistic level.

Basically, a system is an inter-related set of elements that are organized into a structural whole that is often self-regulated. The concept of system implies that if a complex phenomenon has holistic properties, the phenomenon cannot be explored by analyzing the individual components separately. Understanding the inter-relationship among the components and the way they influence one another is also required (Kitao et al., 1987). As previously discussed, in Brown and Levinson's theory (1987), three independent variables are intended to explain the weight of face threatening acts and justify people's choice of politeness strategies. One of the limitations of applying this theory for analyzing the participants' choice of strategies is that it can explain how politeness strategies are chosen one at a time, but does not offer any explanations about how such strategies are adjusted over time. A systemic model of politeness, however, seems to have the potential of showing how different variables can work together to achieve the goal of speech acts while also maintaining the face. The following quotes stated by some of the participants support this systematic model.

> When somebody's rights are violated, then I might do anything to please him… In this scenario I did not pay the money, and I left the bookstore. So I have to go back and apologize, I don't think I would be able to justify my inappropriate behavior in a couple of sentences.
>
> Well I have to know how she is going to react, and then depending on her behavior I would decide what to say and how to say it.
>
> Sometimes you just say I am sorry, and he forgives you though he deserves the rights not to forgive you…

A further relevant problem associated with Brown and Levinson's model concerns the additive values of the three variables indicated by Brown and Levinson (the social distance between speaker and hearer, the power relations between speaker and listener, and the rank of imposition). As previously mentioned, this attempt has been criticized for being over-simplistic. The findings of this study show that the emphasis on the linear calculation of politeness is perhaps misguided, and that a more socially motivated conception of politeness might be more fruitful in the analysis of politeness.

The quotations cited above show that the interlocutor's social motivations are further driving forces that are missing in Brown and Levinson's model of politeness. Sankoff and Laberge (1978) refer to these influential factors as the *speaker investment.* Speaker investment in Sankoff and Leberge's words represents the extent to which speakers intend to appear polite in their social interactions. It can also account for variations of politeness in different cultures.

## 5. Discussion and conclusion

In this paper, Brown and Levinson's model of three contributing factors to politeness was employed to analyze the requests and apologies made by Iranian EFL learners. The study aimed at demonstrating the potential of Brown and Levinson's model in determining politeness strategies employed by the participants. The results of the study revealed that Brown and Levinson's model can account, to a large extent, for people's choice of politeness strategies in making both requests and apologies. In particular, the findings pointed to the ubiquitous presence of three influential factors in the degree of politeness; that is power relations, social distance, and rank of imposition in the participants' choice of politeness strategies. However, the findings of the study indicate that there are a few shortcomings associated with Brown and Levinson's model.

The study suggests that the weight of politeness cannot be simply measured based on a linear, static basis. The findings support that for the individuals who participated in this study politeness is a heavily context-bound and highly dynamic concept. On this basis I hypothesize that a systematic model of politeness can better explain the variations of individuals' choice of politeness strategies. Such a model can also incorporate the interlocutors' *grand* decision concerning their commitment for polite behaviors. Therefore, the findings provide empirical support in favor of Leech's arguments (2007) for the existence of a grand strategy of politeness. Leech discusses that a super constraint governs the whole politeness phenomenon in our daily interactions. Leech (2007) further hypothesizes that a grand strategy of politeness can reflect the social norms and particular concepts of politeness in different speech communities.

Although parts of the findings exhibit certain features of politeness that are in accordance with the criteria of Brown and Levinson's theory of 'politeness', consistent with the literature, the results also partially indicate that there is considerable room for expanding Brown and Levinson's model of politeness by focusing on the broader context where social interactions take place. Granted the fact that Brown and Levinson's politeness formulation is basically a linguistic model rather than a social theory, one may assume that a non-linguistic model of politeness must necessarily take into account the role of context and cultural variations in determining politeness behaviors.

Although this study provides some empirical findings regarding the limitations of Brown and Levinson's theory (1978, 1987), it contains methodological limitations some of which may provide avenues for future research. The scenario items were limited both in number and in scope to investigate interpersonal relationships between a person and the interlocutor. If more variations in the scenarios were incorporated, perhaps more complexities of politeness as a social concept could be revealed, and I would be able to identify even further effects of factors influencing politeness strategies. Additionally, this study merely focused on two types of speech acts, namely requests and apologies, and I collected the data only through questionnaire items and think aloud protocols. Previous research suggests that politeness strategies are more complicated when observing authentic conversation data (Hayashi, 1988; Saito, 2010). Thus, future studies can analyze politeness features of authentic communications. Furthermore, complementary studies can address politeness strategies in other speech acts such as promising, informing, etc.

## 6. References

Brown, J. D. (2001). Pragmatics tests: Different purposes, different tests. In K. Rose , & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 301-325). Cambridge University Press.

Brown, P., & Levinson, S. (1978). Universals in language usage: Politeness phenomena. In E. N. Goody (Ed.), *Questions and politeness* (pp. 56-311). Cambridge University Press.

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

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Cruse, A. (2000). *Meaning in language: An introduction to semantics.* Oxford University Press.

Ellen, G. (2001). *A critique of politeness theories.* St. Jerome Press.

Eshghinejad, S., & Moini, M. R. (2016). Politeness strategies used in text messaging: Pragmatic competence in an asymmetrical power relation of teacher–student. *SAGE Open*, 1-13. doi:https://doi.org/10.1177%2F2158244016632288

Fraser, B. (1990). Perspectives on politeness. *Journal of Pragmatics, 14*, 219-236. doi:https://doi.org/10.1016/0378-2166(90)90081-N

Hayashi, R. (1988). Simultaneous talk from the perspective of floor management of English and Japanese speakers. *World Englishes, 7*, 269-288. doi:https://doi.org/10.1111/j.1467-971X.1988.tb00237.x

House, J., & Kasper, G. (1981). Politeness markers in English and German. In F. Coulmas (Ed.), *Conversational routines: Explorations in standardized communication situations and prepatterned speech* (pp. 157-185). Mouton.

Kitao, K., Munsell, P., Kitao, S., Yoshida, S., & Yoshida, H. (1987). An exploratory study of differences between politeness strategies used in requests by Americans and Japanese. *Educational Resources Information Center (ERIC) Document No. ed. 284426*.

Leech, G. (2007). Politeness: Is there an East-West divide? *Journal of Politeness Research, 3*, 167–206. doi:https://doi.org/10.1515/PR.2007.009

Locher, M. A., & Watts, R. J. (2005). Politeness theory and relational work. *Journal of Politeness Research, 1*, 9-33.

Locher, M. A., & Watts, R. J. (2005). Politeness theory and relational work. *Journal of Politeness Research, 1*, 9-33. doi:https://doi.org/10.1515/jplr.2005.1.1.9

Mao, L. (1994). Beyond politeness theory: "Face" revisited and renewed. *Journal of Pragmatics, 21*, 451-486. doi:https://doi.org/10.1016/0378-2166(94)90025-6

Matsumoto, Y. (1988). Reexamination of the universality of face: Politeness phenomena in Japanese. *Journal of Pragmatics, 12*, 403-426. doi:https://doi.org/10.1016/0378-2166(88)90003-3

Mir-Fernandez, M. (1994). The use of English requests by native Spanish Speakers and its relation to politeness values. *Unpublished doctoral dissertation*. Urbana-Champaign: University of Illinois.

Olshtain, E., & Cohen, A. (1983). Apology: A speech act set. In N. Wolfson, & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 18-35). Newbury House.

Rudy, D., & Grusec, J. E. (2006). Authoritarian parenting in individualistic and collectivistic groups: Associations with maternal emotion and cognition and children's self-esteem. *Journal of Family Psychology, 20*, 68-78. doi:https://psycnet.apa.org/doi/10.1037/0893-3200.20.1.68

Saito, J. (2010). Subordinates' use of Japanese plain forms: An examination of superior-subordinate interactions in the workplace. *Journal of Pragmatics, 42*, 3271-3282. doi:https://doi.org/10.1016/j.pragma.2010.06.014

Sankoff, D., & Laberge, S. (1978). The linguistic marketplace and the statistical explanation of variability. In D. Sankoff (Ed.), *Linguistic variation: Models and methods* (pp. 239-250). Academic Press.

Scollon, R., & Scollon, S. W. (2001). *Intercultural communication: A discourse approach* (2nd ed.). Blackwell.

Vinagre, M. (2008). Politeness strategies in collaborative e-mail exchanges. *Computers & Education, 50*, 1022–1036. doi:https://doi.org/10.1016/j.compedu.2006.10.002

Watts, R. J. (2003). *Politenes.* Cambridge University Press.

Wierzbicka, A. (1991). *Cross-cultural pragmatics: The semantics of human interaction.* Mouton de Gruyter.

## 7. Appendix

## Pragmatics Test: Request and Apology

## Participant Background:

a. First Name: Last Name:
b. University Degree:
c. Major:
d. Gender: ☐Male ☐Female
e. Residence in an English-speaking Country: No☐ Yes ☐ (If yes, for …… years)

**Directions:** Please read each of the following situations in which a person makes a/an apology/request. There are three answers following each situation. While you are reading each situation to decide which one is the **most appropriate** answer, think aloud to show the criteria you use to select the correct answer and reject the incorrect answers. Then circle the letter of the correct answer.

## I. Request

**1. You have a listening class and you cannot hear what is played on T.V. How would you ask your teacher to turn it up?**
A. Pardon me, but I cannot hear.
B. I'll ask you to turn it up.
C. What? Turn it up please.

**2. You need a piece of paper to write a letter. How would you ask your classmate for it?**
A. Can you give me a piece of paper? I need it to write a letter.
B. I wonder if you could possibly give me a piece of paper for my letter.
C. Hey, don't hesitate to give me a piece of paper. I need it right now!

**3. You are now shopping in a department store. You see a beautiful shirt and want to see it. How would you ask the salesperson to show you the shirt?**
A. Oh, sorry, could you pass that shirt to me to have a look? I want to buy it.
B. Lady, I'd like to have a look at that shirt. Would you please do me a favor?
C. Excuse me. Could you show me this shirt please?

**4. You are now discussing your assignment with your teacher. Your teacher speaks very fast. You do not follow what he is saying. How would you ask your teacher to say it again?**
A. You speak so rapidly that nothing can be understood.
B. Sorry, teacher, repeat the point. I didn't get it.
C. Excuse me. May I ask you to explain it again?

**5. You are watching a football game. Your sister comes and stands just in front of you blocking your view. You want to ask her not to block your view. What would you say?**
A. So you are interested in football. So am I. Let me stand beside you and exchange opinions about the game.
B. Sorry, you are blocking my view. Would you please take another place?
C. Hey. You'd better move away or sit down.

## II. Apology

**1. In a bookstore, you accidentally find a book that you have been looking for a long time. You are so excited that you rush out of the bookstore with the book without paying for it. The shop assistant stops you. How would you apologize?**
A. Sorry! I was too happy! I like this book and have been looking for it for a long time.
B. Excuse me. I've been looking for the book for a long time. I hope you can forgive my behavior. I'll be careful next time.
C. Oh, I'm very sorry. I was so excited about finding this book that I just forgot to pay. How much do I owe you?

**2. You are now rushing to the classroom. When you turn a corner, you accidentally bump into a student whom you do not know and the books he is carrying fall onto the ground. How would you apologize?**
A. Oops, sorry, my fault. I'm in such a hurry. Here let me help pick these up for you.
B. I will be late if I'm not in a hurry. I'll pay attention to this when I turn corner next time.
C. Oh, I'm ashamed. I'm going to be late for my class, and if I'm late, I won't be allowed to enter the classroom. But I like this course very much. Take care!

**3. You borrowed a book from a library, but you accidentally spilled a cup of coffee all over it. You return it to the librarian. How would you apologize?**
A. Sorry, it was an accident, chill out.
B. I am deeply sorry. Please allow me to replace the copy.
C. Well, accidents happen, you know?

**4. You are playing on the playground with your classmates. You take a shot and the ball hits a student on the back of the head. You go up to the student. How would you apologize?**
A. Are you all right? I'm sorry I hit you!
B. Dear student, I'm sorry for that! We have to be more careful.
C. I didn't realize you were coming. You'd better move away.

**5. You are almost asleep in the class while the teacher is teaching. The teacher gets very angry when he sees you sleeping in the class. How would you apologize?**
A. I’m sorry; I will try and not let it happen again.
B. I’m sorry, but I didn’t sleep a wink last night.
C. Pardon me. I’m ashamed.

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 83-94*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2023.6289.1014

# Computerized Dynamic Assessment of Incidental Vocabulary Learning: A Case of Iranian ESP Learners

Sepideh Hanifi[1]
Mahdi Nasiri[2]
Seyed Hesamuddin Aliasin[3]

[1]PhD student, English Language Department, Faculty of Humanities, Bu-Ali Sina University, Hamedan, Iran
[2]Assistant Professor, English Language Department, Faculty of Humanities, University of Zanjan, Zanjan, Iran
3Assistant Professor, English Language Department, Faculty of Humanities, University of Zanjan, Zanjan, Iran

**ARTICLE HISTORY**

Received: 27 August 2022
Revised: 27 October 2022
Accepted: 24 January 2023
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: s.hanifi@ltr.basu.ac.ir

**ABSTRACT**

Dynamic assessment (DA) has been widely researched in different linguistic areas, but there is paucity of research on its incorporation into English for Specific Purposes (ESP). Accordingly, this study investigates the effectiveness of DA on incidental words emerging in technical reading textbooks with a focus on electronic engineering students. The research method is a quasi-experimental research design focusing on an intact group of 25 Bachelor of Science students of electronics from the University of Zanjan, Iran. The instruments used are Preliminary English Test (PET), a vocabulary knowledge scale, and a mediation test (a 13-item multiple-choice test). The mediation test was delivered through the researcher-developed website, designed for this purpose. As a triangulation, participants' evaluation of computerized dynamic assessment (CDA), too, was elicited using a survey adopted from Nirmalakhandan (2007). The results of the qualitative and quantitative phases indicated that Electronic students' incidental vocabulary learning promoted dramatically using target CDA, employing structured prompts for the mediation process. The results of this study can inform both teachers and learners by providing a step-by-step procedure for both teaching and assessment of ESP learners' vocabulary.

**KEYWORDS:** Computerized dynamic assessment; Incidental vocabulary learning; English for specific purposes; Zone of proximal development; Mediation

## 1. Introduction

Assessment and why to be assessed is an important everlasting query of most teachers and learners, but the point to be made is that whatever the approach and ideology to learning and teaching are, assessment is an integral and inseparable part of the learning process (Poehner, 2008). Recent approaches and ideologies concerning language learning have led to the transformation of the concepts and bases of assessment, resulting in a shift in the relationship between assessment and teaching. Understanding the inadequacies in traditional assessment, scholars (e.g. Sternberg & Grigorenko, 2002) called for a more comprehensive approach so as to gauge multifarious aspects of learners' capabilities and help them enhance learning. The response to the call was that of integration of assessment and teaching with the goal of enhancing learning.

**Cite this article:** Hanifi, S., Nasiri, M., & Aliasin, S. H. (2023). Computerized dynamic assessment of incidental vocabulary learning: A aase of iranian ESP learners. *Applied Linguistics Inquiry*, *1*(1), 83-94. doi: 10.22077/ali.2023.6289.1014

Built on Vygotsky's sociocultural theory of mind, dynamic assessment (DA), as a uniform integration of teaching and assessment, aimed to address the deficiencies of traditional assessment. Since DA requires the examiner to take into account the examinees' zone of proximal development (ZPD), there are some impracticalities at work. This has made the use of DA for one-session mediation impossible. The idea of computerizing the implementation of DA procedures was a response to overcome the above-mentioned problem. Thus, computerized dynamic assessment (CDA) helped solve this problem by allowing the examiner to dynamically assess a greater number of learners and abilities in a single session (Poehner & Lantolf, 2013; Summers, 2008).

It is established that CDA puts emphasis on learning, and its product is successful learning (Lidz & Gindis, 2003), which is in line with the perspective dominant in English for Specific Purposes (ESP) (Hutchinson & Waters, 1987). By the same token, Hutchison and Waters (1987), for instance, put emphasis on ESP as an approach not a product, i.e. language learning is highlighted over language use. In fact, this is the reason why the present study attempted to introduce and probe the effectiveness of CDA in an ESP context.

Acquiring academic vocabulary is the main goal of ESP courses since university students are required to read English texts in their field of study so as to be able to obtain information regarding the latest technologies and advances, the prerequisite for which is having a good vocabulary knowledge (Nito, 2004). That is why teaching specialized vocabulary to ESP learners has been a concern throughout the history of ESP and language teaching (e.g. Dudley-Evans & St John, 1998; Sarani & Sahebi, 2012; Tangpijaikul, 2014).

The present study aims to answer the flowing questions:

1. Does CDA have an effect on incidental vocabulary learning by Iranian EFL learners in an ESP teaching/learning context?
2. What is the attitude of the participants to deployment of CDA in the ESP teaching/learning context?

## 2. Literature review

### 2.1. Dynamic assessment (DA)

DA challenged the view that assessment is seen as a distinct activity from teaching and argued that assessment and teaching are integrated rather than two distinct activities. In order to interpret learners' abilities and to lead them to higher levels of functioning, intervention is done during the assessment procedure and therefore, the two are integrated (Lidz & Gindis, 2003).

DA is theoretically based on two psychologists' work: it is based on Vygotsky's socio-cultural theory and ZPD and Feuerstein's structural cognitive modifiability theory and mediated learning experience (MLE). Based on Vygotsky's (1978) definition, ZPD is the distance between a child's actual developmental level, independent problem solving and the higher level of potential development, problem solving under adult guidance or in collaboration with more capable peers. ZPD is of great practical significance for education because of identifying how instruction can optimally develop learning by aligning mediation to abilities that are emerging, not those that are emerged (Xi & Lantolf, 2021). Therefore, ZPD is considered as the activity in which instruction leads to development. Moreover, Tzuriel (2001) defines MLE as an interactional process in which parents manipulate adult-child interactions in order to trigger child development.

DA is usually contrasted with static assessment (SA), which is not sensitive to ZPD. SA refers to the conventional forms of assessment in which assessment is performed without any intervention by the assessor, who just records learners' scores (Tzuriel, 2001); thus, DA, as an intervention-oriented approach, came as a result of dissatisfaction with SA as such (Lidz & Gindis, 2003).

There are two approaches toward DA, resulting from the two different interpretations of ZPD (Vafaee, 2011). The first interpretation is quantitative, leading to interventionist DA and the second one is a qualitative interpretation of ZPD, resulting in interactionist DA (Lantolf & Poehner, 2004; Lantolf & Thorne, 2006). Since mediation is the crux of DA, the difference between the two approaches is in terms of the mediation offered during assessment. Sternberg and Grigorenko (2002) differentiated between two formats of the interventionist DA, namely, the "sandwich" and the "cake" formats.

The sandwich format is evocative of traditional test formats and is conducted in the form of pretest-intervention-posttest either individually or in group size. According to Budoff (1968), test results are reported in the form of pre-training scores and post-training scores.

In the cake format, a standardized menu of hints, ranging from implicit to explicit, as for most interventionist approaches, is provided for the examinee during the administration of the assessment itself (Lantolf & Poehner, 2008; Poehner & Lantolf, 2021). Like the layers of a cake, the test items and hints are layered and a menu of hints is available for each question or problem before moving on to the next item on the test.

With the recent advances in technology, computer-based tests are being widely common and used. DA researchers have not lagged behind and are investigating ways to conduct mediation processes via the computer. CDA is quite new and few studies have been reported to use CDA to date. According to Poehner (2008), "CDA has several distinct advantages, including the following: it can be simultaneously administered to large numbers of learners; individuals may be reassessed as frequently as needed; and reports of learners' performances are automatically generated" (p. 177).

In recent years, DA has attracted the attention of many scholars around the world, and Iranian scholars are not exceptions. For example, MovahedFar et al. (2022) carried out a study to investigate the effect of CDA on Iranian EFL learners' performance in writing and their attitude towards CDA. They found that teaching and assessing writing skills through a computer can improve students' performance.

Estaji and Ameri (2020) examined the effect of interventionist approach to DA on Iranian English as a Foreign Language (EFL) learners' grammar achievement at two proficiency levels of pre-intermediate and high-intermediate. They found that the interaction effect of type of assessment and proficiency level was significant, which meant that the effect of DA on grammar achievement was higher for pre-intermediate learners than it was for high-intermediate learners.

Zangoei et al. (2019) consolidated assessment and instruction of L2 pragmatics comprehension through interventionist CDA. Like the present study, this paper provides three actual, mediated, and learning potential scores (LPS). The results of this study showed that the test could improve test takers' pragmatic comprehension competence.

Ebadi and Saedian (2015) investigated the impacts of CDA on promoting at-risk advanced Iranian EFL students' reading skills. Their results showed that this type of assessment may be most useful for individuals requiring a lot of attention, that is, at-risk or retarded learners.

Hessamy and Ghaderi (2014) attempted to investigate the role of DA in vocabulary learning of 50 intermediate Iranian EFL learners. They followed the sandwich format and found that incorporation of DA to standard testing had a positive effect on test performance and vocabulary learning by EFL learners.

Saeidi and Hosseinpour (2013) meticulously investigated the effect of DA on Iranian EFL learners' vocabulary learning. The results suggested that the rate of vocabulary learning can be enhanced using DA. This study is somehow similar to the present study, but there are two main points to be made here. First, their study dealt with DA, while the present study has employed CDA. Second, they conducted their study on general vocabulary learning, but this study was done on ESP vocabulary.

As the literature suggests, in the last two decades, the number of studies conducted on DA has increased significantly, each exploring different aspects of language learning within different frameworks (Qingha & Di, 2015), but to the best of the researchers' knowledge, there is paucity of research on the role of DA in vocabulary learning, in general, and CDA, in particular. Although there are some studies attempting to do so, all of them are in the context of general vocabulary and none has incorporated DA into the context of ESP and, specifically, vocabularies emerging in an ESP context.

## 2.2. Incidental vocabulary

Incidental vocabulary learning has been defined by various researchers from different aspects. For instance, Bogdanov (2013) defines incidental vocabulary learning as the learning of words that are the product of interactive activities, such as reading, listening, and communicating. It should be noted that the definition provided by Hill and Laufer (2003) has been adopted for this study, in which learning vocabulary is the by-product of another activity.

An important point about incidental L2 vocabulary acquisition lies in the fact that the teacher does not direct the learner's attention to the target words. For example, in Laufer (2006), "participants were not told that the purpose of the experiment was vocabulary learning and that they would be tested on vocabulary" (p. 156). But in the end, it is the learner not the teacher who decides whether to attend to the input or not. The empirical studies on incidental vocabulary are presented below.

Tang and Treffers-Daller (2016) studied Chinese EFL students' incidental vocabulary through involvement load hypothesis (ILH). They provided evidence that incidental vocabulary uptake is limited and tried to improve it. In this regard, they employed ILH and found it to be effective in incidental vocabulary learning.

Wang (2013) aimed to investigate whether Taiwanese lower-level EFL learners could increase their word knowledge through extensive reading. The tool for assessing vocabulary knowledge in this study was the table proposed by Paribakht and Wesche (1999). He found that learners achieved significant vocabulary gains after this program, suggesting that extensive reading affects lower-level learners' incidental vocabulary learning.

To the best of researchers' knowledge, there was no study investigating incidental vocabulary learning within the framework of DA. It is noteworthy that the concept of incidental learning has been particularly important in the context of

research on vocabulary acquisition, which is the core of ESP (Nito, 2004) and therefore, both notions will be discussed jointly incorporating DA as well.

### 2.3. ESP revisited

ESP has increasingly grown since its emergence, believed to be in the early 1960s (Swale, 1988). Because of the advances in technology and many other factors such as globalization, overseas scholarship for higher education, economic and political ties among nations, tourism, and international academic gatherings, ESP has become the focus of attention all over the world.

In 1997, the first conference on ESP was held in Japan. In this conference, all participants were asked to give a definition of ESP. Some of them defined ESP as teaching of any English material for any specified purpose. Others defined it more precisely as teaching English in academic lessons or for professional and vocational purposes (Anthony, 1998). Dudley-Evans and St. John (1998), however, offered a characteristic definition of ESP, focusing on absolute and variable characteristics as follows. Regarding the absolute features, an ESP course subscribes to the following essential principles:

1. ESP is defined to meet specific needs of the learners

2. ESP makes use of underlying methodology and activities of the discipline it serves

3. ESP is centered on the language appropriate to these activities in terms of grammar, lexis, register, study skills, discourse and genre.

And the following constitute some of the variable features of an ESP course:

1. ESP may be related to or designed for specific disciplines
2. ESP may use, in specific teaching situations, a different methodology from that of General English
3. ESP is likely to be designed for adult learners, either at a tertiary level institution or in a professional work situation. It could, however, be for learners at secondary school level
4. ESP is generally designed for intermediate or advanced students
5. Most ESP courses assume some basic knowledge of the language systems (p. 4).

Studies conducted on ESP have focused on one dimension of ESP, either the learning process or teachers' and students' attitudes toward ESP. These studies also showed that ESP teachers' lack of sufficient knowledge and training in this area has affected the whole process and has caused it to be in a poor condition. Again as for the incidental vocabulary learning, the researchers could find no international or national studies focusing on the role of DA in incidental vocabulary learning in an ESP course setting. Accordingly, the present study was launched to fill this gap in the related literature.

## 3. Methodology

### 3.1. Participants

The participants of the present study included an intact group of 25 undergraduate electronic engineering students pursuing their studies during the academic year of 2020 – 2021. They all had passed their ESP course. They were both male and female students and all were native speakers of Persian. However, it should be noted that gender wasn't considered as a moderating variable in this study. Their age range was between 19 and 24 years old and they were all informed of the voluntary nature of the study. They were assured that the personal information they provide at the starting page of the website would remain confidential; they were then provided with sufficient information about the procedure and purpose of the study.

### 3.2. Research tools

The instruments and materials used in this study include The English Preliminary Test, a researcher-developed vocabulary knowledge scale (VKS) based on the model of Paribakht and Wesche (1999), a researcher-developed multiple-choice type test of vocabulary. These instruments are elaborated below.

#### 3.2.1. The proficiency test

The PET was used to ensure the homogeneity of the proficiency level of students. This test is composed of two parts: Reading and Writing. The reading part includes five sections with 35 multiple-choice items providing simple written information. The writing section, on the other hand, consists of three parts with 8 items for which students are asked to do sentence completion task, provide specific information, and write a letter with the word limit of 100 words.

#### 3.2.2. Vocabulary knowledge scale

The textbook used in this university for the ESP course was written by Haghani (2013) and published by SAMT organization. The book is composed of twenty short passages on different issues in electronics, followed by some exercises. It is worth mentioning that the first draft of the vocabularies chosen to be included in the VKS contained some compound nouns, but they were excluded for the purpose of the homogeneity of elicitation test. Accordingly, another list without any compound structure was prepared. Finally, thirty out of the forty-three incidental vocabularies were included in the final VKS.

The VKS was based on Paribakht and Wesche's (1999) model and was run to ensure that they were unfamiliar to the participants. Based on this scale, the participants were required to indicate whether they had seen or known each of the thirty target words. The vocabularies, which students didn't know or couldn't recall the meaning, were selected to be included in the mediation-based incidental vocabulary test (MBIVT). Accordingly, thirteen out of the thirty vocabularies were kept for assessment purposes and the others were excluded. Paribakht and Wesche's (1999) model solves the validity concerns since the items to be assessed were based on a needs-analysis basis. Therefore, there was no need to conduct a pilot study for validation indexes.

#### 3.2.3. MBIVT

The development of CDA was inspired by the concept of graduated prompt approach proposed in the research of Brown and Ferrara (1985) and Campione and Brown (1990). In this approach, mediation is provided during a single test administration in the form of prompts arranged from the most implicit to the most explicit. In this form, prompts are offered for each item whenever the learners need. One of the advantages of this approach is that it provides information on the number of prompts required by each learner (Poehner et al., 2015).

The design of the CDA test was based on the cake format of interventionist approach, i.e. test-within-test design. Accordingly, a standardized menu of hints is prepared in this approach and is provided for each item. If the examinee gives a wrong answer in this process, the first mediation prompt appears and the item is re-attempted; otherwise, he or she will be directed to the next item.

The MBIVT was administered to a pilot group of 25 students from the University of Zanjan, who were at the same proficiency level as the participants of the study. Accordingly, some modifications were made to the test items based on the results of the test piloting and comments of the participants. The KR-21 method, as an estimation of reliability, was used to examine the internal consistency of the test. To this end, the calculator at www.cedu.niu.edu/~walker/calculators/kr.asp was used (Poehner et al., 2015).

### 3.3. Student Survey

The qualitative phase of this study, addressed by the second research question, dealt with the participants' evaluation of this CDA-based approach to incidental learning of the target vocabulary items within the ESP course setting. To this end, the survey used in Nirmalakhandan (2007) weighing students' evaluation of CDA was adopted in the present study. This survey consisted of eight items that were rated based on the five-point Likert scale (1 = strongly disagree, 5 = strongly agree). It is of great importance to note that the reliability of this survey was calculated to be .073, which is acceptable. The figure was not included in the original paper and after mailing the author, he assured the reliability of the student survey.

### 3.4. Procedure

#### 3.4.1. Test construction procedure

The first step in preparing the intended test was the selection of passages. First, the professors of electrical engineering department, who had the experience of teaching this course book, were consulted to see which parts of the book were covered during the term. Accordingly, eight passages were randomly selected from this book. Having selected the passages, it was the time to set out to prepare items and their response options based on the passages.

The response options for each item were sentences extracted either from the exercises of the aforementioned book or from the final exams prepared for the electrical students of Payame Noor University. It is noteworthy that five response options

were prepared per item; four response options as the conventional multiple-choice format and one as a distractor in order to provide examinees more freedom and to allow them to make multiple re-attempts (Poehner & Lantolf, 2013).

The next step, after preparing passages, items and response options, was to prepare four hints per item. If the number of response options and hints was equal for each item, the participants would find the answer in the second attempt. In other words, if the test was in a conventional four-option multiple-choice format, one response option would be deleted offering a hint and, thus, the participants would have two alternatives in the second hint. To avoid this problem, the number of the response options was five and the number of the given hints was four per item. The first and the last hints remained fixed for each item, arranged from the most explicit to the most implicit. The format of the first and the second prompts per item was as follows:

Hint 1: That is not the right answer.

Hint 4: The correct answer was.

When the test finished, it was the time to make it a CDA test. Therefore, it was prepared in the form of an online test.

### 3.4.2. Website preparation

The test starts with the first passage. After reading the passage, the first item appears on the next page. If the examinee answers incorrectly, the first prompt, i.e. the most implicit one, is offered and the examinee is given the chance to re-attempt the item. If the second response of the examinee is again incorrect, the second prompt, i.e. more explicit one, is provided. The procedure continues until either the examinee gives the correct answer and is moved to the next item or until the last hint and the right answer, i.e. the most explicit one, is offered.

During the pre-test stage, five professors of the electrical engineering department were asked to comment on the test for the purpose of ensuring the content validity of the MBIVT. This phase was considered as the "expert judgment" for the current test. As a result, the experts confirmed the content validity of the test and then, a pilot group of 25 students, with the same proficiency level as the participants, were asked to take the test. Taking into account the feedback of professors and the pilot group, the researchers modified some problematic response options and prompts.

As was mentioned earlier, the KR-21 method was used as an indicator of reliability estimation. The reliability of dynamic and non-dynamic tests was calculated based on the scores of the pilot group, which turned out to be 0.79 and 0.92, respectively. The reliability of the reading and writing section of the PET is 0.88, the listening section is 0.77, and the speaking section is 0.84. The overall reliability of the test is 0.92. Reliability and validity of dynamic and traditional tests have been calculated in the pre-test stage. The results can be seen in Table 1.

**Table 1.** Validity and reliability of the MBIVT

| | Traditional | Dynamic |
|---|---|---|
| Reliability | 0.921 | 0.796 |
| Validity | | 0.938 |

After the completion of the test procedure, the participants' evaluation of DA procedures was qualitatively reported. To this end, the student survey mentioned earlier was sent to the participants' email and they were asked to rate the degree to which they agreed or disagreed with the statements. Along with this survey, students were given the freedom to comment on the CDA and its procedure in order to have a more comprehensive image of their evaluation of CDA.

### 3.4.3. Scoring procedure

Unfortunately, automatic score file creation was impossible due to the problems and difficulties in designing the website. Consequently, the authors themselves calculated three actual, mediated, and learner potential scores. As it was suggested in the literature (e.g. Poehner et al., 2015), the actual or non-dynamic score is the score to be calculated exactly the same as traditional assessments, i.e. the first response without any mediation. Accordingly, the score is either 0 or 4 in a way that 0 is for incorrect responses and 4 is calculated for correct responses. The mediated or dynamic score ranges from 0 to 4 based on the number of prompts offered to the examinee so that by offering a prompt, one point is reduced. Finally, LPS was calculated based on the formula proposed by Kozulin and Garb (2002) to distinguish the independent and mediated performance of examinees, which captures learners' ZPD.

# 4. Results

## 4.1. Results for the first question

As a prerequisite to the parametric analysis, the normality of data distribution was checked through the Kolmogorov-Smirnov test because of the potential effect that outliers and extremes could have on the results of the study. The results confirmed that the normality assumption was met ($p > 0.05$). Also, paired-sample t-test was meant to answer the first question of the study to see whether there was a significant difference between the actual and mediated scores of the participants. The results of this test revealed a statistically significant difference between the mean actual score ($M = 28.96$, $SD = 10.47$) and mean mediated score [($M = 40.36$, $SD = 6.794$), $t(24) = -11.125$, $p = .000$]. Therefore, the null hypothesis that "CDA has no effect on incidental vocabulary learning by Iranian EFL learners in an ESP teaching/learning context" is rejected and CDA is proved to be effective in improving incidental vocabulary learning of ESP learners. The results appear in Table 2 below.

**Table 2.** Distribution of CF by type

| | | Paired Differences | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval of the Difference | | | | |
| | | Mean | Std. Deviation | Std. Error Mean | Lower | Upper | T | Df | Sig. (2-tailed) |
| Pair 1 | Actual score - Mediated score | -11.400 | 5.123 | 1.025 | -13.515 | -9.285 | -11.125 | 24 | 0.000 |

In addition, the effect size turned out to be quite large (Eta squared = .83). Thus, 83 percent of the shared variance could be explained by the CDA approach. The results of the preceding statistical analyses suggested that mediation helped all examinees without regard to their independent performance.

The scoring profile of four learners was selected to clarify the promotional role of mediational steps and learners' potential for learning. First, two learners with the same actual score, relatively low scores, were selected to show different learning potentials and then, two relatively high scores were selected for further comparison. The scores are presented in Table 3 for further comparison. The classification of LPS according to Kuzolin and Garb (2002) is such that an LPS of greater than or equal to 1 is considered high, an LPS between 0.71 - 0.8 is considered mid, and an LPS of lower than 0.71 is considered low. Therefore, LPS can give an in-depth understanding of learning ability.

**Table 3.** Actual, Mediated, Gain and Learning Potential Scores for a Number of Learners

| | Actual score | Mediated score | Gain score | LPS |
|---|---|---|---|---|
| Learner 1 | 16 | 24 | 8 | 0.61 |
| Learner 2 | 16 | 36 | 20 | 1.07 |
| Learner 3 | 44 | 49 | 5 | 1.03 |
| Learner 4 | 36 | 45 | 9 | 1.03 |

The first obvious point that can be seen in Table 3 above is that all learners benefited from mediational prompts. Accordingly, the first two learners had the same actual scores, showing that they were at the same level of independent performance, with different mediated scores and LPSs. This suggests that learners can benefit differently from the mediation procedure and they have different learning abilities. In other words, learner 1 has an actual level of 16, which is the same as learner 2. Going through the mediational steps, learner 1 was promoted by 8 scores and got a mediated score of 24, which was

much less than that of learner 2, who was promoted by 20 scores and got 36. Regarding their LPSs, learner 1, with a score of 0.61, was in a lower range and learner 2, with a score of 1.07, was in a higher range of LPS, suggesting their different ZPDs or learning capabilities. In other words, despite the same actual scores, they needed different instructional support.

The next point to be made in this regard is the case of learners 3 and 4. These two learners got different actual and mediated scores, but had the same LPSs. This indicates that unlike their different actual and mediated scores, the two learners' potential for learning or further progress was the same. Their high LPSs suggest that learners who have high independent performance can also benefit more than those who perform poorly.

### 4.2. Results for the second research question

The second research question was meant to investigate the participants' evaluation of the CDA program. To this end, the survey questionnaire adopted from Nirmalakhandan (2007) was adopted and mailed to the participants. They were also asked to comment on the CDA program. The quantified results of the survey were analyzed through frequency analysis. As mentioned earlier, anchor 1 in this survey indicated strongly disagree and anchor 5 indicated strongly agree. The results of frequency analysis are presented in Table 4 below.

**Table 4.** The Emerging Percentages for the Survey Items

| Items | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
|---|---|---|---|---|---|
| Item 1 | 0 (0%) | 2 (8%) | 3 (12%) | 10 (40%) | 10 (40%) |
| Item 2 | 0 (0%) | 1 (4%) | 4 (16%) | 12 (48%) | 8 (32%) |
| Item 3 | 0 (0%) | 1 (4%) | 4 (16%) | 11 (44%) | 9 (36%) |
| Item 4 | 5 (20%) | 7 (28%) | 7 (28%) | 6 (24%) | 0 (0%) |
| Item 5 | 8 (32%) | 8 (32%) | 2 (8%) | 7 (28%) | 0 (0%) |
| Item 6 | 0 (0%) | 0 (0%) | 5 (20%) | 15 (60%) | 5 (20%) |
| Item 7 | 0 (0%) | 5 (20%) | 3 (12%) | 12 (48%) | 5 (20%) |
| Item 8 | 0 (0%) | 0 (0%) | 8 (32%) | 7 (28%) | 10 (40%) |

Accordingly, from the eight items constituting the student survey, items 4 and 5 were somehow against the use of CDA. The frequency of responses for items other than these two items shows that respondents mostly favored the CDA and its use. That is to say, 40% strongly agreed and 40% agreed with the first item. Regarding item 2, 32% strongly agreed and 48% agreed. With respect to item 3, 36% strongly agreed and 44% agreed. For item 6, 20% strongly agreed and 60% agreed. Regarding items 7 and 8, 20% and 40% strongly agreed and 48% and 28% agreed, respectively.

With respect to the two items that were against CDA, 20%, 28%, and 28% strongly disagreed, disagreed, and were neutral, respectively. For the second item, 32% strongly disagreed, 32% disagreed, and 8% were neutral. The results suggest that most of the participants strongly agreed or agreed with the items favoring CDA and most of them strongly disagreed or agreed with the items against CDA. Also, most of the participants commented positively on the CDA program and only a few of them mentioned some negative points in this regard. Taken together, students' evaluation of CDA was positive.

## 5. Discussion

The present study aimed to investigate whether CDA was effective in assisting learners' incidental vocabulary learning in the context of ESP. Regarding the first research question, the results of the paired-sample t-test helped reject the null hypothesis

that CDA has no effect on incidental vocabulary learning of Iranian EFL learners in an ESP teaching/learning context. In other words, the CDA program proved to be significantly effective in improving the participants' incidental vocabulary learning within the ESP course context. This finding is in line with the findings of Saeidi and Hosseinpour (2013) who showed that learners' vocabulary learning rate can be enhanced through DA. Overall, the results of the present study are in line with the main body of the literature (e.g. MovahedFar et al., 2022; Estaji and Ameri, 2020; Poehner & Lantolf, 2021; Khomijani et al., 2019; Zangoei et al., 2019; Ajideh & Nourdad, 2013; Hessamy & Ghaderi, 2014; Mardani & Tavakoli, 2011; Orikasa, 2010; Pishghadam et al., 2011; Poehner et al., 2015), demonstrating the positive role of DA and CDA in improving L2 learning skills and sub-skills.

The results also indicate that the LPS can be used to predict learners' abilities and the amount of help received. Literature suggests that DA and CDA are powerful learning tools to enhance learning, and LPS can be used to distinguish learners who have similar actual scores and gain useful information about learners' ability to learn.

The second research question aimed to investigate the participants' evaluation of the CDA program. The results of the survey along with the comments of the participants on the CDA program and its use showed that their evaluation was in favor of the program and its use as an assessment tool, integrating instruction and assessment. This finding of the study was in line with the findings of Nirmalakhandan (2007) and Hidri (2014), who investigated the benefits of the use of CDA in improving students' achievement and learning.

## 6. Conclusion

To sum it up, DA, being underpinned by Vygotsky's sociocultural theory, is not an assessment tool per se, rather it focuses on promoting learning through determining learning potentials. L2 CDA is an inchoate concept and does not have a long history. The present study endeavored to apply interventionist CDA on ESP incidental vocabulary learning, which is scarce in the literature. Accordingly, both qualitative and quantitative procedures were applied through a mixed method triangulation research design to find possible answers to the research questions.

The findings of this study suggested that CDA is effective in improving incidental vocabulary learning among students of electrical engineering and also it can encourage lackadaisical students to engage more in the learning process. Furthermore, CDA can help overcome some difficulties of DA so that it is less time- and labor-consuming and can be applied in group size.

In addition, a majority of ESP students are not satisfied with their courses and their outcomes and believe that these courses do not meet their academic and occupational needs and, thus, ask for an urgent shift and transformation in ESP teaching and testing (e.g. Alibakhshi et al., 2011; Moslemi et al., 2011). Therefore, the use of CDA for vocabulary learning, as the core of ESP, can serve as a possible solution to the current problems faced by the students. CDA is specifically efficient because of its view of instruction and assessment as a uniform activity, which can respond to ESP students' needs in both teaching and assessment areas.

## 7. References

Ajideh, P., & Nourdad, N. (2013). Dynamic assessment revealing individual differences in EFL reading comprehension ability. *International Journal of Social Sciences and Educations*, 3(2), 340-350.

Alibakhshi, G., Ghand Ali, H., & Padiz, D. (2011). Teaching and Testing ESP at Iranian Universities: A Critical View. *Journal of Language Teaching and Research*, *2*(6), 1346-1352.

Anthony, L. (1998). Defining English for specific purposes and the role of the ESP practitioner. *Center for Language Research 1997 Annual Review*, 115-120.

Bogdanov, S. (2013). One-time treatment for incidental vocabulary learning: Call for discontinuation. *Teaching English with Technology, 13*(2), 75-82.

Brown, A. L., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 273–305). New York: Cambridge University Press.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136-163.

Budoff, M. (1968). Learning potential as a supplementary testing procedure. In J. Hellmuth (ed.) *Learning Disorders. Vol. 3*. Seattle, WA: Special Child.

Campione, J. C., & Brown, A. L. (1990). Guided learning and transfer: Implications for approaches to assessment. In N. Frederikson, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 141–172). Hillsdale, NJ: Erlbaum.

Dudley-Evans, T., & St. John, M. J. (1998). *Developments in English for Specific Purposes: A Multi-disciplinary Approach*. Cambridge: Cambridge University Press.

Ebadi, S., & Saeedian, A. (2015). The Effects of Computerized Dynamic Assessment on Promoting At-Risk Advanced Iranian EFL Students' Reading Skills. *Issues in Language Teaching*, *4*(2), 26-1. DOI: 10.22054/ilt.2015.7224

Estaji, M. & Ameri. A.H. (2020). Dynamic assessment and its impact on pre-intermediate and high-intermediate EFL learners' grammar achievement. *Cogent Education*, 7:1, 1740040, DOI: 10.1080/2331186X.2020.1740040

Haghani, M. (2013). *English for the students of Power, Electronics, Control & Communications*. Tehran: SAMT.

Hessamy, GH. & Ghaderi, E. (2014). The role of dynamic assessment in vocabulary learning of Iranian EFL learners. *Procedia - Social and Behavioral Sciences,* 98, 645-652.

Hidri, S. (2014). Developing and evaluating a dynamic assessment of listening comprehension in an EFL context. *Language Testing in Asia*, 4(4).

Hill, M., & Laufer, B. (2003). Type task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *IRAL*, 41, 87-106.

Hutchison, T., & Waters, A. (1987). *English for Specific Purposes: A Learning-Centered Approach.* Cambridge: Cambridge University Press.

Khomijani Farahani, A. & Mirzaei, Gh.R. (2019). Investigating the Effect of Dynamic Assessment on Learning Congruent and Non-Congruent Collocations by Iranian EFL Learners. *Journal of Foreign Language Research*, 9(3), 789-814.

Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International*, *23*, 112-127.

Lantolf J. P., & Poehner M. E. (Eds.). (2008). *Sociocultural theory and the teaching of second languages*. London: Equinox.

Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, *1*(2), 49-72.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural Theory and the Genesis of Second Language Development.* Oxford: Oxford University Press.

Laufer, B. (2006). Comparing Focus on Form and Focus on Forms in Second-Language Vocabulary Learning. *Canadian Modern Language Review*, 63(1):149.

Lidz, C. S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In Kozulin, A., Gindis, B., Ageyev, V. S. & Miller, S. M. (Eds.), *Vygotsky's educational theory in cultural context* (pp. 99–118). Cambridge: Cambridge University Press.

Mardani, M., & Tavakoli, M. (2011). Beyond reading comprehension: The effect of adding a dynamic assessment component on EFL reading comprehension. *Journal of Language Teaching and Research,* 2(3), 688-696.

Moslemi, F., Moinzadeh, A., & Dabaghi, A. (2011). ESP needs analysis of Iranian MA students: A case study of the University of Isfahan. *Canadian Center of Science and Education,* 4(4), 121-129.

MovahedFar, S., Abbasian, G., & Ameri, A. (2022). Computer-Based Dynamic Assessment of EFL Learners' Writing Performance: Evidence from Both Cognitive and Emotive Domains. *Journal of Language and Translation*, *12*(3), 1-19. DOI: 10.30495/ttlt.2022.692048

Nirmalakhandan N. (2007). Computerized adaptive tutorials to improve and assess problem-solving skills. *Computers & Education*, 49 (4), 1321-1329

Nito, R. (2004). Vocabulary & ESP: Teaching Business Vocabulary. Retrieved 25th October 2011, from http://www.scribd.com/doc/46984761/ESP

Orikasa, M. (2010). Interactionist dynamic assessment in L2 learning: A case study of tutoring L2 English oral communication. Retrieved from http://scholarspace.manoa.hawaii.edu/handle/10125/20258

Paribakht, T. & Wesche, M. (1999). Incidental L2 vocabulary acquisition: Theory, current research and instructional implications. *Thematic Issue: Studies in Second Language Acquisition*, 21(2).

Pishghadam, R., Barabadi, E., & Mehri Kamrood, A. (2011). The differing effect of computerized dynamic assessment of L2 reading comprehension on high and low achievers. *Journal of Language Teaching and Research*, *2(6)*, 1353-1358.

Poehner, M.E., & Lantolf J.P. (2021). The ZPD, Second Language Learning, and the Transposition ~ Transformation Dialectic. *Cultural-Historical Psychology*, 17(3), 31—41. DOI: https://doi.org/10.17759/chp.2021170306

Poehner, M. E. & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized Dynamic Assessment. *Language Teaching Research 17*(3): 323-342.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development.* New York: Springer.

Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (CDA): Diagnosing L2 development according to learner responsiveness to mediation. *Journal of Language Testing.* 32(3), 337-357. [Downloaded from ltj.sagepub.com at Hacettepe University on June 28, 2015]

Qingha, L., & Di, L. (2015). A review of empirical studies in L2 dynamic assessment. *Chinese Journal of Applied Linguistics*, 38(1), 55-73.

Saeidi, M., & Hosseinpour, A. (2013). The effect of dynamic assessment as an instructional tool on Iranian EFL learners' vocabulary learning. *Journal of Basic and Applied Scientific Research.* 3(10), 421-429.

Sarani, A., & Sahebi, L. (2012). The impact of task-based approach on vocabulary learning in ESP courses. *Canadian Center of Science and Education*, 5(10), 118-128.

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.

Summers, R. (2008). *Dynamic assessment: Towards a model of dialogic engagement*. (Unpublished doctoral dissertation), University of South Florida, Tampa.

Swales, J. (1988). *Episodes in ESP*. Prentice Hall.

Tang, C. & Treffers-Daller, J. (2016). *Assessing incidental vocabulary learning by Chinese EFL learners: a test of the Involvement Load Hypothesis*. In: Yu, G. & Jin, Y. (eds.). Assessing Chinese learners of English: language constructs, consequences and conundrums. Palgrave, London, pp. 121149. ISBN 9781137449771. DOI: https://doi.org/10.1057/9781137449788. Available at http://centaur.reading.ac.uk/39023/

Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *Regional Language Centre Journal*, 45(1), 51-65.

Tzuriel, D. (2001). *Dynamic assessment of young children*. New York: Plenum Publishers.

Vafaee, P. (2011). Dynamic assessment: A dialectical integration of assessment and instruction. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 11(1), 63-66.

Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Cambridge, MA: Harvard University Press.

Wang, Y. H. (2013). Incidental vocabulary learning through extensive reading: A case of lower level EFL Taiwanese learners. *The Journal of ASLA TEFL*, 10(3), 59-80.

Xi, J., & Lantolf, J. P. (2021). Scaffolding and the zone of proximal development: A problematic relationship. *Journal for the Theory of Social Behaviour, 51*(1), 25–48. https://doi.org/10.1111/jtsb.12260

Zangoei, A., Zareian, G., Adel, S.A.M., & Amirian, S.A.M. (2019). The impact of computerized dynamic assessment on Iranian EFL learners' interlanguage pragmatic development. *Journal of Modern Research in English Language Studies*, 6(4), 139-165. DOI: 10.30479/jmrels.2019.11536.1433

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 95-112*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2023.6241.1011

# An Exploration into EFL Learners' Vocabulary Learning in Flipped Classrooms

Nouroddin Yousofi[1]
Saba Bashiri[2]

[1]Associate Professor, Department of English Language and Literature, Faculty of Persian Literature and Humanities, Razi University, Kermanshah, Iran
[2]Ph.D. Candidate, Department of English Language and Literature, Faculty of Persian Language and Humanities, Razi University, Kermanshah, Iran

**ARTICLE HISTORY**

Received: 27 August 2022
Revised: 27 October 2022
Accepted: 24 January 2023
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: nyousofi@yahoo.com

**ABSTRACT**

Utilizing a sequential exploratory mixed-methods design, this study investigated the effects of flipped classrooms on Iranian EFL learners' vocabulary learning. Thirty Iranian EFL learners, who were enrolled in an upper-intermediate class at a private language school in Iran, were randomly divided into two groups: the experimental group (the flipped learning group) and the control group (the traditional learning group). Pre-tests and immediate post-tests were administered via the DIALANG online diagnostic test which was also utilized to assess the learners' current vocabulary level to explore the effects of flipped classrooms on the learners' vocabulary learning. Semi-structured interviews along with students' weekly journals were used to collect the qualitative data. The results of paired- and independent-sample t-tests showed that the experimental group performed better in the post-test both compared to their pre-test and the control group's post-test. Thematic analysis of the qualitative data indicated that the flipped classroom enhanced the experimental groups' knowledge of vocabulary, class participation, interaction, and engagement, although some of the participants were not accustomed to such an instructional method and preferred being instructed by the teacher and do the activities individually in the classroom. The study implications will be discussed.

**KEYWORDS:** flipped learning; EFL vocabulary learning; EFL learners

## 1. Introduction

One of the most important aims of language teaching is to help learners to improve different skills and language components. The field is being penetrated by the different technologies which are increasing and advancing rapidly (Chakowa, 2018). In line with the technological advancements, in both mainstream and language education, stakeholders have been propelled to take advantage of either e-learning or flipped classrooms. Flipped classrooms combine face-to-face and distance education to render what is called blended learning (Santikarn & Wichadee, 2018) and shift the learning atmosphere into a dynamic interactive environment (Bergmann & Sams, 2012) by requiring the learners to do pre-and post-class tasks in order to take advantage of in-class time (Tawfik & Lilly, 2015). An important point regarding flipped classrooms is that both learning types, i.e. face-to-face in-class and online, are integrated and none of them is prioritized over or marginalized for the sake of others (Garrison & Kanuka, 2004). Information technologies intervene in the flipped classrooms to complete knowledge teaching after the face-to-face class time (Wei, 2019), but knowledge internalization takes place with the aid of the teacher and the classmates during the class (Jinlei et al.,2012). Positive results of flipped learning in both L2 and mainstream education are

**Cite this article:** Yousofi, N., & Bashiri, S. (2023). An exploration into EFL learners' vocabulary learning in flipped classrooms. *Applied Linguistics Inquiry*, *1*(1), 95-112. doi: 10.22077/ali.2023.6241.1011

reported in the literature, including more individualized learning and more learners' satisfaction (Cho & Lee, 2016) significantly increased academic achievement and better performance (Adnan, 2017; Bang, 2017; Pyo, 2017; Wu et al., 2019) more academic self-efficacy (Pyo, 2018) higher teaching efficiency (Knežević et al., 2020), more engagement, learning motivation, and interaction (Ahmed et al., 2022; Challob, 2021; Chuang, 2018; Zainuddin & Halili, 2016). Positive students' perceptions of flipped classrooms are also reported in several studies (Hung, 2015; Lee, 2021; Nugroho & Fitriati, 2021; Teng, 2017; Webb et al., 2014; Yang, 2017).

Previously ignored compared to grammar and structure, vocabulary is now an important aspect of English language teaching (ELT) (Kim, 2018). As one of the challenging and key language components to be mastered by L2 learners (Alghamdi, 2019; Lin & Lin, 2019) vocabulary learning is the foundation for mastering an additional language (Schmitt, 2010). According to Burston "Without grammar very little can be conveyed, without vocabulary nothing at all can be conveyed." (as cited in Katemba, 2019, p.2). L2 vocabulary acquisition is facilitated by the numberless language resources and apps introduced by language technologies (Ma, 2017). The literature on technology-aided language learning is replete with studies the main target of which is vocabulary development (Stockwell, 2007). For instance, L2 vocabulary experiences can be facilitated by virtual environments and virtual reality technologies (Tai et al., 2020; Tseng et al., 2020), media and multimedia technologies (Alhamami, 2016; Arndt & Woore, 2018; Montero Perez, 2020; Peters & Webb, 2018; Ramezanali & Faez, 2019; Shadiev et al., 2020; Teng, 2020; Wang, 2019; Wong & Samudra, 2019), and the social media can also intervene to aid both learners and instructors in the process of vocabulary learning (Amer, 2014; Chen et al., 2018; Dehghan et al., 2017; Khansarian-Dehkordi & Ameri-Golestan, 2016; Ko, 2019; Ma & Yodkamlue, 2019; Terantino, 2016).

More specifically, investigating the role of flipped classrooms on vocabulary learning is also burgeoning across different language learning courses. According to Knežević et al. (2020), English academic vocabulary can be instructed more efficiently compared to the conventional approach. In Turkey, Özkal (2019) found that pre-intermediate EFL learners benefitted more from learning vocabularies through flipped classrooms. To the researchers' best knowledge, few studies were conducted on the effectiveness of flipped vocabulary learning among Iranian EFL learners (Fahandezh & Mohammadi 2021; Jalili, et al., 2020; Rezaei Fard et al., 2021). As there seems to be little information about the effects of flipped classrooms on upper-intermediate Iranian EFL learners' vocabulary learning, this study aimed at exploring the issue. Vocabulary was selected as the target component to be explored as the participants were on the verge of entering the advanced classes in which they would need to produce and comprehend a good number of vocabulary and lexical items in order to gain mastery of other language skills; furthermore, they were eager to attend a vocabulary course before attending advanced classes. The aspects of vocabulary knowledge being measured to trace the learners' improvement were word combination, word formation, semantic relations, and meaning. These criteria were selected based on the DIALNG diagnostic online test, taken by the participants, as they seemed to be troublesome for them. The following research questions guided this study:

1. To what extent do flipped classrooms affect the EFL learners' vocabulary learning?

2. What are the EFL learners' perspectives and perceptions about flipped vocabulary learning in terms of its perceived benefits and associated challenges?

## 2. Review of literature

The advent of technologies across teaching practices, along with all spheres of modern life, led the practitioners to embrace new instructional methods one of which was flipped learning introduced by Bergmann and Sams (2012), two chemistry teachers in the USA. Despite being a novel instructional approach, FL is not an original term (Berrett, 2012) and different terms have been used to refer to it, including inverted classroom (Lage & Platt, 2000) just-in-time teaching (Novak, 2011) and inverted learning (Davis, 2013). Bergmann and Sams (2012) stated that FL is a class in which "which is traditionally done in class is now done at home, and that which is traditionally done as homework is now completed in class" (p. 13). Therefore, flipped classrooms, as opposed to traditional classrooms, have two phases: the pre-class learning phase and the in-class student-centered phase (Network, 2014).

The most prevalent learning theory underlying FL is Vygotsky's (1978) sociocultural theory (SCT) of mind which states that all human learning is social, not individual. The main focus of this theory is on engaging students in practical activities (Lantolf & Poehner, 2014) during which interaction occurs and leads to human mental and behavioral development and functioning (Lantolf & Thorne, 2006). It has been argued that this development occurs within the person's zone of proximal development (ZPD) via mediation, which are two main terms in SCT (Lantolf et al., 2015). According to Vygotsky (1978, 1998), mediation helps the still-developing abilities to come to the surface. One of the mediating tools in flipped classrooms is the pre-class content shared by the teacher which then is discussed in pairs/groups during the in-class phase and leads scaffolding to happen (Loucky &Ware, 2016). Scaffolding happens in a social situation in which a more knowledgeable person (in this case, peers or teachers) helps the beginner to develop his/her ZPD using language and available tools (Donato, 1994; Turuk, 2008). This scaffolded functioning later becomes internalized by the learners and he/she can function independently and reach autonomy. Autonomy is defined as a learner's ability in goal-setting, self-generating opportunities for practice and learning, and self-evaluating their improvement (Cotterall, 1995).

The positive impacts of FL have been well substantiated in teaching different language skills and components. The effects of FL on Cambodian pre-university EFL learners' listening skills were explored by Roth and Suppasetseree (2016). The results showed the outperformance of the group who received the instructional material in the flipped classrooms and who also thought positively about the effectiveness of the method. In the same vein, Ahmad (2016)'s study on 34 third-year Egyptian EFL learners in a one-group pre-posttest study design showed the positive effects of flipped learning on their listening comprehension. Santikarn and Wichadee (2018) assessed the speaking and writing performance and perceptions of 40 advanced Thai EFL students who were instructed in an English course via flipped learning. Using multiple data collection instruments, they realized that the learners' performance improved significantly after the flipped course as revealed via their scores. In addition, they perceived the course to be satisfactory in making them autonomous learners. Yuliani et al. (2018) investigated the effectiveness of project-based flipped learning on 40 Indonesians' writing performance. They concluded that the flipped classroom was more successful in rendering competent writers, besides bringing about engagement and enjoyment in the flipped learners. Huang and Hong (2016) and Brown (2018) found that English learners enhanced their reading comprehension as a result of attending flipped reading classes and were satisfied with what happened to their reading abilities and autonomy. Moran (2014) opted to investigate not only 49 English Language Arts (ELA) students' emotional and behavioral engagement but also 143 teachers' pedagogical practices with flipped classrooms. The result was that the learners in the flipped classrooms increased their intrinsic and extrinsic motivation, and became more hardworking and organized compared to those who were in the traditional classrooms. Some of them reported that they experienced demotivation during the final flipped sessions, although they enjoyed them at the outset. More knowledgeable ELA students felt more comfortable with the nature of the flipped classes than those who were lesser successful, in other words, less competent ELA students felt frustration. The teachers asserted that their attitudes towards and their comfort with the method were dependent upon the class size, such that small classes size was more successfully instructed via FL.

During the history of foreign language teaching, different methods waxed and waned each of which treated each language skill and component, including vocabulary, differently. According to Carter & McCarthy (2014) and Schmitt (2010) until the 1980s, vocabulary was not paid considerable attention and after then L2 researchers got interested in vocabulary teaching (1997). Zhang et al. (2016) explored the adjustment and effects of vocabulary teaching strategies in a flipped classroom. Chinese EFL freshmen were divided into experimental and control groups in a mixed-methods study in order to compare traditional and flipped vocabulary learning methods. The flipped classroom performed better than its counterpart in the post-test. Furthermore, they had positive attitudes toward the classes in terms of better vocabulary learning, more class discussions to overcome problems, and more activity engagement. On the other hand, the control group deemed traditional vocabulary instruction time-consuming and not conducive to word memorization. Fifty-seven Korean EFL learners were the participants of a study by Kim (2018) which investigated the impact of FL on receptive and productive vocabulary learning. Based on the immediate post-test scores, the experimental group performed better than the control group. To explore the true impact of the treatment, delayed post-tests were also administered the results of which showed the endurance of the flipped classrooms' positive impact. Another intermediate EFL vocabulary class in Saudi Arabia was flipped by Alnuhayt (2018) in order to trace the impacts on the 45 students. In line with the last two studies, FL was conducive to vocabulary knowledge improvement and students' positive perceptions of them.

Jalili et al. (2020) focused on intermediate Iranian EFL learners' vocabulary development via flipped classrooms in a mixed-methods study. To the researchers' surprise, the traditional method of vocabulary instruction was superior the flipped model as the control group performed better on the post-test, and the rather obvious consequence was to the experimental groups' dissatisfaction with these classes. In a novel trend utilizing gamified flipped classrooms, Fahandezh and Mohammadi (2021) found that Iranian female EFL learners enhanced their vocabulary knowledge as a result of being provided the content in an interesting and creative context. To the best knowledge of the researchers, there are still issues to be explored regarding flipped vocabulary learning in the Iranian EFL context which should be fulfilled via conducting more empirical studies. This research aims to explore the same issue by targeting upper-intermediate EFL learners in a mixed-methods study design the first phase of which is experimental and the latter is an interview- and reflective-journals-based.

## 3. Method

### 3.1. Design

Quantitative and qualitative data were used in this study as we followed a sequential exploratory mixed-methods design. QUAN →QUAL mixed-methods type (Dornyei, 2007) was used based on which quantitative data is collected followed by the qualitative data. The quantitative data was gathered via pre-and post-tests and the qualitative data was collected through follow-up semi-structured interviews and students' weekly journals. We adopted this study design in order to elaborate more on the quantitative results through an in-depth analysis of the qualitative data (Dornyei, 2007; Riazi, 2016).

### 3.2. Participants

The sampling method used to select the study participants was the convenience sampling method (Dornyei, 2007). This study selected two intact upper-intermediate classes at a private language institute in Iran. There were 15 students in each class one of which received flipped vocabulary instruction and the other received traditional vocabulary instruction. They were all female learners whose ages ranged from 15 to 25 and their proficiency level was assured via the institute's standards and the DIALANG online diagnostic test. Based on the Common European Framework of Reference for the Languages (CEFR), both groups' vocabulary skills were at the B2 level. Pseudonyms are used in reporting the qualitative data results in order to protect the participants' privacy and assure their anonymity. Debriefing sessions were held before the treatment sessions during which the participants were provided with enough information about the study procedure and were given the freedom to withdraw without any penalties or need for compensation (Mackey & Gass, 2015). No participants in the study had prior experience of attending flipped classrooms and all of them were native Kurdish speakers, attending the private English school for almost five years, were at the same proficiency level, and all had access to the internet, computers, and personal email accounts.

## 3.3. Instrumentation

Research instruments in this study consisted of the DIALANG online diagnostic test, multimedia material for the treatment sessions, students' weekly journals, and semi-structured interviews. The DIALANG test diagnoses the language proficiency level in 14 European languages and delivers the results based on the Common European Framework of Reference for Languages (CEFR) from A1 (the least proficient) to C2 (the most proficient) in different language skills and components. The participants' proficiency level and their pre-and post-test scores were collected via this online test which is available at https://dialangweb.lancaster.ac.uk/.

The multimedia materials were prepared by one of the researchers who was also the course instructor of both groups. Vocabulary items from the book *504 absolutely essential words* (Bromberg et al., 6th edition) were selected to be taught in the treatment sessions, such that each lesson that included 12 words was covered during one single session. The lessons for the flipped classroom were instructed via PowerPoint slides prepared by the instructor. Each PowerPoint file contained the 12 vocabularies of a single session and each item included the following parts: part of speech, pronunciation (teachers' voice), a picture depicting the word, English definition, and contextualizing the word in sentences. (Appendix A). This group was also provided with a tutorial screen recording on how to search for online dictionaries and how to work with dictionary apps.

The students in the experimental group were asked to keep weekly journals and reflect on what aspects of vocabulary knowledge they had mastered more in this way and what affordances and challenges they faced during the flipped classes. Follow-up semi-structured interviews were conducted with eight students in the experimental group to tap into their perceptions towards attending flipped vocabulary learning classrooms. A total of four guiding questions were asked in Persian after being piloted by asking three non-participant students in the same institute in order to remove any ambiguities (Appendix B). All interview sessions were held individually in the institution and recorded, during which any popping-up points were clarified by both parties.

## 3.4. Procedures

This study was conducted for a total of 18 sessions during nine weeks (two 90-minute sessions per week). Right before starting the actual instructional sessions, all the students took the English vocabulary test of the DIALNG online test to determine their proficiency level (session one). The scores were recorded as the pre-test scores and no feedback was given to the students on their performance on the test. During the 15-session experimental period, the experimental group (flipped group, FG) received vocabulary instruction in flipped classrooms, while the control group (traditional group, TG) received instructions on the same vocabulary in traditional classrooms (sessions 2-16). The PowerPoint slides were shared with the FG one day before the face-to-face class via their emails. They were required to practice the pronunciation and part of speech and learn the word definition/meaning. Furthermore, they were asked to find synonyms and antonyms, find different collocations and combinations, and practice word-formation techniques to change the word into different parts of speech, e.g. via using affixes. These all had to be done in the slides to which specific spaces were allocated and then sent back to the teacher by the determined time. The same tasks formed the basis of class practices and discussions. During face-to-face classes, they formed three groups of five members (group members were subject to change each session) to answer the instructors' questions and take part in interactive activities to demonstrate what they have learned at home. After sharing their ideas on each lesson, each group had to write a short essay containing the lexical items covered in that session which was then shared with the teacher. All the participants' weekly journals were collected in the last treatment session.

The same vocabulary items were instructed to the TG following the conventional method which included: reading each word aloud in class and asking the students to repeat it, practicing its pronunciation and definition, thinking about synonyms, antonyms, different parts of speech, and collocation/combinations in a teacher-centered manner. This group was also asked to compose a short essay containing whatever they had learned in that session. The difference was that this task was done individually at home and was then submitted to the teacher the next session.

After finishing the treatment session, all the participants took an immediate post-test (three days after the last session using the DIALING online test [session 17]). Both pre-and post-tests were conducted in the institution based on a schedule and under specific time limitations. Finally, follow-up interviews were conducted individually in the 18th session. Each interview session lasted about 15-20 minutes and was audio-recorded for further analysis.

### 3.5. Data analysis

SPSS version 26 was used to conduct the quantitative data analysis. To compare the differences within and between groups, paired-sample t-tests, and independent-sample t-tests were calculated, respectively. The data obtained via participants' weekly journals and semi-structured interviews were analyzed via thematic analysis.

## 4. Results

### 4.1. Quantitative results

#### *4.1.1. Pre- and post-test differences within each group*

A paired-sample t-test was carried out to compare the pre-and post-test scores of each group. As Table 1 indicates, the TG's pre-and post-test scores were different from each other (sig<0.05), but as shown in Table 2 this difference was not so significant as the pre-test mean was (21.07) and the post-test mean was (22.60).

The FG's pre-and post-test were also significantly different from each other (sig<0.05) as shown in Table 3. But here the difference was larger and showed that the FG performed better in the post-test as the mean scores in Table 4 show.

**Table 1.** Paired samples t-test for the TG

| | | | | 95% Confidence Interval of the Difference | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Std. Error Mean | Lower | Upper | t | df | Sig. (2-tailed) |
| pre-test.T post-test.T | -1.533 | 1.506 | .389 | -2.367 | -.700 | -3.944 | 14 | .002 |

**Table 2.** Paired samples statistics for the TG

| | | Mean | N | SD | Std. Error Mean |
|---|---|---|---|---|---|
| Pair | pre-test.T | 21.07 | 15 | 3.369 | .870 |
| | post-test.T | 22.60 | 15 | 2.613 | .675 |

**Table3.** Paired samples t-test for the FG

| | | | | 95% Confidence Interval of the Difference | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Std. Error Mean | Lower | Upper | t | df | Sig. (2-tailed) |
| pre-test.F | -4.067 | 3.575 | .923 | -6.046 | -2.087 | -4.406 | 14 | .001 |

post-test. F

**Table 4.** Paired samples statistics for the FG

| | | Mean | N | SD | Std. Error Mean |
|---|---|---|---|---|---|
| Pair | pre-test.F | 22.73 | 15 | 3.595 | .928 |
| | post-test.F | 26.80 | 15 | 2.624 | .678 |

### *4.1.2. Pre- and post-test differences between groups*

We carried out two independent samples t-tests in order to find out whether there were any significant differences between the two groups in terms of their pre-and post-test scores. Table 5 shows that sig<0.05, so there was a difference between the groups' pre-tests but the mean scores (FG mean score= 22.73; TG mean score= 21.07) indicated that this difference was meager and insignificant (Table 6).

**Table 5.** Independent samples t-test for the pre-tests

| | | | | 95% Confidence Interval of the Difference | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Std. Error Mean | Lower | Upper | t | df | Sig. (2-tailed) |
| pre-test.F pre-test.T | 1.667 | 2.895 | .747 | .063 | 3.270 | 2.230 | 14 | .043 |

**Table 6.** Independent samples statistics for the pre-tests

| | | Mean | N | SD | Std. Error Mean |
|---|---|---|---|---|---|
| Pair | pre-test.F | 22.73 | 15 | 3.595 | .928 |
| | pre-test.T | 21.07 | 15 | 3.369 | .870 |

Although there was not a great difference between the FG and the TG in their pre-test scores, the independent samples t-test statistics of the post-tests revealed that there was a significant difference between their post-tests (sig<0.05, as shown in Table 7). Furthermore, Table 8 indicated that the FG performed significantly better (mean score= 26.80) than the TG (mean score= 22.60) in their post-test.

**Table 7.** Independent samples t-test for the post-tests

| | Mean | SD | Std. Error Mean | 95% Confidence Interval of the Difference | | t | Df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|---|
| | | | | Lower | Upper | | | |
| post-test.F post-test.T | 4.200 | 3.167 | .818 | 2.446 | 5.954 | 5.137 | 14 | .003 |

**Table 8.** Independent samples statistics for the post-tests

| | | Mean | N | SD | Std. Error Mean |
|---|---|---|---|---|---|
| Pair | post-test.F | 26.80 | 15 | 2.624 | .678 |
| | post-test.T | 22.60 | 15 | 2.613 | .675 |

## 4.2. Qualitative results

For answering the second research question, the data obtained via conducting semi-structured interviews and weekly journals, kept by the students, were analyzed following a thematic analysis procedure. Out of the transcribed interviews and written journals, a set of themes emerged. Table 9 depicts a detailed elaboration of the results. Students' quotations will be reported anonymously by using pseudonyms including F1, F2, F3, F4, F5, F6, F7, and F8.

**Table 9.** Themes, codes, and segments extracted from the qualitative data

| Themes | Codes | Segments |
|---|---|---|
| **Perceived benefits** | *-Pedagogical benefits* | -Different and more activities to do |
| | | -Feeling more self-responsibility and autonomy for learning |
| | | -Enjoying learning from discovering |
| | | - More preparation for in-class participation |
| | | -More engagement in the learning process |
| | | -Spending in-class time more efficiently |
| | | -Gaining more profound knowledge of words |
| | | -Better memorization of the lessons |
| | *-Social benefits* | -Not being forced to answer without preparation |
| | | -Not feel frustrated to take part in classroom activities |
| | | - Sharing knowledge with classmates fostered social interaction |

| | | |
|---|---|---|
| **Associated challenges** | *-Pedagogical challenges* | - Preferring teacher's explanations<br>-Not suitable for those who prefer individual activities<br>- Not being accustomed to such instruction<br>-Preferring to think within classrooms confines |
| | *-Material-related challenges* | -Problems with downloading the file<br>-Preferring video materials instead of audio<br>-Time-consuming |

Table 9 shows that the FG experienced both benefits and challenges during the vocabulary courses instructed via flipped classrooms. With regard to the perceived benefits, data analysis revealed two broad categories: pedagogical and social. According to F3 requiring the students to do more activities was a good instructional method:

*Indeed, we did activities in our previous courses, but the nature and amount of activities differed here, as we had to surf different resources to find what was required.*

Being responsible to solve the tasks at home helped learners to practice autonomy:

*I think the assignments to be done at home before class brought about several positive characteristics for us. For instance, as an EFL learner, I felt more responsibility for my learning and enjoyed learning without being dependent upon in-class instructions (F8).*

Some students thirst for discovering some aspects of what they are learning on their own, instead of being passive knowledge receivers. The material used for FG in this study provided these kinds of students with such an opportunity:

*I always enjoyed being familiar with different uses and combinations of words via searching different resources such as dictionaries or online materials. In traditional classrooms, we always had a shortage of time to cover all those aspects and I was not prompted to discover them, but this experience was a great opportunity to do what I always enjoyed in my learning practices (F1).*

As the students were required to read the material and do the tasks, they were more prepared to participate in face-to-face classes, hence more engagement in the learning process was assured. F4's comment is revealing in this regard:

*We had no choice except to receive the teacher's file and do what we were asked to …. At first, it was somehow cumbersome for me but as time went on I felt more prepared for taking part in class activities to which I was not so eager previously.*

As the actual class time is limited to cover all that is needed for better learning and consolidation of the instructional materials, FC compensated for this problem to a great extent as also pointed to by F2:

*The whole class time was devoted to doing the task by interacting with the teacher and our classmates. Previously, we lacked enough time to proceed based on our syllabus, but this instructional method compensated for this problem. Another good point was that we did not feel the passage of time.*

As the students were asked to find different parts of speech, word combinations/collocations, and contextualize all of those in a text, they asserted gaining a deeper knowledge of the lexical items which assisted their long-term memory as well. For instance, F6 said:

*According to a famous quotation from Franklin, when you are engaged in doing something, you will understand and learn it better… this was exactly what we experienced during the course, such that instead of rote memorization of the instructed words, we were able to manipulate a word in different ways meaningfully and for different purposes.*

Along with the aforementioned pedagogical benefits, some social benefits were also perceived by the FG. Doing assignments in front of several peers can be frustrating for some students, especially those who are shy and apprehend explicit performance and those who cannot perform without prior preparation. Both points were stated by F5:

*I always feel frustrated to take part in class interactions which causes my teachers to perceive me as an inactive person … In fact, I am somehow an introverted person who does not prefer to express public comments. … The preparation we were required to gain before classes, helped me overcome this characteristic, as I was sure that I have something to say*.

The social interaction invoked by practicing what was learned in the form of groups was very appealing for those students who enjoyed doing activities via social interaction. F1 was a case in point:

*I think group practicing was a strong point of this experience… because I was able to share what I knew with my peers and vice versa … I learned a lot from my classmates because we all had practiced and were somehow sure of the points raised and the solutions provided.*

Although these benefits and affordances experienced during the flipped classrooms were of great appeal to the participants, some negative points were also present as perceived by the same students which cannot be neglected. These challenges were either pedagogical or related to what was embedded in the flipped material. In terms of the pedagogical problems faced by the participants, F7 and F8 asserted preferring to listen to teachers' explanations and take notes for later recourse instead of looking for what was asked them, as they were not accustomed to such an instructional method.

*I have been accustomed to listening to my teacher's voice in class while she is explaining the lesson points and taking notes to look them up later. So, this class format was not so appealing to me.*

*During the very first sessions, I was totally demotivated to continue the course, because this kind of classroom was totally bizarre for me…. But I gradually adapted myself to it and I found how it would be of aid to practice autonomous learning. I think it would not be so easy for somebody else to be positively adapted to new instructional situations.*

Some other participants were complaining about the kind of group activities going on during in-class time, indeed they were susceptible about their classmates who would rather do the activities individually as expressed in the following extract by F1:

*I prefer group or pair work instead of individual work, especially when doing instructional activities, but this was not the case with some classmates… They either did not want to share what they knew or were unable to do so … I think this issue is related to their personality type.*

F7 was more in favor of doing the assignments either in the classroom or after it, as she stated:

*As I said, I learn more from teacher explanations instead of browsing different sources to find out what is there, … You know, for me the teacher is the best source who triggers me to think and answer the questions or solve the posed problems … in such [flipped] classrooms I lacked the teacher's prompt in a voice which caused some difficulties for me during the course… in other words, the classroom atmosphere propels me to think more.*

Three members of the FG had problems that were related to the material used in the classroom. According to them, some other classmates had complaints about the same problems by and large. These problematic points are evident in the following extracts:

*Sometimes I lost access to the internet and got frustrated about not being able to download the file shared by my teachers, thus being blamed for not sending the answers on time* (F2).

*I found this kind of assignment and class preparation time-consuming. We had to devote additional time to out-of-class activities both before and after the determined class time… I was totally busy with the course requirements (F4).*

*I think the PowerPoint file could be created more appealing. The audio pronunciation and the pictorial depiction of the word were also good ideas but additional materials can also be added…. I and two of my classmates usually watch English movies or animations and they have helped us a lot to improve our listening. … I think inserting small instructional videos or animations can be very helpful for the learners…. As some students have a powerful photographic and acoustic memory (F7).*

## 5. Discussion and conclusion

The aim of this study was to explore the impact of flipped classrooms on Iranian EFL learners' vocabulary improvement. A sequential mixed-methods study design was adopted to explore the issue the results of which will be reiterated here. The quantitative findings indicated that flipped vocabulary learning was beneficial for the experimental group as they performed better in the post-test, not only in comparison with their pre-test but also compared to the control group results. In addition, the control group did not perform significantly better in the post-test. Although these results differ from some published studies (Alhamami & Khan, 2019; Fraga & Harmon, 2014; Jalili et al., 2020; Mori et al., 2016; Oh, 2017), they are consistent with those of Fethi and Marshall (Fethi & Marshall, 2018), Han (2018), Kim (2018), Knežević et al. (2020), and Özkal (2019). Better achievement in flipped classrooms is due to the before-class preparation required of the students (Kim, 2018) and using online resources (McLaughlin & Rhoney, 2015). Besides, the interactive classroom atmosphere is a great aid for the students to master what has been encountered out-of-class (Zhang et al., 2016). The interactional patterns going on in face-to-face

classrooms were perceived by the participants to be useful and conducive to better learning. As stated by Kumaravadivelu (2008) interpersonal activities have the pedagogical benefit of more learner-learner interaction, as a social process which is deemed as important as cognitive ones in the route to successful L2 achievement.

With regard to the qualitative findings, it goes without saying that most of the students held positive attitudes toward the flipped classrooms which corroborates those of Kirmizi and Kömeç (2019), Moran (2014), Prefume (2015), Santikarn and Wichadee (2018), and Webb et al. (2014). The slides used in the flipped classrooms contained different input modalities including text, audio, and picture. Using visual input is a powerful way of gaining efficiency in language instructions (Özkal, 2019; Kang, 2015). The participants believed that they were more active and engaged in the class activities. According to Ahmed (2016) and Moran (2014), flipped classes foster learners' engagement and motivation to a great degree, however just flipping the classrooms does not and cannot guarantee more learner engagement (Moran, 2014).

In the post-method era, both learners and instructors are encouraged to follow some macro strategies (Kumaravadivelu, 2003) some of which were evident among this study's participants, for instance, the learners promoted their autonomy in vocabulary learning, contextualized the language input, and took part in negotiated interactions (Kumaravadivelu, 2008). In this case, both in-class and before-class activities resulted in such benefits, in other words, flipped classrooms by the use of technology are conducive to active learning, provide a social context for learning, and cause engaging learning materials to be accessed easily (Richards, 2015). In addition, interaction in the form of small groups in the in-class phase of the experimental group was another conducive factor to better learning and positive attitudes. Ellis (2008) asserted that group activities increase social interaction, enjoyment, motivation, learning, and independence more than teacher-centered classrooms.

Apart from these positive attitudes, some technological and material-related challenges were reported by the experimental group. Technological aspects such as system and internet characteristics are among the important factors to take into account during the implementation of any kind of learning which involves a connection to the web (Al-araibi, 2019; Kanwal & Rehman, 2017). The eagerness of this study's participants to include videos is an indication of the facilitative role of audio-visual material in improving vocabulary (Çakmak & Erçetin, 2018; Kabooha & Elyas, 2018) and overall learning performance (Montero Perez, 2020; Ramezanali & Faez, 2019), however, the teachers' role in selecting the most pertinent audio-visual input to the learners' needs should not be neglected (Stempleski, 2002).

Overall, the results of this study showed the positive impact of flipped vocabulary learning among upper-intermediate EFL learners. Moreover, it was found that they held positive attitudes towards such an experience in terms of leading to more learning engagement, more interaction, better preparation, less frustration, and more efficient face-to-face classes. On the other hand, it was revealed that some learners are not already ready to accept the instructional mode as favorable and either prefer traditional classrooms or want their classes to be flipped via different material types. One of the limitations of this study is related to the kind of material used for the experimental group and the rather short span of time devoted to the treatment sessions. As suggested by the participants, the used material could be more influential if teacher-created or available instructional videos were also embedded in the slides. As this was the first time that these students experienced attending flipped classrooms, the trajectories of vocabulary development and their perceptions towards these classes could be traced more successfully if the treatment sessions would be held for more than one semester. Other data collection instruments could also add to the study findings by adding other variables to find relationships among them, for instance, the role of the students' learning and cognitive styles on their performance in flipped classrooms could be explored via surveys and classroom observations. The impact of gender differences on learners' performance and attitude can also be explored by comparing classes with different genders as it is one of the important biological variables affecting learning in general and language learning in particular. This study has valuable implications for those EFL teachers who want to introduce flipped classrooms to their students and e-content developers, as there are many factors to consider when preparing such materials including the students' needs, preferences, learning styles, etc.

## 6. Data availability

The data are available upon request with the corresponding author.

## 7. Conflicts of interests

The authors have no conflicts of interest to declare.

## 8. References

Adnan, M. (2017). Perceptions of Senior-Year ELT Students for Flipped Classroom: A Materials Development Course. *Computer Assisted Language Learning, 30*(3–4), 204–222. https://doi.org /10.1080/09588221.2017.1301958

Ahmad, S. Z. (2016). The Flipped Classroom Model to Develop Egyptian EFL Students' Listening Comprehension. *English Language Teaching*, *9*(9), 166-178. : http://dx.doi.org/10.5539/elt.v9n9p166

Ahmed, A. A. A., Keezhatta, M. S., Khair Amal, B., Sharma, S., Shanan, A. J., Ali, M. H., & Farooq Haidari, M. M. (2022). The Comparative Effect of Online Instruction, Flipped Instruction, and Traditional Instruction on Developing Iranian EFL Learners' Vocabulary Knowledge. *Education Research International*, *2022*.

Ahmed, M. A. E. A. S. (2016). The Effect of a Flipping Classroom on Writing Skill in English as a Foreign Language and Students' Attitude towards Flipping. *US-China Foreign Language*, 14(2), 98-114. doi:10.17265/1539-8080/2016.02.003

Al-araibi, A. A. M., Naz'ri bin Mahrin, M., & Yusoff, R. C. M. (2019). Technological Aspect Factors of E-Learning Readiness in Higher Education Institutions: Delphi Technique. *Education and Information Technologies*, *24*(1), 567-590. https://doi.org/10.1007/s10639-018-9780-9

Alghamdi, H. H. (2019). Exploring Second Language Vocabulary Learning in ESL Classes. *English Language Teaching*, *12*(1), 78-84. doi: 10.5539/elt.v12n1p78

Alhamami, M. (2016). Vocabulary Learning through Audios, Images, and Videos: Linking Technologies with Memory. *CALL-EJ*, *17*(2), 87-112. http://callej.org/journal/17-2/Alhamami2016.pdf

Alhamami, M., & Khan, M. R. (2019). Effectiveness of Flipped Language Learning Classrooms and Students' Perspectives. *Journal on English as a Foreign Language*, *9*(1), 71-86. doi: 10.23971/jefl.v9i1.1046

Alnuhayt, S. S. (2018). Investigating the Use of the Flipped Classroom Method in an EFL Vocabulary Course. *Journal of Language Teaching and Research*, *9*(2), 236-242. http://dx.doi.org/10.17507/jltr.0902.03

Amer, M. (2014). Language Learners' Usage of a Mobile Learning Application for Learning Idioms and Collocations. *Calico Journal*, *31*(3), 285-302. doi: 10.11139/cj.31.3.285-302

Arndt, H. L., & Woore, R. (2018). Vocabulary Learning from Watching YouTube Videos and Reading Blog Posts. *Language Learning & Technology*, *22*(3), 124-142. https://scholarspace.manoa.hawaii.edu/bitstream/10125/44660/1/22_03_arndt_10125-44660.pdf

Bang, Y. (2017). The Effect of Flipped Learning in an EFL Classroom. *Modern English Education*, *18*(2), 87-107. doi: 10.18095/meeso.2017.18.2.05

Bergmann, J., & Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*. International Society for Technology in Education

Berrett, D. (2012). How "Flipping" the Classroom Can Improve the Traditional Lecture. *The Chronicle of Higher Education*, 58 (25), 16-18. https://people.ok.ubc.ca/cstother/How_Flipping_the_Classroom_Can_Improve_the_Traditional_Lecture.pdf

Brown C.J. (2018). Flipping the ESL/EFL Academic Reading Classroom: A Group Leader Discussion Activity. In: Mehring J., Leis A. (eds) *Innovations in Flipping the Language Classroom*. Springer, Singapore.

Carter, R., & McCarthy, M. (2014). *Vocabulary and language teaching*. Routledge.

Chakowa, J. (2018). Enhancing Beginners' Second Language Learning through an Informal Online Environment. *Journal of Educators Online*, *15*(1), 1-14. doi: 10.9743/JEO2018.15.1.7

Challob, A. I. (2021). The effect of flipped learning on EFL students' writing performance, autonomy, and motivation. *Education and Information Technologies*, 1-27. https://doi.org/10.1007/s10639-021-10434-1

Chen, C. M., Chen, L. C., & Yang, S. M. (2018). An English Vocabulary Learning App with Self-Regulated Learning Mechanism to Improve Learning Performance and Motivation. *Computer Assisted Language Learning*, *32*(3), 237-260. doi: 10.1080/09588221.2018.1485708

Cho, Y. S., & Lee, Y. J. (2016). Can Flipped Learning Be Applied to a Test-Based English Classroom? *Modern English Education*, *17*(4), 181-200.

Chuang, H. H., Weng, C. Y., & Chen, C. H. (2018). Which Students Benefit Most from a Flipped Classroom Approach to Language Learning? *British Journal of Educational Technology*, *49*(1), 56–68. doi:10.1111/bjet.12530

Cotterall, S. (1995). Readiness for Autonomy: Investigating Learner Beliefs. *System*, *23*(2), 195–205. doi: 10.1016/0346-251X (95)00008-8

Çakmak, F., & Erçetin, G. (2018). Effects of Gloss Type on Text Recall and Incidental Vocabulary Learning in Mobile-Assisted L2 Listening. *ReCALL*, *30*(1), 24-47. https://doi.org/10.1017/S0958344017000155

Davis, C. (2013). Flipped or inverted learning: Strategies for course design. In E.G. Smyth & J.X. Volker (Eds.), *Enhancing instruction with visual media: Utilizing video and lecture capture* (pp. 241-265). IGI Global.

Dehghan, F., Rezvani, R., & Fazeli, S. (2017). Social Networks and their Effectiveness in Learning Foreign Language Vocabulary: A Comparative Study Using WhatsApp. *CALL-EJ*, *18*(2), 1-13. http://callej.org/journal/18-2/Dehghan-Rezvani-Fazeli2017.pdf

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian Approaches to Second Language Research*. Ablex Publishing.

Dornyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies.* Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition*. Oxford University Press.

Fahandezh, F., & Mohammadi, A. (2021). The Effect of Gamified Flipped Classroom on the Improvement of Vocabulary Learning of Iranian Pre-Intermediate EFL Learners. *Vision: Journal for Language and Foreign Language Learning*, *10*(2), 83-98.

Fethi, K., & Marshall, H. W. (2018). Flipping movies for dynamic engagement. In *Innovations in Flipping the Language Classroom* (pp. 185-202). Springer.

Fraga, L. M., & Harmon, J. (2014). The Flipped Classroom Model of Learning in Higher Education: An Investigation of Pre-service Teachers' Perspectives and Achievement. *Journal of Digital Learning in Teacher Education*, *31*(1), 18-27. https://doi.org/10.1080/21532974.2014.967420

Garrison, D. R., & Kanuka, H. (2004). Blended Learning: Uncovering its Transformative Potential in Higher Education. *Internet and Higher Education*, 7, 95-105. https://doi.org/10.1016/j.iheduc.2004.02.001

Han, Y. J. (2018). Flipping Tech-Enhanced, Content-Based EAP Courses with Online Content. In *Innovations in Flipping the Language Classroom* (pp. 203-220). Springer.

Hung, H.-T. (2015). Flipping the Classroom for English Language Learners to Foster Active Learning. *Computer Assisted Language Learning*, *28*(1), 81–96. https://doi.org/10.1080/09588221.201 4.967701

Huang, Y. N., & Hong, Z. R. (2016). The Effects of a Flipped English Classroom Intervention on Students' Information and Communication Technology and English Reading Comprehension. *Educational Technology Research and Development*, *64*(2), 175-193. https://doi.org/10.1007/s11423-015-9412-7

Jalili, S., Khalaji, H., & Ahmadi, H. (2020). Vocabulary Learning in the Mobile-Assisted Flipped Classroom in an Iranian EFL Context. *Teaching English with Technology*, *20*(4), 82-95. http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-643dc4d5-fe8f-44b0-b8d8-35dd846560db/c/ARTICLE5.pdf

Jinlei, Z., Ying, W., & Baohui, Z. (2012). Introducing a New Teaching Model: Flipped Classroom. *Journal of Distance Education*, *4*(8), 46-51. http://en.cnki.com.cn/Article_en/CJFDTotal-YCJY201204007.htm

Kabooha, R., & Elyas, T. (2018). The Effects of YouTube in Multimedia Instruction for Vocabulary Learning: Perceptions of EFL Students and Teachers. *English Language Teaching*, *11*(2), 72. http://doi.org/10.5539/elt.v11n2p72

Kang, N. (2015). The Comparison between Regular and Flipped Classrooms for EFL Korean Adult Learners. *Multimedia-Assisted Language Learning*, *18*(3), 41-72. doi:10.15702/MALL.2015.18.3.41

Kanwal, F., & Rehman, M. (2017). Factors Affecting E-Learning Adoption in Developing Countries– Empirical Evidence from Pakistan's Higher Education Sector. *IEEE Access*, *5*, 10968-10978. doi: 10.1109/ACCESS.2017.2714379.

Katemba, C. (2019). Students' Vocabulary Enhancement at Grade 10: A Comparative Study Using CALL & MALL in Indonesia. *CALL-EJ 20* (1), 87-114. https://www.researchgate.net/publication/331834170_Students'_vocabulary_enhancement_at_grade_10_A_comparative_study_using_CALL_MALL_in_Indonesia

Khansarian-Dehkordi, F., & Ameri-Golestan, A. (2016). Effects of Mobile Learning on Acquisition and Retention of Vocabulary among Persian-Speaking EFL Learners. *CALL_EJ*, *17*(2), 43-56. https://www.mdpi.com/2227-7102/9/3/179/pdf

Kim, W. Y. (2018). Effects of Flipped Learning on the Learning of English Vocabulary. *Korean Journal of English Language and Linguistics*, 18(4), 470-485. doi: 10.15738/kjell.18.4.201812.470

Kirmizi, Ö., & Kömeç, F. (2019). The Impact of the Flipped Classroom on Receptive and Productive Vocabulary Learning. *Journal of Language and Linguistic Studies, 15*(2), 437-449. doi: 10.17263/jlls.586096

Knežević, L., Županec, V., & Radulović, B. (2020). Flipping the Classroom to Enhance Academic Vocabulary Learning in an English for Academic Purposes (EAP) Course. *SAGE Open*, *10*(3), 1-15, doi: 10.1177/2158244020957052

Ko, M. H. (2019). Students' Reactions to Using Smartphones and Social Media for Vocabulary Feedback. *Computer Assisted Language Learning*, *32*(8), 920-944. https://doi.org/10.1080/09588221.2018.1541360

Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. Yale University Press.

Kumaravadivelu, B. (2008). *Understanding language teaching: From method to post-method*. Lawrence Erlbaum Associates.

Lage, M.J., & Platt, G. (2000). The Internet and the Inverted Classroom. *The Journal of Economic Education*, 31(1), 11-11. doi: 10.1080/00220480009596756

Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.

Lantolf, J.P., & Thorne, S.L. (2006). *Sociocultural theory and the genesis of second language development.* Oxford University Press.

Lantolf, J., Thorne, S. L., & Poehner, M. (2015). Sociocultural Theory and Second Language Development. In B. van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition* (pp. 207-226). Routledge.

Lee, L. (2021). Exploring Learners' Attitudes towards Technology-Enhanced Flipped Language Instruction. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *11*(1), 81-98. doi: 10.4018/IJCALLT.2021010106

Lin, J. J., & Lin, H. (2019). Mobile-Assisted ESL/EFL Vocabulary Learning: A Systematic Review and Meta-Analysis. *Computer Assisted Language Learning*, *32*(8), 878-919. doi: 10.1080/09588221.2018.1541359

Loucky, J.P. & Ware, J.L. eds. (2016). *Flipped Instruction Methods and Digital Technologies in the Language Learning Classroom*. IGI Global.

Ma, Q. (2017). Technologies for Teaching and Learning L2 Vocabulary. In *The handbook of technology and second language teaching and learning*, (pp.45-61). doi: 10.1002/9781118914069.ch4

Ma, X., & Yodkamlue, B. (2019). The Effects of Using a Self-developed Mobile App on Vocabulary Learning and Retention among EFL Learners. *PASAA*, 58. 166-205. https://files.eric.ed.gov/fulltext/EJ1227379.pdf

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design.* Routledge.

McLaughlin, J., & Rhoney, D. (2015). Comparison of an Interactive E-Learning Preparatory Tool and a Conventional Downloadable Handout Used within a Flipped Neurologic Pharmacotherapy Lecture. *Currents in Pharmacy Teaching and Learning*, *7*(1), 12-19. doi:10.1016/j.cptl.2014.09.016

Montero Perez, M. (2020). Incidental Vocabulary Learning through Viewing Video: The Role of Vocabulary Knowledge and Working Memory. *Studies in Second Language Acquisition*. 1-25. doi: 10.1017/S0272263119000706

Moran, C. (2014). *Changing Paradigms: A Mixed Methods Study of Flipping the English Language Arts Classroom*. (Doctoral Dissertation, North Carolina State University)

Mori, Y., Omori, M., & Sato, K. (2016). The Impact of Flipped Online Kanji Instruction on Written Vocabulary Learning for Introductory and Intermediate Japanese Language Students. *Foreign Language Annals, 49*(4), 729-749. https://doi.org/10.1111/flan.12222

Nation, I. S. P. (1997). Bring today's vocabulary research into tomorrow's classrooms. In Jacobs, G. M. (Ed.) *Language Classrooms of Tomorrow: Issues and Responses*, pp. 170-182. SEAMEO Regional Language Centre

Network, F. L. (2014). The Four Pillars of FLIP. [Online] Available at: http://www. flippedlearning. Org/cms/lib07/VA01923112/Centricity/Domain/41. *FLIP_ha ndout_FNL_Web. Pdf. [accessed 12.9. 2020]*.

Novak, G.M. (2011). Just-in-Time Teaching. *New Directions for Teaching and Learning*, 2011(128), 63-73. https://doi.org/10.1002/tl.469

Nugroho, A., & Fitriati, R. (2021). Flipped Learning Instruction and Pragmatic Competence: A Case of English for Accounting Students. *English Learning Innovation*, *2*(1), 1-9. doi: 10.22219/englie.v2i1.14646

Oh, E. (2017). The Effect of Peer Teaching via Flipped Vocabulary Learning on Class Engagement and Learning Achievements. *Multimedia-Assisted Language Learning, 20*(3), 105-127. http://journal.kamall.or.kr/wp-content/uploads/2017/10/MALL_v20_n3_Oh_05.pdf

Özkal, C. (2019). *Flipped vocabulary learning among Turkish learners of English as a foreign language: a sequential explanatory mixed method study* (Master's Thesis, Middle East Technical University, Ankara).

Peters, E., & Webb, S. (2018). Incidental Vocabulary Acquisition through Viewing L2 Television and Factors that Affect Learning. *Studies in Second Language Acquisition*, *40*(3), 551-577. doi: 10.1017/S0272263117000407

Prefume, Y. (2015). *Exploring a Flipped Classroom Approach in a Japanese Language Classroom: a Mixed Methods Study*. (Doctoral Dissertation, Baylor University)

Pyo, K. H. (2017). Variables Influencing Learner Achievement and Satisfaction in a Flipped Classroom. *Studies in Linguistics* (43), 345-373. doi: 10.17002/sil..43.201704.345

Pyo, K. H. (2018). Effects of Flipped Classroom on University Students' Academic Self-Efficacy. *Studies in Linguistics* (49), 553-581. doi:10.17002/sil..49.201810.553

Ramezanali, N., & Faez, F. (2019). Vocabulary Learning and Retention through Multimedia Glossing. *Language Learning & Technology*, *23*(2), 105–124. https://doi.org/10125/44685

Rezaei Fard, Z., Shahrokhi, M., & Talebinejad, M. R. (2021). The effect of flipped classroom on Iranian ESP students' vocabulary learning, retention and attitude. *International Journal of Foreign Language Teaching and Research*, *9*(35), 115-129.

Riazi, A. M. (2016). *The Routledge encyclopedia of research methods in applied linguistics*. Routledge.

Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.

Roth, C. &Suppasetseree, S. (2016). Flipped Classroom: Can it Enhance English Listening Comprehension for Pre-University Students in Cambodia? *Proceedings of CLaSIC* 2016, 255-264. *https://www.Fas.Nus.Edu.Sg/cls/CLaSIC/clasic2016/PROCEEDINGS/Roth cranny*.Pdf255-263

Santikarn, B., & Wichadee, S. (2018). Flipping the Classroom for English Language Learners: A Study of Learning Performance and Perceptions. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(09), 123-135. https://doi.org/10.3991/ijet.v13i09.7792

Schmitt, N. (2010). *Researching vocabulary.* Palgrave Macmillan.

Shadiev, R., Wu, T. T., & Huang, Y. M. (2020). Using Image-to-Text Recognition Technology to Facilitate Vocabulary Acquisition in Authentic Contexts. *ReCALL*, *32*(2), 195-212. doi: 10.1017/S0958344020000038

Stempleski, S. (2002). Video in the ELT Classroom: The Role of the Teacher. *In Methodology in language teaching. An anthology of current practice*, (pp. 364-367. https://doi.org/10.1017/CBO9780511667190.052

Stockwell, G. (2007). Vocabulary on the Move: Investigating an Intelligent Mobile Phone-Based Vocabulary Tutor. *Computer Assisted Language Learning*, *20*(4), 365-383. doi: 10.1080/09588220701745817

Tai, T. Y., Chen, H. H. J., & Todd, G. (2020). The Impact of a Virtual Reality App on Adolescent EFL Learners' Vocabulary Learning. *Computer Assisted Language Learning*, 1-26. https://doi.org/10.1080/09588221.2020.1752735

Tawfik, A. A., & Lilly, C. (2015). Using a Flipped Classroom Approach to Support Problem-Aased Learning. *Technology, Knowledge and Learning*, *20*(3), 299-315. doi 10.1007/s10758-015-9262-

Teng, F. (2020). Vocabulary Learning through Videos: Captions, Advance-Organizer Strategy, and their Combination. *Computer Assisted Language Learning*, 1-33. https://doi.org/10.1080/09588221.2020.1720253

Teng, M. F. (2017). Flipping the Classroom and Tertiary Level EFL Students' Academic Performance and Satisfaction. *The Journal of Asia TEFL*, *14*(4), 605–620. https://doi.org/10.18823/asiatefl.2017.14.4.2.605

Terantino, J. (2016). Examining the Effects of Independent MALL on Vocabulary Recall and Listening Comprehension: An Exploratory Case Study of Preschool Children. *Calico journal, 33*(2).260-278. doi: 10.1558/cj.v33i2.26072

Tseng, W. T., Liou, H. J., & Chu, H. C. (2020). Vocabulary Learning in Virtual Environments: Learner Autonomy and Collaboration. *System*, *88*, 102-190. https://doi.org/10.1016/j.system.2019.102190

Turuk, M. (2008). The Relevance and Implications of Vygotsky's Sociocultural Theory in the Second Language Classroom. *Annual Review of Education, Communication & Language Sciences*, *5*, 244–262. https://www.academia.edu/download/37336864/vygotsky.pdf

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

Vygotsky, L. S. (1998). In R. W. Rieber (Ed.). *The collected works of L. S. Vygotsky*. Volume 5, Child psychology. Plenum.

Wang, Y. T. (2019). Effects of L1/L2 Captioned TV Programs on Students' Vocabulary Learning and Comprehension. *Calico Journal*, *36*(3). 204-224. https://doi.org/10.1558/cj.36268

Webb, M., Doman, E., & Pusey, K. (2014). Flipping a Chinese University EFL Course: What Teachers and Students Think of the Model. *The Journal of Asia TEFL, 11*(4), 53–87. http://www.asiatefl.org/main/download_pdf.php?i=389&c=1420009720&fn=11_4_03.pdf

Wei, J. (2019, March). Feasibility Analysis on Application of Dynamic Assessment in Flipped Classroom of College English. In *2018 8th International Conference on Education and Management (ICEM 2018)*. Atlantis Press.

Wong, K. M., & Samudra, P. G. (2019). L2 Vocabulary Learning from Educational Media: Extending Dual-Coding Theory to Dual-Language Learners. *Computer Assisted Language Learning*, 1-23. https://doi.org/10.1080/09588221.2019.1666150

Wu, W. V., Yang, J. C., Chen Hsieh, J. S., & Yamamoto, T. (2019). Free from Demotivation in EFL Writing: The Use of Online Flipped Writing Instruction. *Computer Assisted Language Learning, 33*(4), 353–387. https://doi.org/10.1080/09588221. 2019.1567556

Yang, J. (2017). Effects of Flipped Learning on University Students' Acquisition of Collocation and Attitude. *New Korean Journal of English Language and Literature* 59(1), 145-161. http://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE07113193

Yuliani, K. A., Made Hery Santosa, S. P., & Saputra, I. N. P. H. (2018). The Effect of Project-Based Flipped Learning on Writing Competence of the Eleventh-Grade Students of SMA N 4 Singaraja in Academic Year 2017/2018. *Jurnal Pendidikan Bahasa Inggris undiksha, 4*(2). https://ejournal.undiksha.ac.id/index.php/JPBI/article/view/13274

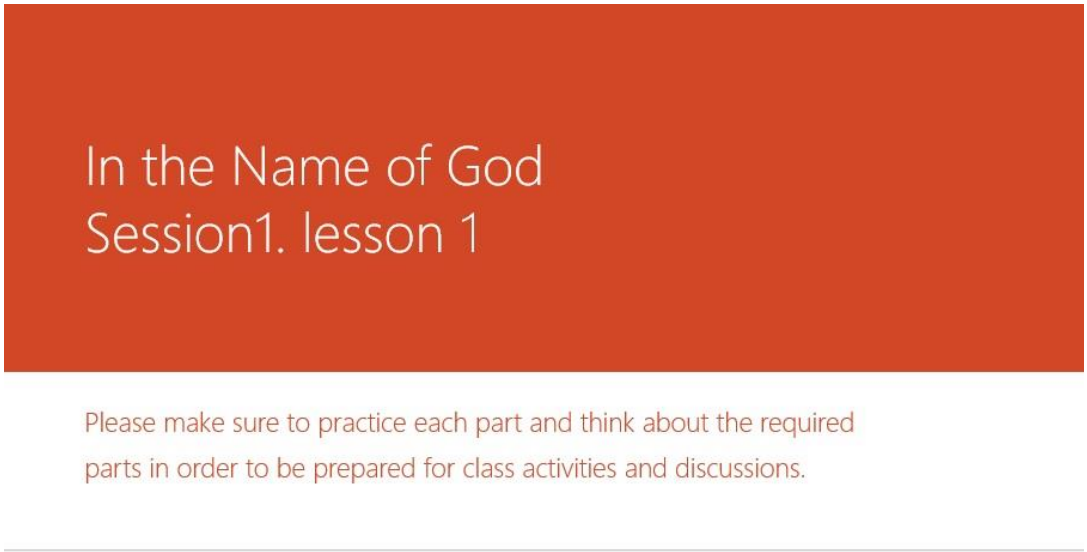
Zainuddin, Z., & Halili, S. H. (2016). Flipped Classroom Research and Trends from Different Fields of Study. *The International Review of Research in Open and Distributed Learning, 17*(3). doi:10.19173/irrodl.v17i3.2274

Zhang, H., Li, J., Jiao, L., Ma, W., & Guan, C. (2016). The Adjustment and Effects of Vocabulary Teaching Strategies in Flipped Classroom. *Creative Education. 7*(14), 1966-1973. doi: 10.4236/ce.2016.714199

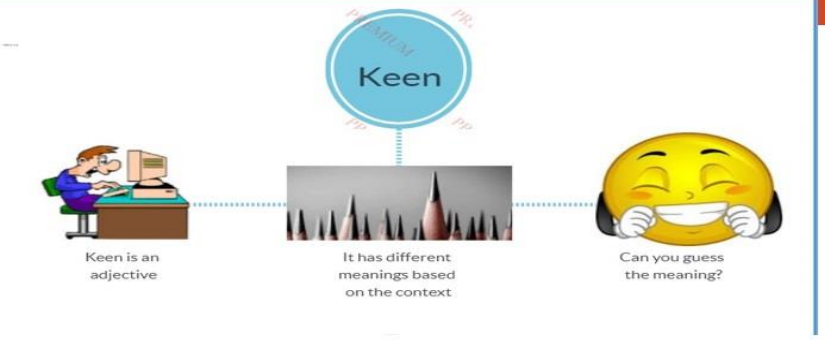
## 9. Appendices

## Appendix A

Screenshots of the slides for the FG



1.




## 1. Keen

-Wanting to do something or wanting something to happen very much.

**-A keen knife is very sharp.**

✓ He told me that he was keen to help.

Can you think of any synonyms/antonyms? Write them.

……..

Find out what words and prepositions combine with this word.

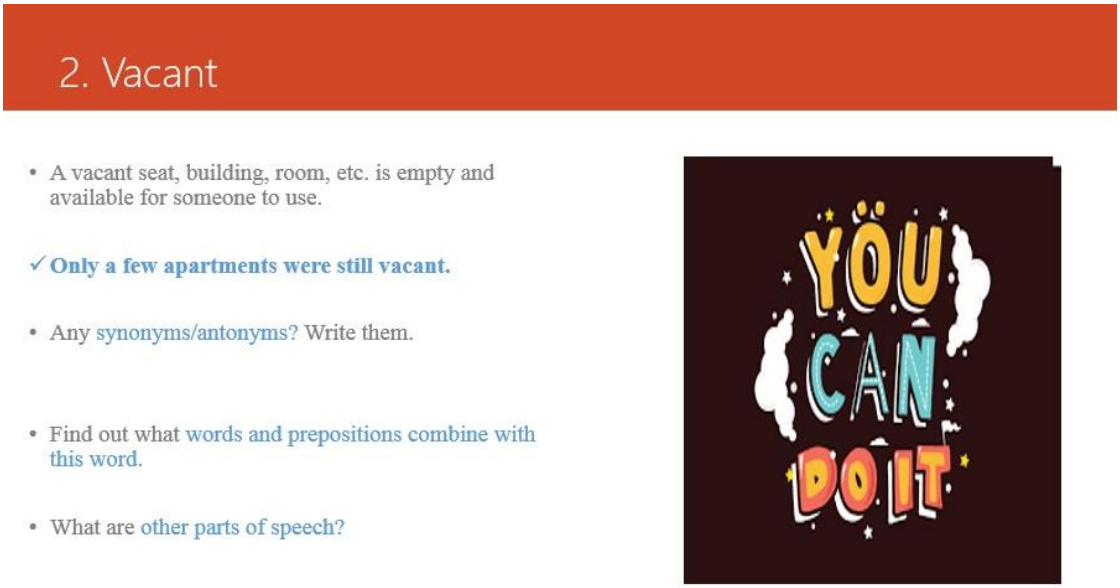
What are other parts of speech?

2.

Vacant

Adjective

Did you guess the meaning?




## Appendix B

Interview guiding questions

**1. Did you enjoy vocabulary learning via this method? Why and why not?**

**2. What do you think about the benefits of vocabulary learning using this method?**

**3. What do you think about the challenges of vocabulary learning using this method?**

**4. What are your suggestions for improving future courses instructed this way?**

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 113-128*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2023.6374.1016

# Investigating the Relationship Between Collective Teacher Efficacy and Teachers' Withdrawal Intention

Fateme Fadaeian[1]
Mohammad Aliakbari[2]

[1]M.A. Graduate, Department of English Language and Literature, University of Ilam, Ilam, Iran
[2]Professor in Applied Linguistics, Department of English Language and Literature, University of Ilam, Ilam, Iran

**ARTICLE HISTORY**

Received: 27 August 2022
Revised: 27 October 2022
Accepted: 24 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail:
fatemeh.fadaeeyan1998@gmail.com

**ABSTRACT**

This study aims at examining the link between collective instructor efficacy and withdrawal intention. Data was gathered through an online survey to collect information from Iranian EFL teachers using two questionnaires. One questionnaire asked them about their collective efficacy perception and another regarding their withdrawal intent or lack thereof. Participants included 208 male and female educators who were working in either public schools, private institutions or both participated in the scrutiny. Through the medium of the SPSS software, the data were scrutinized. The findings of this analysis noted that there exists a negative association between instructors' collective efficacy and withdrawal intention. As collaborative competence perceptions elevate in teachers, their withdrawal intent is likely to drop. Furthermore, another research question was whether it was possible for collective efficacy subscales to predict withdrawal intention or not. Instructional strategies and student discipline are subsets of collective efficacy. Student discipline was reported to predict to some extent variability in withdrawal intention. It was also concluded that headmasters and educational leaders need to focus on collective efficacy as an asset to diminish the unfavorable attrition of staff members.

**KEYWORDS:** Collective efficacy; Self-efficacy; Withdrawal intention; EFL teacher; Teacher efficacy; Teacher turnover

## 1. Introduction

According to Tella (2008), the educators' perceptions and conceptions affect the leaners in direct and indirect ways. How educators observe, think, and act influence their instruction and their learners' attainments of teaching materials. An important issue that might occupy educators' minds is withdrawal intention. In general, intent to leave is choosing to withdraw from a professional position (Lee & Mitchell, 1994). Teachers' withdrawal intention causes complications that have special educational significance since generally, superior teachers contemplate the possibility of leaving, which endangers the quality of teaching at schools (Ingersoll, 2001). Teachers might leave their workplaces for various reasons. They might feel they are not valued or that their contribution is not meaningful. According to Kaplan et al., (2009), educators with withdrawal intention have decreased work participation. Decreased job satisfaction, increased emotional exhaustion, and reduced commitment (Da'as et al., 2020; Shin et al., 2020) all associated with enhanced withdrawal intention. The current study on teachers' withdrawal intention focuses on its relationship with Collective Teacher Efficacy (CTE). CTE is characterized as the common conviction of instructors working in the same organization that how much their cumulative efforts can impact all students' achievement inclusive of disinterested or deprived ones positively (Hattie, 2016). Therefore, CTE is closely related to how teachers working in the same organization will perceive to what extent their efforts have constructive effects on their learners. Cumulative efficacy positively influences the self-efficacy of educators, student achievement, teacher leadership, and school

**Cite this article:** Fadaeian, F., & Aliakbari, M. (2023). Investigating the relationship between collective teacher efficacy and teachers' withdrawal intention. *Applied Linguistics Inquiry, 1*(1), 113-128. doi: 10.22077/ali.2023.6374.1016

improvement (Donohoo, 2018; Loughland & Ryan, 2022). Increasing educators' collective efficacy has been reported to result in lowering their burnout, boosting their job satisfaction, and advancing their perceived competence (Dixon et al., 2014).

Workers who contemplate leaving their job are inclined to decrease the quality of their work (Maertz & Campion, 1998). When employees start considering leaving their job, they start to shift their focus and energy to find alternative jobs and this in turn, affects their performance in their current profession. Qadach et al. (2020) worked on the link between instructional leadership and educators' withdrawal intention as well as mediators such as collective efficacy and joint conception. Shapira-Lishchinsky and Rosenblatt (2009) proposed an approach to the investigation of organizational ethics and teachers' withdrawal behavior. In the local context of Iran, an analysis of the influence of business ethical values on withdrawal was executed by Abzari et al. (2015).

Insight into teachers' withdrawal intention and its relationship with instructors' collective efficacy contributes to the existing area of analysis and increases communal awareness of the problems teachers encounter on a daily basis. Results from this study are beneficial to policymakers to potentially modify regulations in order to avoid the drop-out of teachers. In addition, results can assist teacher educators to fortify teachers with the kinds of skills and mindsets that would help them along the way in facing obstacles in the classroom or the organization. Moreover, teachers themselves can acknowledge different aspects of the story that they might currently live with; such as what they might be able to do in their existing workplace to enhance their circumstances.

The current study supplies explanations for the subsequent queries:

1. Is there any relationship between collective teacher efficacy and teachers' withdrawal intention among Iranian EFL teachers?
2. Do any of the subscales of collective teacher efficacy predict withdrawal intention among Iranian EFL teachers?
3. Is there a significant role of gender in teachers' collective efficacy?
4. Is there a significant role of teaching experience in EFL teachers' withdrawal intention?

## 2. Literature review

### 2.1. Theoretical background

The present-day views of personal and collective efficacy owe much to the theoretical background of social cognitive theory (Bandura, 1986). Social cognitive theory suggests that humans' accomplishments revolve around their individual characteristics and their surrounding milieu (Bandura, 2002). Bandura (1977) worked on a theoretical framework to investigate psychological transformations accomplished by various treatments. Rotter (1966) carried out an analysis on internal opposed to external control of intensification to investigate the group differences in the behavior of subjects when they receive reinforcement. Findings on Americans taking part in this study showed that if they perceive that a particular situation is dependent on external factors such as chance, they do not intend to have high expectations for future reinforcement. In other words, since they perceived chance factors that they cannot predict, significant in that situation they most likely do not see their success happening again. Therefore, their failure shakes their confidence less than a situation they discern to be dependent upon their own behavior. Results suggested that subjects' expectancies and perceptions of the situation whether learning situations or in general, predict their behavior.

In the major study by Mobley (1977), it is suggested that when employees are discontented with their job, they search for other professional positions and their search leads to their eventual departure from or continuation of their current job. There are studies strongly supporting the association between withdrawal intention and withdrawal behavior (Bluedorn, 1982; Porter & Steers, 1973).

According to Abelson and Baysinger's (1984) study, optimal turnover is defined as the equilibrium between the expenses of employee attrition and the expenses linked with employee retention. Furthermore, if there is an imbalance between the costs of the staff members' attrition and retention, this asymmetry would be regarded as dysfunctional turnover. When the cost of keeping a job holder is too high, employers gladly receive high turnover rates. The other extreme situation is when the cost of keeping an employee is relatively low, the company can survive reduced turnover rates. Companies look forward to a balance and equilibrium between the two situations. Very high turnover rates are associated with high costs of looking for and educating recent job holders and at the same time, very low turnover rates are associated with costs of keeping employees with unsatisfactory performances which eventually trigger the turnover rate of other employees. Moreover, individual, organizational, and environmental factors impact the inclination of employees toward quitting. Therefore, the eventual turnover rates of companies are impressed by the aforementioned factors' influence on withdrawal tendency.

### 2.2. Empirical framework

In the analysis of Glassman et al. (2021), an instrument of collective efficacy was developed and validated. The data was gathered using a sample that included 634 educators. The results of the convergence validity analysis indicated that this proposed instrument had significant and moderate links to the other well-established instrument of similar constructs. The model proposed by this study points out that there is a cognitive filter at two levels: individual and group. In other words, an educator has some experience as an independent human being and some as a team member. These two levels shape actions or behaviors of educators and behavior in turn lead to feedback from the surroundings. Feedback eventually provides resources for cognitive filters. Collective efficacy is generated by the encounters people have with their surroundings both as individuals and as team members.

In an attempt made by Da'as et al. (2021), the perceived collective teacher efficacy across three nations and four cultures was explored. The sample included 4,216 Arab, Jewish, Turkish, and American instructors in total. Multigroup confirmatory factor analysis was employed for the purpose of testing configural invariance. The results of the study point out a sufficient fit of the framework over these nations with the use of configural invariance. Another emphasis of this analysis was to check whether the collaborative teacher efficacy instruments held separate meanings for people in these three countries. The differences in the collective efficacy results among the countries could be traced back to their cultural differences.

Meyer et al. (2022) make the case for the association between headmaster administration and instructor partnership. Additionally, the mediating part of collaborative instructor efficacy is investigated thoroughly. The sample for the data collection included 630 primary and secondary school educators in 29 institutions in Germany. The results from analyzing the structural equation modeling revealed that teacher collaboration is indirectly influenced by principal leadership with the mediating role of cumulative competence impression. To rephrase it, at schools in which principals are trying to promote alliance and cooperation among educators, there is a good chance that educators find this atmosphere a good place to put in effort for a common goal of enhancing student achievement. Where there is ongoing involvement of educators in such valuable endeavors, these educators are more inclined to strongly believe that they have what it takes to impact the academic performance of pupils.

Schechter et al. (2020) worked on the model of the associations among principals' cognitive complexity, school's absorptive capacity, educators' affective commitment, and withdrawal intention. A total of 1,664 elementary school teachers were surveyed and structural equation modeling was utilized. The results showed that schools' absorptive capacity and educators' affective commitment are mediators between withdrawal intent and principals' cognitive complexity. Principals' cognitive complexity reduces withdrawal intention among educators. Understanding the operations within a work environment helps establish better regulations or planning inside the school to promote a pleasing rate of intent to stay.

Qadach et al. (2020) carried out an investigation to scrutinize the association between headmasters' educational administration and withdrawal intention. The go-between roles of collaborative instructor efficacy and collaborative insight were also inquired into. A combined number of 1,830 elementary school instructors were surveyed to gather the data needed. The first result of this analysis was support for the mediation of collective instructor competence and shared vision in the association among principals' instructional leadership and withdrawal intention. There was found to be a clear correlation between instructional management and collaborative instructor competence. There was also a negative link between collective teacher efficacy and withdrawal intention. In other words, increments in instructional leadership through promoting a common foresight among educators and promoting their perceived competence to influence students' achievement will coincide with a reduction in contemplations of withdrawal and departure.

Due to the importance of instructor variables, this investigation will scrutinize the association between educators' collective efficacy and teachers' withdrawal intention among Iranian EFL educators. Qadach et al. (2020) analyzed the two variables of collective efficacy and intent to leave with a different lens than the current research. The focus was on the role of headmasters on education and the mediating role of collective efficacy on the association between instructional leadership and intent to leave. In similar studies only school teachers were scrutinized. The probability of predicting withdrawal intention by data from the collective efficacy of teachers is also examined here. It also attempts to capture the gender gap in instructors' collective efficacy and the role of experience in the withdrawal intention rate. These issues are under-researched in the EFL context. The current analysis aspires to straighten out the conclusions of previous research on teachers, hopefully, helping to explain the results of those analyses, and perhaps giving us more credible results.

## 3. Method

### 3.1. Design

The present analysis used a quantitative approach and a single-group correlational research design. It was conducted using an online survey. Data was gathered through purposive sampling by distributing the questionnaires via the internet for Iranian EFL educators to fill out. Present work intended to measure the correlation between the aforementioned variables by calculating a correlation coefficient to show the degree of the relationship. Collective teacher efficacy and withdrawal intention are the two

variables observed to examine if there is a correlation, the direction of the correlation, and its magnitude. Iranian nationality and teaching English as a foreign language are two controlling variables in this study. Overall, gender and experience were moderators of the association between collective efficacy and withdrawal intention.

## 3.2. Participants

The sample of this investigation incorporated Iranian EFL teachers consisting of both male and female participants. Approximately 208 teachers, 96 women, and 112 men participated in the analysis. The entire sample of attendees were native speakers of Persian. They came from different cultural backgrounds and differed in their first language. Participants ranged from undergraduate (B.A.) students to Doctor of Philosophy (Ph.D.) holders. Their age range was 18 to 63 years old, with most of the teachers aged from 25 to 31 years old. Participants of this study consisted of those who teach at schools, in private institutions, or in both.

**Table 1.** The Demographic Characteristics of the Participants

| ***Demographic Variables*** | ***Frequency*** | ***Percentage*** |
|---|---|---|
| ***Age groups*** | N=208 | 100 |
| 15-25 | 36 | 17% |
| 26-35 | 95 | 46% |
| 36-45 | 50 | 24% |
| 46-55 | 24 | 12% |
| 56-65 | 3 | 1% |
| ***Academic Degree*** | N=208 | 100 |
| *Diploma* | 2 | 1% |
| *Associate* | 1 | 0.5% |
| *Bachelor* | 72 | 34.6% |
| *Master* | 115 | 55.2% |
| *Ph.D.* | 18 | 8.7% |
| ***Years of Experience*** | N=208 | 100 |
| *1-8* | 93 | 44.7% |
| *9-16* | 63 | 30.3% |
| *17-24* | 22 | 10.6% |
| *25-32* | 27 | 13% |
| *33-40* | 3 | 1.4% |

As is presented in Table 1, the most frequent age range is between the ages of 26 to 35 and the least frequent age range belongs to ages 56 to 65 which is only three people. The youngest instructor taking part in this study was 18 years old and the oldest was 63 years old. The least frequent academic degree was the associate degree (0.5%) and the most frequent one was the master's degree (55.2%). The participants holding the diploma (between 18 to 25 years old) and associate degree (between 26 to 35 years old) were among the younger teachers. About 98.5 % had at least a B.A. degree. Most of the teachers (44.7%) participating in this investigation had between 1 to 8 years of teaching background.

### 3.3. Instruments

Two scales were translated into Persian and employed in this study. The first measure that was used in the current work is the 12-item Collective Teacher Beliefs scale created by Tschannen-Moran and Barr (2004). This measure consists of two subscales. Six items assess Collective efficacy for instructional strategies and six other items assess Collective efficacy for student discipline. Both subscales are five-point Likert type ranging from *nothing* to *a great deal*.

The reliability of the whole measure reported by Tschannen-Moran and Barr (2004) is .97 and factor loadings reported at least .58 to a maximum of .79. The six-item subscale of instructional strategies demonstrates reliability of .96 and its factor loadings range from .78 to .67. The other six items assessing collective teacher efficacy for student discipline have a reliability of .94 and factor loading range from .78 to .64 (Tschannen-Moran and Barr, 2004).

The present study used another survey consisting of five items developed by Shapira-Lishchinsky (2012). This measure asks teachers to rate the five-point Likert-type items from *strongly disagree* to *strongly agree* for the purpose of determining their withdrawal intention. Qadach et al. (2020) who used this scale reported its reliability to be .92. Using confirmatory factor analysis Qadach et al. (2020) on the survey of teachers' intent to leave developed by Shapira-Lishchinsky (2012), reported an incremental fit index of .95, confirmatory fit index .96, and Tucker-Lewis index of .94.

To avoid any ambiguity of the questions for the Iranian teachers, the two questionnaires were translated into Persian. Then, the translated questionnaires were back-translated into English by an expert in translation to ensure the clarity of the translated versions. Next, the original English versions and the back-translated versions were compared which showed a high similarity between them.

### 3.4. Procedures for data collection and analysis

In this study questionnaires were distributed via the internet among Iranian EFL teachers. Since the aim is to test whether there exists an association between collective instructor efficacy and teachers' withdrawal intention, a statistical procedure for correlation is needed. The Pearson product-moment correlation coefficient (Pearson r) is a practical statistic and the direction and magnitude of the association are also demonstrated through this medium (Ary et al, 2018, p.152). Statistical Package for the Social Sciences (SPSS) software was used in order to scrutinize the data and be able to generalize the results from the sample to the population. Multiple regression was used to determine whether the subscales of collective efficacy can predict withdrawal intention. Independent samples t-tests were utilized to uncover if there is a gender gap in collaborative instructor efficacy and if experien ce plays a significant role in the withdrawal intention of instructors.

## 4. Results

### 4.1. Test of normality

The Kolmogorov-Smirnov test was applied to try out the normality of data apportionment. The inspection was employed for the purpose of checking if the allocation deviates from a normal apportionment. The first possible outcome is when the *p*-value is not significant ($p>.05$), this shows that the allocation of data in this study is not considerably dissimilar to a normal apportionment. Consequently, it can be considered a normal distribution. The second outcome is when the *p*-value is significant ($p<.05$), this means that the allocation of data in this study is remarkably dissimilar to a normal apportionment.

**Table 2.** The Results of K-S Test for Withdrawal Intention and Collective Efficacy

| | **Kolmogorov-Smirnov[a]** | | |
|---|---|---|---|
| | Statistic | Df | Sig. |

| | | | |
|---|---|---|---|
| **Withdrawal intention** | .075 | 208 | .185 |
| **Collective efficacy** | .078 | 208 | .144 |

According to table 2, the value acquired from the instruments in this study (withdrawal intention and collective efficacy) is greater than .05. Thus, it can be deduced that the data is normally spread throughout the two variables.

## 4.2. Descriptive statistics

Almost in every research project, the basic descriptive statistics and the common statistics such as mean, standard deviation, minimum, and maximum of the tallies are provided.

**Table 3.** Descriptive Statistics of Collective Efficacy and its Comprising Factors

| | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|
| **Instructional strategies** | 208 | 7.00 | 30.00 | 20.12 | 4.73 |
| **Student discipline** | 208 | 6.00 | 30.00 | 19.95 | 4.58 |
| **Collective efficacy** | 208 | 15.00 | 60.00 | 40.07 | 9.03 |
| **Withdrawal intention** | 208 | 5.00 | 25.00 | 13.72 | 5.53 |

Table 3 indicates descriptive statistics of teachers' collective efficacy along with its two sub-factors: instructional strategies and student discipline. The educational tactics factor has a higher mean than student discipline (M = 20.12). The descriptive statistics of collective efficacy are as follows: (*M*= 40.07, *SD*=9.03). Teachers' collective efficacy scale comprised 12 five-Likert type items so that the possible range of score could be between 12 to 60. As table 3 represents, the minimum, maximum, mean, and Std. Deviations are 5.00, 25.00, 13.72, and 5.53 respectively.

The outcomes of descriptive statistics of the two groups are represented in Table 4. As the table indicates, the mean scores of participants are different in the two groups (1: male, 2: female). The mean score of male participants (M= 118.03) is more than the mean score of female participants (M= 98.22).

**Table 4.** Descriptive Statistics of CTE and Teaching Experience in Two Groups

| | degree | N | Mean | Std. Deviation | Std. Error Mean |
|---|---|---|---|---|---|
| **CTE** | 1 | 114 | 118.03 | 7.13 | 3.02 |
| | 2 | 94 | 98.22 | 7.02 | 2.71 |
| **Teaching Experience** | 1 | 115 | 87.14 | 5.27 | 4.32 |
| | 2 | 93 | 79.22 | 5.00 | 3.14 |

In addition, the mean scores of participants are different in the other two groups (1: more than 10 years of experience, 2: less than 10 years of experience) and the mean score of the first group (M= 87.14) is more than the mean score of the second group (M= 79.22).

## 4.3. Inferential statistics

### 4.3.1. Correlation

Correlation is a procedure in which the researcher analyses the potential association between the variables. Considering that the data gathered from these investigations are interval, the Pearson product-moment formula was adopted.

The first research inquiry of the current analysis was whether there exists a significant association between EFL instructors' collective efficacy and withdrawal intention. The null hypothesis based on the questions is:

$H_{01}$: There is not any significant association between withdrawal intention and collaborative efficacy among EFL teachers.

**Table 5.** The Correlation between Teachers' Collective Efficacy and Withdrawal Intention

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **1.Instructional strategies** | 1 | .87** | .97** | -.39** |
| **2. Student discipline** | .87** | 1 | .96** | -.43** |
| **3. Collective efficacy** | .97** | .96** | 1 | -.42** |
| **4. Withdraw intention** | -.39** | -.43** | -.42** | 1 |

**. Correlation is significant at the 0.01 level (2-tailed)

As table 5 indicates, withdrawal intention associated moderately with both student discipline ($r$ = -0.43, $p$ < 0.01) and instructional strategies ($r$ = -0.39, $p$ < 0.01). Collective efficacy and withdrawal intention correlate moderately with each other. (r = -0.42, p < 0.01).

### 4.3.2. Multiple regression analysis

Regression is a procedure that is applied to the data in order to anticipate variability in the dependent variable in accordance with one or more independent variables. Regression was utilized for the purpose of realizing whether any of the subscales of collective efficacy could predict withdrawal intention.

The second research question is whether EFL teachers' withdrawal intention can be predicted by teachers' collective efficacy. The null hypothesis in accordance with this inquiry is:

$H_{02}$: No subscales of collective efficacy can predict EFL teachers' withdrawal intention.

**Table 6.** Variability in EFL Teachers' Withdrawal Intention due to its Predictors

| Model | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | 1189.065 | 1 | 1189.065 | 47.504 | .000b |
| Residual | 5156.315 | 206 | 25.031 | | |
| Total | 6345.380 | 207 | | | |

a. Dependent Variable: Withdrawal intention

b. Predictors: (Constant)

The analysis of variance (ANOVA) is a valuable approach in order to measure whether the regression model, with scores on student discipline as its predictor, has a linear relationship to withdrawal intention scores. Fisher statistic is used to ascertain whether the formula is significant. The F value is compared to the Fisher's table which has all the established criteria. A concept can only be included in the formula if the likelihood linked to the F tests is smaller than or equal to the level of the specified significance. Table 7 illustrates that F (1,206) = 47.50, p ≤ .05. Consequently, it can be induced that there exists a significant association between the two concepts, this model is adequately proper, and student discipline is considered a fine predictor of teachers' withdrawal intention.

**Table 7.** Correlation Coefficient of Withdrawal Intention and the Predictor

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 24.151 | 1.552 | | 15.563 | .00 |
| | Student discipline | -.522 | .076 | -.433 | -6.892 | .00 |

The magnitude of the coefficient for any of the independent variables exhibits the magnitude of the impact that each variable has on the dependent variable in linear regression. Furthermore, the direction of the coefficient manifests the direction of this influence. The coefficient is liable to rise if it is positive and inclined to drop if it is negative. Table 7 encompasses the regression formula. As demonstrated in column B, the magnitude of the regression coefficient is submitted. Moreover, these values offer details on the noteworthiness of each of the variables.

As shown in Table 8, the formula was: Prediction to withdrawal intention= 24.151 + (-.522) * student discipline

**Table 8.** Model Summary of the R Square of the Correlation Coefficient between Withdrawal Intention and Predictors

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .43 | .18 | .18 | 5.00 |

a. Predictors: (Constant), student discipline
b. Dependent Variable: Withdrawal intention

The correlation coefficient is hereby demonstrated within this table by the R-value. The R square is the fraction of the variation that is predicted by independent variables. As Table 8 shows, that R amounts to .43 and R square is .18. R square is possible to be taken as the magnitude of the expected alteration for the purpose of informing that the student discipline marks are able to anticipate about 18% of the alteration in withdrawal intention which is not a small amount. The column denoted as adjusted R square amounts to .18. Adjusted R square is computed in order to avoid the overvaluation of R square. The final column shows the standard error of estimate equals 5.00.

**Table 9.** Residuals Statistics[a]

| | Minimum | Maximum | Mean | Std. Deviation | N |
|---|---|---|---|---|---|
| **Predicted Value** | 8.47 | 21.01 | 13.72 | 2.39 | 208 |
| **Residual** | -12.88 | 15.47 | .00 | 4.99 | 208 |
| **Std. Predicted Value** | -2.19 | 3.04 | .00 | 1.00 | 208 |
| **Std. Residual** | -2.57 | 3.09 | .00 | .99 | 208 |

a. Dependent Variable: Withdrawal intention

Table 9 includes the non-standardized predicted and residual magnitudes. In addition, the standardized (std.) predicted and residual magnitudes are also provided. The dissimilarity between the acquired and anticipated values is called the residual.

### 4.4.3. T-test

With the aim of inspecting the significance of the dissimilarity between gender sets in their collective teacher efficacy, an independent samples *t*-test was employed. The outcomes are indicated in Table 10 In accordance with the chart below, the mean of the male set is more than the mean of the female set. As is illustrated, the mean of the teachers with more than 10 years of experience is more than the teachers with less than 10 years of experience.

**Table 10.** The Results of Independent Samples T-tests

| | Levene's Test for Equality of Variances | | | | | | |
|---|---|---|---|---|---|---|---|
| | F | Sig. | T | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference |
| **Equal variances assumed (Gender)** | .052 | .5 | -2.14 | 114 | .032 | -3.31 | .78 |
| **Equal variances not assumed (Gender)** | | | -2.11 | 107.32 | .032 | -3.31 | .78 |
| **Equal variances assumed (Experience)** | .049 | .5 | -4.23 | 87 | .012 | -4.45 | 1.47 |
| **Equal variances not assumed (Experience)** | | | -4.18 | 84.32 | .012 | -4.45 | 1.38 |

As Table 10 indicates, there exists a considerable dissimilarity in teachers' collective efficacy between men and women in favor of male participants: ($t$= -2.14, $p$<0.05). So, it is possible to state that the third null hypothesis is dismissed. For the purpose of exploring the significance of the dissimilarity among EFL teachers' withdrawal intention based on their teaching experience, an independent samples $t$-test was employed. Table 10 also demonstrates a considerable dissimilarity in teachers' withdrawal intentions between the two groups in favor of the first group (teachers with more than 10 years of experience): ($t$= -4.23, $p$<0.05). So, it is possible to state that the fourth null hypothesis is dismissed.

## 5. Discussion

It can be concluded that there was a noteworthy and reverse link between cumulative instructor efficacy and withdrawal intention among Iranian EFL teachers. Mattingly (2007) and Mawhinney et al. (2005) notified the same finding. Results indicated that student discipline as one of the subscales of collective teacher efficacy can predict withdrawal intention. However, instructional strategies which was the other subscale of collective efficacy could not predict withdrawal intention. According to Jensen et al. (2011), collective efficacy was supported to be a moderator between physical workload and withdrawal behavior. Moreover, collective efficacy has been supported to predict interpersonal behavior (Tasa et al., 2011), professional commitment (Ware & Kitsantas, 2007), and group performance (Carroll et al., 2005). Furthermore, withdrawal intention has been found to be predicted by job embeddedness (Mitchell et al., 2001), work environment (Breau & Rhéaume, 2014). And also cooperating in decision-making and sharing information in the workplace can predict intent to stay (Boyle et al., 1999).

Another objective of the present investigation was to examine the difference between collective efficacy in male and female educators. It was disclosed that male instructors are expected to have a higher perception of their potential to influence student achievement than female instructors. The results are similar to that of Skaalvik and Skaalvik (2007), supporting, that men had higher teacher efficacy than their female associates. However, Brennen et al. (1996) and Fives and Looney (2009) concluded that women have higher teacher efficacy than men. The reason for the opposing results could be the difference between the aspect of efficacy dealt with in each study and the occupational context of the studies. It could be speculated that female teachers can better identify with the role of teachers and perceive their potential to impose advancements (Fives & Looney, 2009). On the other hand, it seems that male educators are more equipped with team perceptions of potential enhancement. Another explanation could be that female educators feel more able to influence their environments in some contexts than others (Blau et al., 1998; Haydel, 1997; Wittmann, 1992).

In addition, the results showed that instructors who have spent more than 10 years instructing their students have higher withdrawal intentions than those instructors who have spent less time in this profession. The outcomes of this analysis did not corroborate those of Hill and Hirshberg (2013) and Ost and Schiman (2015), which reported that the turnover rate was greater among instructors who have taught for less than 3 years and more than 20 years in comparison to their coworkers. Another study by Nogueras (2006) on the withdrawal intention of nurses found that experience is a strong predictor of withdrawal intention. Additionally, more experienced employees exhibited lower withdrawal intent rates. In a study by Knani and Fournier (2013), it was reported that work experience positively influences withdrawal intention. The dissimilarity among

the findings could have several reasons such as the difference between the concepts of withdrawal determination and eventual physical withdrawal, the groupings of years of experience not being identical, and the behavior of the employees in those contexts. It can be assumed that there are contexts in which as the staff members gain experience, they become less inclined to leave their position as a result of various factors. There are also contexts in which more experienced job holders are more inclined toward leaving their position than staying at their job. Some of the factors that could come together to form a pattern of behavior of the staff members with regard to their added experience are decision latitude (Knani & Fournier, 2013), secondary traumatic stress (Christian-Brandt et al., 2020), and coping strategies (Wunnenberg, 2020).

According to Goddard et al. (2015), instructors are more inclined to involve in group work interactions with their peers provided that their leaders have some expertise in classroom practice. Several analyses have emphasized the part executed by communication and the involvement of teachers in constructing the prospective agenda of the school (Lambersky, 2014; Yahaya and Ebrahim, 2016). Since teachers with scarce group work experience tend to only conjecture on the cumulative competencies of themselves and their coworkers (Parker et al., 2006), individualistic contexts are more inclined to have discrepancies in collective efficacy estimates. Instruction is facilitated in an environment where instructors believe and behave in such a way that they can influence their students' achievement. Developing a culture of participation and cooperation among teachers would set a satisfying example for students' tendency toward teamwork. Teachers who are more inclined towards leaving the school are unlikely to try their best at educating students or enhancing their ambition. Other faculty members are also influenced by teachers. Working alongside colleagues who have their minds occupied with intentions to leave may not be uplifting or inspirational.

This present investigation achieves development of some compelling issues; however, a few limitations are essential to be mentioned with respect to the current analysis. The corona virous pandemic prevented feasible and face-to-face access to teachers. Another limitation of the researcher was the restricted available time. Moreover, the skill, knowledge, and experience of the researcher were limited as well. Conducting a mixed-method, or cross-national study would demand resources unavailable to the researcher. To obtain more precise results on the issue and also to extend our knowledge of teacher-related variables a qualitative approach would benefit the researcher in collecting additional data. Further research could also look into the same relationship with university professors and investigate the difference between their perceptions and their academic rank, annual income, or related and unrelated academic backgrounds.

## 6. Conclusion

This study made an effort to test the hypothesis that collective teacher efficacy subscales could predict variability in withdrawal intention and the relationship between CTE and WI. It has come to light through multiple regression analysis that student discipline which is one of the two subscales of collective teacher efficacy is able to predict withdrawal intention. The findings clearly indicate that the aggregate mentality of teachers regarding their power to make changes in the classroom for the benefit of the pupils would be in accordance with the determination of educators to commit to their profession.

This analysis has provided deeper insight into the factors leading instructors to decide to stay put rather than depart from their organization. In general, therefore, it seems that there exist both personal and organizational components involved in the issue. As it was reviewed in the literature of the field, organizational elements such as principals' organizational leadership, organizational justice, workplace ethics, and workplace environment influence the withdrawal intention of teachers. On the other hand, individual elements such as self-efficacy and burnout are associated with alterations in withdrawal intent. This fresh awareness enhances the chance of making proper predictions of the association between other organizational or individual components with the voluntary turnover of instructors.

The present study has supplied additional endorsement for the reverse association between withdrawal intention and withdrawal intention. This research has several practical applications. Firstly, it points to the importance of teacher mentality. How teachers perceive their capacity to ameliorate student performance will eventually give rise to that improved performance. Another closely related issue in regard to the job holders at the same company. Colleagues can mirror each other and the teacher mentality is able to be reshaped through the interactions among teachers and their reciprocation of wisdom gained from experience.

## 7. References

Abelson, M. A., & Baysinger, B. D. (1984). Optimal and dysfunctional turnover: Toward an organizational level model. *Academy of Management review*, *9*(2), 331-341.

Abzari, M., Kabiripour, V., & Saeidi, A. (2015). The effect of business ethical values on turnover intention: Mediating role of person-organization fit dimensions. The case of Iranian teachers in public school. *Academic Journal of Economic Studies*, *1*(1), 65-82.

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2018). *Introduction to research in education*. Cengage Learning.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of social and clinical psychology*, *4*(3), 359.

Bandura, A. (2002). Social cognitive theory in cultural context. *Applied psychology*, *51*(2), 269-290.

Blau, F. D., Simpson, P., & Anderson, D. (1998). Continuing progress? Trends in occupational segregation in the United States over the 1970s and 1980s. *Feminist economics*, *4*(3), 29-71.

Bluedorn, A. C. (1982). A unified model of turnover from organizations. *Human relations*, *35*(2), 135-153.

Boyle, D. K., Bott, M. J., Hansen, H. E., Woods, C. Q., & Taunton, R. L. (1999). Manager's leadership and critical care nurses' intent to stay. *American Journal of Critical Care*, *8*(6), 361.

Breau, M., & Rhéaume, A. (2014). The relationship between empowerment and work environment on job satisfaction, intent to leave, and quality of care among ICU nurses. *Dynamics*, *25*(3), 16-24.

Brennan, M. D., Robison, C. K., & Shaughnessy, M. F. (1996). Gender comparison of teachers' sense of efficacy. Psychological Reports, 78, 122

Carroll, J. M., Rosson, M. B., & Zhou, J. (2005, April). Collective efficacy as a measure of community. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1-10).

Christian-Brandt, A. S., Santacrose, D. E., & Barnett, M. L. (2020). In the trauma-informed care trenches: Teacher compassion satisfaction, secondary traumatic stress, burnout, and intent to leave education within underserved elementary schools. *Child Abuse & Neglect*, *110*, 104437.

Da'as, R., Qadach, M., Erdogan, U., Schwabsky, N., Schechter, C., & Tschannen-Moran, M. (2021). Collective teacher efficacy beliefs: testing measurement invariance using alignment optimization among four cultures. *Journal of educational administration*, *60*(2), 167-187.

Da'as, R., Watted, A., & Barak, M. (2020). Teacher's Withdrawal Behavior: Examining the Impact of Principals' Innovative Behavior and Climate of Organizational Learning. *International Journal of Educational Management*, *34*(8), 1339-1355.

Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted*, *37*(2), 111-127.

Donohoo, J. (2018). Collective teacher efficacy research: Productive patterns of behaviour and other positive consequences. *Journal of educational change*, *19*(3), 323-345.

Fives, H., & Looney, L. (2009). College Instructors' Sense of Teaching and Collective Efficacy. *International journal of Teaching and Learning in Higher education*, *20*(2), 182-191.

Glassman, M., Kuznetcova, I., Peri, J., & Kim, Y. (2021). Cohesion, collaboration and the struggle of creating online learning communities: Development and validation of an online collective efficacy scale. *Computers and Education Open*, *2*, 100031.

Goddard, R., Goddard, Y., Sook Kim, E., & Miller, R. (2015). A theoretical and empirical analysis of the roles of instructional leadership, teacher collaboration, and collective efficacy beliefs in support of student learning. *American journal of education*, *121*(4), 501-530.

Hattie, J. (2016). *Third annual visible learning conference (subtitled Mindframes and Maximizers)* Washington, DC, July,

Haydel, J. B. (1997). *The relationships between evaluative culture, teacher efficacy, and student efficacy*. University of New Orleans.

Hill, A., & Hirshberg, D. (2013). Alaska Teacher Turnover, Supply, and Demand: 2013 Highlights.

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American educational research journal*, *38*(3), 499-534.

Jensen, J. N., Holten, A. L., Karpatschof, B., & Albertsen, K. (2011). Does collective efficacy moderate the associations between physical work load and intention to leave or sickness absence? *Journal of advanced nursing*, *67*(11), 2425-2434.

Kaplan, S., Bradley, J. C., Luchman, J. N., & Haynes, D. (2009). On the role of positive and negative affectivity in job performance: a meta-analytic investigation. *Journal of applied psychology*, *94*(1), 162.

Knani, M., & Fournier, P.-S. (2013). Burnout, Job Characteristics, and Intent to Leave: Does Work Experience Have Any Effect. *Journal of Emerging Trends in Economics and Management Sciences (JETEMS)*, *4*(4), 403-408.

Lambersky, J. J. (2014). *Understanding the human side of school leadership: Improving teacher morale, efficacy, motivation, and commitment*. University of Toronto (Canada).

Lee, T. W., & Mitchell, T. R. (1994). An alternative approach: The unfolding model of voluntary employee turnover. *Academy of Management review*, *19*(1), 51-89.

Loughland, T., & Ryan, M. (2022). Beyond the measures: the antecedents of teacher collective efficacy in professional learning. *Professional Development in Education*, *48*(2), 343-352.

Maertz Jr, C. P., & Campion, M. A. (1998). 25 years of voluntary turnover research: A Review. *Management Journal,(35)*, 1036-1054.

Mattingly, J. W. (2007). *A study of relationships of school climate, school culture, teacher efficacy, collective efficacy, teacher job satisfaction and intent to turnover in the context of year-round education calendars*. University of Louisville.

Mawhinney, H., Haas, J., & Wood, C. (2005). Teachers' collective efficacy beliefs in professional learning communities. *Leading and Managing*, *11*(2), 12-45.

Meyer, A., Richter, D., & Hartung-Beck, V. (2022). The relationship between principal leadership and teacher collaboration: Investigating the mediating effect of teachers' collective efficacy. *Educational Management Administration & Leadership*, *50*(4), 593-612.

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of management journal*, *44*(6), 1102-1121.

Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of applied psychology*, *62*(2), 237.

Nogueras, D. J. (2006). Occupational commitment, education, and experience as a predictor of intent to leave the nursing profession. *Nursing economics*, *24*(2), 86.

Ost, B., & Schiman, J. C. (2015). Grade-specific experience, grade reassignments, and teacher turnover. *Economics of Education Review*, *46*, 112-126.

Parker, K., Hannah, E., & Topping, K. J. (2006). Collective teacher efficacy, pupil attainment and socio-economic status in primary school®. *Improving Schools*, *9*(2), 111-129.

Porter, L. W., & Steers, R. M. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. *Psychological bulletin*, *80*(2), 151.

Qadach, M., Schechter, C., & Da'as, R. a. (2020). Instructional leadership and teachers' intent to leave: The mediating role of collective teacher efficacy and shared vision. *Educational Management Administration & Leadership*, *48*(4), 617-634.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, *80*(1), 1.

Schechter, C., & Qadach, M. (2020). From principal cognitive complexity to teacher intent to leave: Exploring the mediating role of school absorptive capacity and teacher commitment. *Journal of educational administration*.

Shapira-Lishchinsky, O. (2012). Teachers' withdrawal behaviors: integrating theory and findings. *Journal of educational administration*.

Shapira-Lishchinsky, O., & Rosenblatt, Z. (2009). Organizational ethics and teachers' intent to leave: An integrative approach. *Educational Administration Quarterly*, *45*(5), 725-758.

Shin, Y., Hur, W.-M., Park, K., & Hwang, H. (2020). How managers' job crafting reduces turnover intention: The mediating roles of role ambiguity and emotional exhaustion. *International journal of environmental research and public health*, *17*(11), 3972.

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of educational psychology*, *99*(3), 611.

Tasa, K., Sears, G. J., & Schat, A. C. (2011). Personality and teamwork behavior in context: The cross-level moderating role of collective efficacy. *Journal of Organizational Behavior*, *32*(1), 65-85.

Tella, A. (2008). Teacher variables as predictors of academic achievement of primary school pupils' mathematics. *International Electronic Journal of Elementary Education*, *1*(1), 16-33.

Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and policy in schools*, *3*(3), 189-209.

Ware, H., & Kitsantas, A. (2007). Teacher and collective efficacy beliefs as predictors of professional commitment. *The journal of educational research*, *100*(5), 303-310.

Wittmann, L. Z. (1992). The new woman as a European phenomenon. *Neohelicon*, *19*(2), 49-68.

Wunnenberg, M. (2020). Psychosocial bullying among nurse educators: exploring coping strategies and intent to leave. *Journal of Nursing Scholarship*, *52*(5), 574-582.

Yahaya, R., & Ebrahim, F. (2016). Leadership styles and organizational commitment: literature review. *Journal of management development*.

## 8. Appendices

**Appendix A:** Collective Teacher beliefs scale developed by Tschannen-Moran and Barr (2004)

This questionnaire has been developed to collect research data with the purpose of completing an M.A. thesis in TEFL. We ask all Iranian EFL teachers working in either: **public schools**, **private institutions**, as well as **freelancers working in teams** to fill out this questionnaire. Thank you for your time and patience.

A- How old are you? .....

B- How would you describe your gender? Male ☐ Female☐

C- What is your latest educational degree? Diploma ☐ Associate ☐

Bachelor's degree☐ Master's degree ☐ Doctor of Philosophy ☐

D- How many years have you been teaching? .....

*Sub-scale1: Instructional Strategies*

1- How much can teachers in your school do to produce meaningful student learning?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

2- How much can your school do to get students to believe they can do well in schoolwork?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

3- To what extent can teachers in your school make expectations clear about appropriate student behavior?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

4- To what extent can school personnel in your school establish rules and procedures that facilitate learning?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

5- How much can teachers in your school do to help student's master complex content?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

6- How much can teachers in your school do to promote deep understanding of academic concepts?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

*Sub-scale2: Student Discipline*

7- How well can teachers in your school respond to defiant students?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

8- How much can school personnel in your school do to control disruptive behavior?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

9- How much can teachers in your school do to help students think critically?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

10- How well can adults in your school get students to follow school rules?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

11- How much can your school do to foster student creativity?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

12- How much can your school do to help students feel safe while they are at school?

Nothing ☐ very little☐ some degree☐ quite a bit☐ a great deal☐

**Appendix B: "Withdrawal Intention" developed by Shapira Lishchinsky (2012)**

**Please answer these questions with the option that best describes your response.**

1- I'm considering contacting professional bodies about other work options.

Totally disagree☐ disagree☐ neutral☐ agree☐ totally agree☐

2- Lately I'm looking for work options in other places.

Totally disagree☐ disagree☐ neutral☐ agree☐ totally agree☐

3- I plan on leaving school.

Totally disagree☐ disagree☐ neutral☐ agree☐ totally agree☐

4- I've begun inquiring among friends/acquaintances about other work options.

Totally disagree☐ disagree☐ neutral☐ agree☐ totally agree☐

5- I'm thinking about leaving my job.

Totally disagree☐ disagree☐ neutral☐ agree☐ totally agree☐

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 129-141*

Applied Linguistics Inquiry

https://doi.org/10.22077/ALI.2023.6268.1013

# Effects of Corrective Feedback on Iranian EFL Learners' Uptake and Attention

Sara Ziaei[1]
Azizollah Dabaghi[2]

[1]Ph.D. Candidate, Department of Foreign Languages, University of Isfahan, Isfahan, Iran
[2]Associate Professor, Department of Foreign Languages, University of Isfahan, Isfahan, Iran

**ARTICLE HISTORY**

Received: 27 August 2022
Revised: 27 October 2022
Accepted: 24 December 2022
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: s.ziaei93@yahoo.com

**ABSTRACT**

The purpose of the present research was to investigate if the type of feedback significantly affected EFL learners' uptake. In addition, it aimed to search if feedback could lead to more learners' attention in EFL classes. The participants were 60 EFL learners. The data were obtained from classroom observations in addition to running a test. In order to investigate the effects of teachers' CF on learners' uptake, the current study employed Lyster and Mori's (2006) taxonomy of CF types and learners' immediate uptake moves. All episodes which included learner errors were identified by the researchers. Then, instances with teacher feedback were extracted. In order to investigate the amount of uptake by the participants, an uptake sheet was utilized. The uptake sheet was facile and easy to understand. Corrective feedback and uptake were coded and tabulated. It was shown that recasts were the most highly frequent type of feedback followed by explicit correction and prompts. In addition, the frequency of uptakes followed by recast was higher than the uptakes followed by explicit correction and prompts. The results of ANOVA revealed a significant difference among various amounts of uptake followed by different types of feedback ($p > .034$).

**KEYWORDS:** Corrective feedback; Explicit explanation; Prompt; Recast; Uptake

## 1. Introduction

As indicated by Nassaji (2016), uptake is the immediate student answers that happen after feedback. Chaudron (1977) was one of the main specialists who emphasized the significance of student reactions following input. Lyster and Ranta (1997) later utilized the term uptake to discuss such reactions. Lyster and Ranta (1997) also stated that uptake is any reaction following input, the reactions might go from those that changed the student's original output to expressions which just admitted the receipt of the criticism. In addition, Lyster and Ranta (1997) categorized uptake considering learner repair and stated that uptake shows that learners attempt to accomplish something through feedback.

The idea of uptake has been explored in various studies. Nonetheless, there are contrasts among studies as to what sort of student reactions ought to be considered as uptake. According to some scholars (e.g., Nassaji, 2011a), uptake isn't just a student reaction, yet it is an endeavor including some level of change of the output following feedback. This comprehension of uptake is by all accounts more significant in that the term uptake, connotes that students have to some degree benefitted from the input, though reactions of affirmation don't really do as such.

Cite this article: Ziaei, S., & Dabaghi, A. (2023). Effects of corrective feedback on iranian EFL learner's uptake and attention. *Applied Linguistics Inquiry*, *1*(1), 129-141. doi: 10.22077/ali.2023.6268.1013

Uptake has been explored in various studies about interactional feedback. A few contrasts exist among these investigations about the sort of student reactions which ought to be considered as uptake. A few researchers accept that uptake isn't a straightforward reaction of students, yet it includes some level of changes of the first result following criticism (Nassaji, 2011b). Effective uptake conducts students to see the error between their non-target-like output and their target-like production (Egi, 2010), which acts as an impetus for their interlanguage change (Gass, 1997). Real uptake incorporates students' alteration of their structures which are mistaken. In this way, uptake is distinguished as a solid indicator of learning as shown by the discoveries of common posttests (e.g., Loewen, 2005; Nassaji, 2011a).

Oral corrective feedback (CF) has been defined as providing information “about the success of their [learners’] utterances and giving additional opportunities to focus on production and comprehension” (Gass & Selinker, 2008, p. 329). Feedback in language classes promotes students' awareness. "Teacher can use feedback in a motivating way to reduce the gap between the student’s understandings and also how the teacher wants him to develop (Carvaldho, et al., 2014, p.170). Gower and Phillips (2004) claimed that "the aim of feedback is to bring improvement and raise self-awareness" (as cited in Rahman, 2012, p.9). According to Ur (1995) "feedback has two main components, error correction and assessment" (as cited in Akter 2010, p. 9).

According to Ellis et al. (2006), "Corrective feedback takes the form of responses to learner utterances that contain error" (p. 340). Mendez and Cruz (2012) believed that the responses can range from "(a) an indication that an error has been committed, (b) provision of the correct target language form, or (c) information about the nature of the error, or any combination of these" (p. 64). Feedback has a role in most of the theories of second language learning and language pedagogy. In both behaviorism and cognitive theories of language learning, feedback is attributed to language learning. And also it is seen as a means of increasing student motivation in both structural and communicative language teaching (Ellis, 2009).

Feedback is an important part of an instructional model. Reigeluth (1999) contends that feedback is a method of instruction that can nurture cognitive learning. Reigeluth (1999) gives an example of instructional design theory called "Theory One" that was defined by Perkins (1992) and explains that teaching should include informative feedback as well as other methods like clear information, and strong motivation. According to Merrill (1994), feedback has a crucial position within an instructional design theory.

## 2. Literature review

Merril's (1994) Component Display Theory considers feedback as the most important part of Secondary Presentation Forms which are "information added to the Primary Presentation Forms to enhance the learning that occurs" (p. 150) in an instructional activity. According to this theory, feedback may happen during practice and/or explanation stages. According to Merrill (2002) feedback is the most essential form of learner guidance. To show the important position of feedback, Andrews and Goodson (1980) claimed that feedback is included in one of the purposes of instructional design that is to improve evaluation process "by means of the designated components and sequence of events, including feedback and revision events, inherent in models of systematic instructional design" (p. 4).

The instructional designs have been affected by various learning theories like behavioral learning theory and cognitive information processing theory of instruction. All of these theories consider feedback as an important part of learning and instruction. Driscoll (2002) believes that these theories of learning and instruction have provided a strong foundation for current practices of instructional design; for example, concepts such as reinforcement and feedback as presented by Skinner's behavioral learning theory introduced in Skinner’s (1958) Teaching Machines. According to the behavioral view of learning, reinforcement and feedback can have important instructional effects on student learning as they can shape learner behavior by reinforcing correct responses or providing corrective feedback for incorrect responses. In behavioral learning theory, a learner's behavior is checked before and after an instruction. An instruction is not effective if the expected behavioral changes that are related to instruction do not happen. According to Driscoll (2007) "these observations (providing feedback) are part of formative evaluation, which is used to collect information about whether instruction has resulted in learning and how it might be improved to result in better learner performances" (p. 38). According to Lockee, et al. (2007), the idea of "reinforcement through evaluation and feedback" (p. 192) in systematic instructional design is based on Skinnerian theory.

According to cognitive information processing theory, feedback has a significant value in instructional design. According to Information processing theory learning is an internal process within the learner (Driscoll, 2007), where the learner processes input from the environment to become necessary output as a result of learning. Driscoll (2007) believed that feedback serves two functions in learning process. First, feedback provides learners with information about the correctness of their responses or performance. Second, feedback provides corrective information that can be used by the learners to modify their performance.

Feedback is also an integral part of Gagne's systematic instructional design model. Gagne's (1985) model of instructional design includes gaining attention, informing the learner of the objective, stimulating recall of prerequisite learning, presenting the stimulus material, providing learning guidance, providing feedback, assessing performance, and enhancing

retention and transfer. Gagne et al. (1992) confirm the important function of feedback in an instructional program and emphasize that one important characteristic of feedback is its function, which is to provide information to learners relating to the correctness of their performance.

As the present study aimed to investigate the effect of feedback on EFL learners' uptake, the following section is devoted to uptake. A number of studies have so far been conducted on uptake (Mackey et al., 2000; Oliver 2000; Ellis et al., 2001a; Panova & Lyster 2002; McDonough 2005). Some studies have shown that learner uptake plays a positive role in learners' second language development (McDonough, 2005; Loewen, 2005). On the other hand, some others are not sure about whether uptake leads to long-term learning (e.g. Nabei & Swain 2002; Morris & Tarone 2003). It is argued that "learners' uptake may not fully represent their cognitive processing of the feedback" (Nabei & Swain 2002, p. 45). According to Morris and Tarone (2003) "uptake, in the form of recast repetition, may not be a reliable indicator of acquisition" (p. 328). However, Nabei and Swain (2002) claim that a "learner's immediate response after recast feedback might not be appropriate evidence for evaluating its effect [i.e. for assuming that acquisition has taken place]" (p. 45). Still, some studies suggest that uptake may have a longer-term effect. As a case in point, Iwashita (2003) showed that general learner accuracy increased significantly as the number of recasts was provided during a treatment period, and the accuracy rate was constant.

According to McDonough (2005) uptake "may contribute to target language development by strengthening knowledge representation that learners already have stored and by encouraging automatic retrieval of linguistic forms" (p. 83). McDonough (2005) found that when learners produced more complex forms in uptake or modified output, they were more likely to produce these forms in later utterances, again suggesting sustained impact. Similarly, Loewen (2005) observed that successful uptake is considered a significant predictor of correct test scores. This is still another clue that supports the assumption that uptake leads to better language development.

In a recent study, Gholami and Gholami (2018) investigated the degree to which incidental focus-on-form episodes (FFEs) with formulaic sequences occur, and the extent to which they result in uptake in EFL classes. To gather the data, 36 hours of communicative classroom interactions from three advanced EFL classes were audio-recorded. It was shown that learners and teachers regularly draw their attention to formulaic sequences. FFEs with formulaic focus as well as collocations led to more uptake than those with other linguistic foci. Finally, among types of incidental FonF, the highest rate of (successful) uptake was observed in student-initiated FFEs with a formulaic focus.

In a study, Ellis et al. (2001a) explored student uptake in the center around structure episodes in an informative ESL instructing class. The results revealed that student uptake was higher than whatever was mentioned about submersion classrooms. In a new report, Gholami and Gholami (2018) explored how much incidental focus on form episodes (FFEs) with formulaic sequences occur, and the degree to which they bring about uptake in EFL classes. To gather the data, 36 hours of communicative classroom interactions from three progressed EFL classes were sound recorded. It was shown that students and instructors usually pay attention to formulaic sequences. FFEs with conventional concentration just as collocations prompted more uptake than those with other semantic foci. Finally, among kinds of incidental FonF, the most noteworthy pace of (effective) uptake was seen in student-initiated FFEs with standard concentration.

Gholami et al. (2017) examined three classifications of formulaic sequences (FSs), in particular, collocations, lexical packs, and figures of speech in incidental focus on form (F on F). 30 hours of recorded interaction between the instructor and the class in three EFL classes were analyzed. A sum of 1102 occurrences of FFEs were recognized, 38% of which were of formulaic sequences. The outcomes showed more non-formulaic FFEs than formulaic ones. Preemptive formulaic FFEs and student-initiated episodes were utilized more as often as possible than other kinds. Among the FSs, collocation was the category with the highest frequency.

Gholami and Bassirian (2011) researched learner- and teacher-initiated focus on form episodes (FFEs), just like the following rate of uptake moves. To do this, 18 meetings of a moderate EFL class were noticed, sound recorded, and examined. A think-aloud protocol incidental on structure sheet was used to assemble occurrences of student-and-teacher-produced FFES. The results showed that teacher-initiated FFEs were higher and more effective. In dealing with the existing gap, the present descriptive study aimed to investigate the effects of various types of feedback on EFL learners' uptake using an instrument called 'uptake sheet', along with the traditional procedure of measuring uptake through audio-recorded data, so as to gain a clearer picture of uptake by cross-checking the results from both procedures. Another purpose of this study was to measure the level to which the feedback leads to attention. A review of the existing literature during recent decades shows that corrective feedback improves linguistic knowledge. On the other hand, not much research has so far investigated the interplay between uptake and corrective feedback; therefore, the present study intended to compare the uptake and amount of attention after feedback in EFL classes. Therefore, the following research questions were addressed:

Q1: Does the type of feedback significantly affect EFL learners' uptake?

Q2: Does feedback lead to more learners' attention in EFL classes?

# 3. Method

## 3.1. Data source

The data were from a study conducted at a language institute in Isfahan, Iran. The data were obtained from classroom observations of 9 classes. The observation included 80 h of classroom observations. To make sure that the data were manageable. of data, only four classes were selected, with two from each data set. Accordingly, the data for the current study consisted of 18 h of recordings from each setting, totaling 36 h. The participants all were at intermediate proficiency level according to a placement test. The test at the beginning of the study was a proficiency test which was an adapted version of the proficiency test, with different test items but the same question types and scoring rubrics.

In addition to class observation, an online test (test of EFL learners' self-correction) was run on Big Blue Button platform. The test was an open-ended one in a structure similar to semi-structured interviews. Learners answered each question and wrote down what they do to rectify themselves in such situations.

## 3.2. Instruments

### 3.2.1. Uptake sheet

Uptake sheet was first introduced as a method of data collection by Allwright (1984) on learners' perceptions about what they learned in their language classes. Allwright (1984) gathered learners' reports about their learning which he called 'uptake'. According to Mackey and Gass (2005), in classroom research, through uptake, the learners are asked to mark or note things on which the researcher or teacher is focusing. A copy of the uptake sheet is put in Appendix A.

### 3.2.2. The test of attentional and interpersonal style (TAIS) (Greher, 2000)

The TAIS has 144 items and is comprised of 17 sub-scales, nine of which attempt to evaluate how one might behave during certain interpersonal situations. The other six sub-scales attempt to evaluate attentional processes, while two others attempt to evaluate behavioral and cognitive control. The TAIS is a measure, designed using a rational-intuitive approach toward test construction.

The present study implemented the short form of the TAIS (see Appendix B) because of the complicated nature of the standard version and its 17 subscales. This version of the TAIS is comprised of 12 questions concerning attentional skills, each of which is rated based on the following scale: 0 = never, 1 = rarely, 2 = sometimes, 3 = 18 frequently, and 4 = all the time. High numeric value responses indicate effective attention, while low numeric value responses indicate ineffective attention.

The twelve items constitute six subscales, with two items per scale. These six subscales are identical to those intended to assess attention in the full-length version of the TAIS and include (1) the Broad-External subscale (BET), (2) the Broad-Internal subscale (BIT), (3) the Narrow Effective Focus subscale (NAR), (4) the External Overload subscale (OET), (5) the Internal Overload subscale (OIT), and (6) the Errors of Under inclusion subscale (RED).

### 3.2.3. Test of EFL learner's self-correction

The test was an open-ended one which was answered by learners via Big Blue Button platform. The items asked learners what they would do in the situations provided. They answered the questions and revealed themselves. The answerers were scrutinized by the researcher and categorized into themes similar to FB types namely surface correction, deep correction, reformulation,

## 3.3. Participants

The participants of the present study included a number of 60 EFL learners who were within the age range of 16 to 24, with 30 learners in each class. The classes met twice a week, and each meeting lasted 1.5 h. The teacher was the same person who taught the three classes. The students had an average age of 19.1, all with Farsi as their mother tongue and English as a foreign language. None had ever visited an English-speaking country.

## 3.4. Data collection

The data for the present thesis were obtained from classroom observations. During the observations, the lessons were recorded and the author took notes. In order to obtain the data, classroom interactions were both audio-recorded using an MP3 recorder.

The MP3 recorder was put in a position so that it was convenient for the teacher. The focus was on teacher feedback and learner uptake with the same coding scheme.

### 3.5. Data coding scheme

In order to investigate the effects of teachers' CF on learners' uptake, the current study employed Lyster and Mori's (2006) taxonomy of CF types and learners' immediate uptake moves. All episodes which included learner errors were identified by the researcher. Then, instances with teacher feedback were extracted. An episode containing feedback was a sequence in which a learner made a mistake, then the teacher gave feedback, and it ended with the learner's reaction (adapted from Lyster & Mori, 2006).

In order to investigate the amount of uptake by the participants, an uptake sheet was utilized. The uptake sheet used in this research was the one employed by Gholami and Bassirian (2010). A sample of the used uptake sheet is put in the appendix. In the present research, the teacher's interactions with the whole class were audio recorded, and the learner-learner interactions were not recorded.

The uptake sheets were distributed among the learners at the beginning of every session and they were gathered at the end of each session. This was done to make sure of immediate, on-the-site uptake moves. In addition, the learners were taught to write only those language forms which they observed in the class, no matter if they were presented by the teacher. The learners were also taught how to fill in the uptake sheet. CF and uptake were coded and tabulated. CF was coded for its type, linguistic focus and emphasis. Based on Lyster and Mori (2006), it was coded into three major types: recasts, prompts, and explicit correction. Recasts refer to teachers' reformulation of learners' problematic utterances in part or full without changing their meanings. Prompts refer to techniques teachers use to elicit the target form from learners, such as clarification requests and linguistic cues. Explicit correction consists of two elements: a message alerting the learner to the presence of an error plus a provision of the correct form. Having completed the data collection phase, the audio-recorded data were transcribed.

Based on Loewen (2004), the linguistic focus of CF was coded into four categories: grammar (e.g. incorrect tense), vocabulary (e.g. inappropriate word choice), pronunciation (e.g. inaccurate stress), and other (including pragmatic and spelling errors).

Results of the EFL test of self-correction were categorized in themes mentioned in the previous section. They were subsequently compared with CF types.

## 4. Results

The first research question of the present research intended to investigate whether the type of feedback significantly would affect EFL learners' uptake. The results regarding this research question are presented in the following section. In so doing, first, the frequency of all types of feedback is presented.

**Table 1.** Distribution of CF by Type

| Feedback type | Recast | Prompts | Explicit correction | Total |
|---|---|---|---|---|
| Frequency | 135 | 31 | 89 | 255 |
| Percentage | 52.95 | 12.15 | 34.90 | 100% |

Totally, several 255 feedback types were identified in the corpus drawn from the observation of classes from which 135 (52.95%) were recasts, 89 (34.9%) were explicit corrections and 31 (12.15%) were prompts.

**Table 2.** Uptake Following each Type of Feedback obtained from Uptake Sheets

| Feedback Type | Recast | Prompts | Explicit Correction | Total |
|---|---|---|---|---|
| Uptake(Frequency) | 113 | 21 | 45 | 179 |
| Uptake(Percentage) | 63.12% | 11.73% | 25.13 | 100% |

Table 2 shows the frequency of uptakes followed by different types of feedback. As shown in Table 2, the frequency of uptakes followed by recast was 113 which counted for 63.12% of the total number of uptakes. Furthermore, a percentage of 25.13% of the uptakes were followed by explicit correction. Finally, a frequency of 21 uptakes (11.73%) was observed after

prompts. In order to investigate whether the type of feedback significantly affects the amount of uptake, an ANOVA was employed.

**Table 3.** ANOVA

| | frequency |
|---|---|
| Chi-Square | 2.000 |
| df | 2 |
| Asymp. Sig. | .034 |

As shown in Table 3, a significant difference was observed among various amounts of uptake followed by different types of feedback (p >.034). In order to have a clearer picture of the issue under study, the number of uptakes in different language skills and subskills were compared and contrasted.

**Table 4.** Uptake in Different Language Skills and Subskills

| **Feedback Type** | **Grammar** | **Pronunciation** | **Vocabulary** | **Other** | **Total** |
|---|---|---|---|---|---|
| Uptake(Frequency) | 83 | 45 | 41 | 10 | 179 |
| Uptake(Percentage) | 46.36% | 25.13% | 22.90% | 5.58% | 100% |

As shown in Table 3, grammar included the highest percentage of uptake in the corpus with a percentage of 46.36% of the corpus. The next type of uptake was uptake in pronunciation which accounted for 25.13% of the corpus. The last type of language skill which followed feedback was uptake in vocabulary knowledge which accounted for 22.9% of the corpus. The rest 5.58% of the corpus included other types of language skills.

**Table 5.** ANOVA Statistics

| | frequency |
|---|---|
| Chi-Square | 1.953 |
| df | 3 |
| Asymp. Sig. | .044 |

As presented in Table 4, a significant difference was observed among the various subcategories of language skills as the effects of different types of feedback.

The second question of the present study aimed to investigate if feedback leads to more learners' attention in EFL classes. In so doing, the Test of Attentional and Interpersonal Style was employed. Then, the scores obtained from the questionnaire were compared and contrasted. Table 6.

**Table 6.** One-way ANOVA on the Attention Scores Based on Feedback Type

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 917.200 | 2 | 458.600 | 5.052 | .010 |
| Within Groups | 5174.200 | 57 | 90.775 | | |
| Total | 6091.400 | 59 | | | |

Table 6 gives both between-groups and within-groups sums of squares, degrees of freedom, F value, etc. The significant value is smaller than .05 (.010< .05), so there is a significant difference somewhere among the attention level of EFL learners due to various types of feedback.

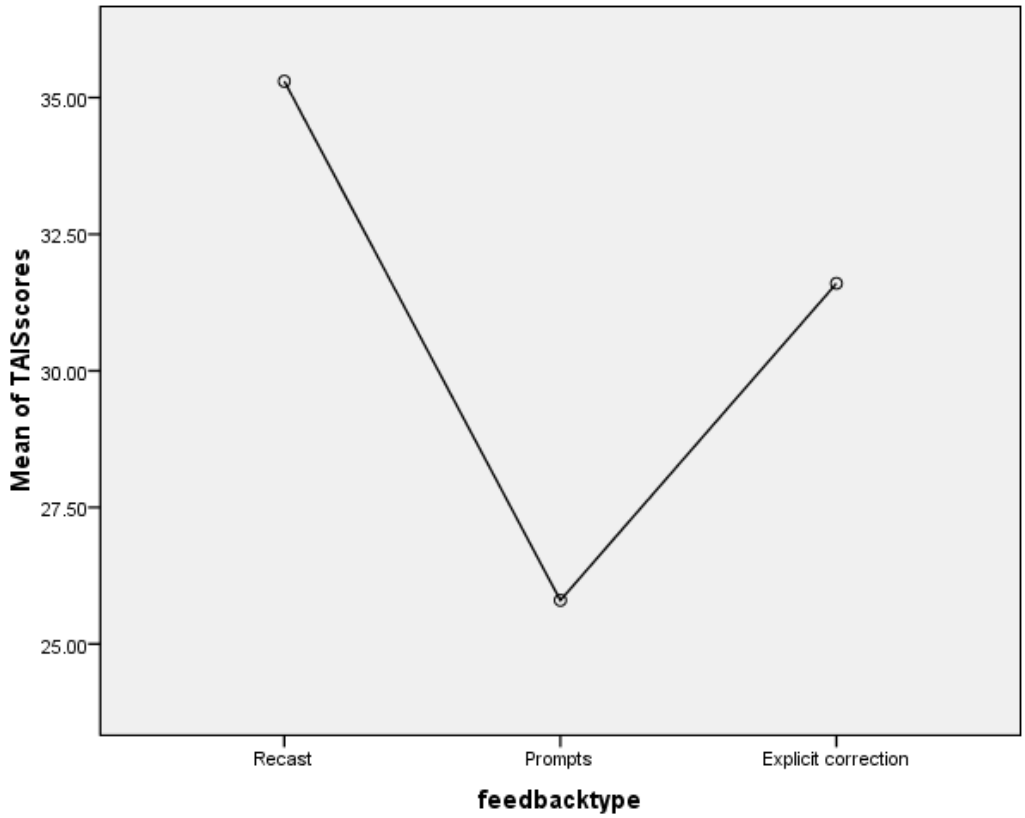



**Figure 1.** Mean Comparison of Attention Level Based on Type of Feedback

Figure 1 compares the attention level of the participants based on three types of feedback. The mean difference among the attention level based on the three types of feedback is clear. It is difficult, however, at this point to tell if this significant difference occurred between which type of feedback, since an ANOVA provides information on whether or not these groups differ, but it provides no information as to the location or the source of the difference. Having received a statistically significant difference, we can now look at the results of the post-hoc test provided in Table 7 to be able to locate the source of significance in our data.

**Table 7.** Post-hoc Scheffe Test (Post-test, Multiple Comparisons)

| | | | | | | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | **(I) feedback type** | **(J) feedback type** | **Mean Difference (I-J)** | **Std. Error** | **Sig.** | **Lower Bound** | **Upper Bound** |
| Tukey HSD | Recast | Prompts | 9.50000* | 3.01290 | .007 | 2.2497 | 16.7503 |
| | | Explicit correction | 3.70000 | 3.01290 | .442 | -3.5503 | 10.9503 |
| | Prompts | Recast | -9.50000* | 3.01290 | .007 | -16.7503 | -2.2497 |
| | | Explicit correction | -5.80000 | 3.01290 | .141 | -13.0503 | 1.4503 |
| | Explicit correction | Recast | -3.70000 | 3.01290 | .442 | -10.9503 | 3.5503 |
| | | Prompts | 5.80000 | 3.01290 | .141 | -1.4503 | 13.0503 |
| LSD | Recast | Prompts | 9.50000* | 3.01290 | .003 | 3.4668 | 15.5332 |
| | | Explicit correction | 3.70000 | 3.01290 | .224 | -2.3332 | 9.7332 |
| | Prompts | Recast | -9.50000* | 3.01290 | .003 | -15.5332 | -3.4668 |
| | | Explicit correction | -5.80000 | 3.01290 | .059 | -11.8332 | .2332 |
| | Explicit correction | Recast | -3.70000 | 3.01290 | .224 | -9.7332 | 2.3332 |
| | | Prompts | 5.80000 | 3.01290 | .059 | -.2332 | 11.8332 |

*. The mean difference is significant at the 0.05 level.

As presented in Table 7, there are some asterisks besides some values in a column titled mean differences. As the caption of the table shows, the asterisks determine that the values are significant. The only significant difference exists between EFL learners' attention due to recasts and prompts. This means that the attention level of EFL learners due to recast was significantly higher than their attention level due to prompts and explicit correction; however, there was no significant difference between the attention level of EFL learners due to prompts and explicit correction.

**Table 8.** The results of the self-correction test analysis

| **Feedback type** | **Recast** | **Prompts** | **Explicit correction** | **Total** |
|---|---|---|---|---|
| Frequency | 135 | 31 | 89 | 255 |
| Percentage | 52.95 | 12.15 | 34.90 | 100% |

## 5. Discussion

The purpose of the present research was to investigate if the type of feedback significantly affected EFL learners' uptake. In addition, it aimed to search if feedback could lead to more learners' attention in EFL classes. It was shown that recasts were the most highly frequent type of feedback followed by explicit correction and prompts. In addition, the frequency of uptakes followed by recast was higher than the uptakes followed by explicit correction and prompts.

The results of a *Kruskal Wallis* Test revealed a significant difference among various amounts of uptake followed by different types of feedback ($p > .034$). Further analysis of the results revealed that uptake level which was observed in grammar significantly more than in pronunciation and vocabulary. The results of an ANOVA also showed a significant difference in EFL learners' attention levels of EFL learners due to various types of feedback. The findings of the present study can be justified in light of Li (2020), and Wang and Loewen (2016) who noted that the utility of CF is primarily grounded in the interaction hypothesis and the noticing hypothesis. According to the interaction hypothesis (Gass, 1997; Long, 1996), feedback (especially recasts) embedded in negotiated interaction prompts the learner to notice the gap between the target form and the erroneous utterance and to make adjustments to his/her interlanguage. According to the noticing hypothesis (Schmidt, 1990), which is

closely related to the interaction hypothesis, language learning is mainly a conscious process and that only input that has been noticed can be converted into intake. CF facilitates the noticing of linguistic forms, especially structures that are often not attended to by the learner.

The high frequency of recasts is in line with the findings of previous studies (e.g. Llinares & Lyster, 2014; Sheen, 2004). This can be attributed to the fact that EFL lessons focused on learners' speaking skills which emphasized fluency over accuracy, and this made implicit feedback such as recasts, a better choice for the learners. The EFL classes focused more on listening and speaking, and most group works were meaning-focused. In such activities, it was not suitable for the teacher to break the interaction permanently. Therefore, recast, which is usually unobtrusive and does not interrupt the flow of communication, as noted by Long (2007) and Lyster et al., (2013), was chosen as the teacher's preferred feedback strategy.

## 6. Implications

The attentional system is a complicated construct involving various influences related to the instruction and the learner. The results of the present study showed that various types of feedback do not let language learners notice a specific construct. In fact, the findings of the present study suggest that most students do not pay continuous attention after different types of feedback; however, there are specific factors which can be controlled by the teacher. Language teachers can control the type of feedback that they exert to their learners, and give for example, recast more often, as this type of feedback leads to more uptake and more attention level.

## 7. Conclusion

Uptake is a learner's reaction to the CF on his/her utterance. Researchers (e.g. Ellis et al., 2001) claim that uptake may lead to acquisition since it helps learners to automatize the linguistic form and integrates the target form in immediate L2 production. Other researchers (e.g. Long, 2007) believe that uptake is a discourse phenomenon and is not a guarantee for acquisition as it may not be produced by the learners. In addition, topic continuation sometimes provides no opportunity for learners to respond to CF. On the other hand, uptake has been researched as an important observable source for understanding the effectiveness of CF (e.g. Bao & Du, 2015).

## 8. Limitations and suggestions for further research

Similar to other studies, there are some limitations for this study. First of all, considering the sample size employed it should be known that the sample was limited to EFL students studying at a language institute, the results may be different with university students. The generalizability of the present study is also limited by the fact that there were more female participants than male participants. In addition, the participants fell within the age range of 16 to 24 years old. Finally, the data collection tools were limited to questionnaires and uptake sheets. Further research can consider the mentioned limitations and can make use of other instruments and employ other participant groups to investigate if similar results will be obtained.

## 9. References

Akter, L. (2010). *Teacher Talk Time in ESL classrooms in Bangladesh*. Unpublished doctoral dissertation. BRAC University, Bangladesh.

Allwright, R. (1984). The importance of interaction in classroom language learning. *Applied Linguistics, 5*, 156-171.

Andrews, S. & Goodsman, J. (1980). Just like instant noodles: L2 teacher and their beliefs about grammar pedagogy. *Teachers and Teaching: Theory and Practice, 9(4),* 351-375.

Bao, R., & Du, X. (2015). Implementation of task-based language teaching in Chinese as a foreign language: Benefits and challenges, *Language. Cult. Curric.*, *28*(3), 291-310.

Carvaldho, J.M., Santos, S., Conboy, J., & Martins, J. (2014). A qualitative exploration of the factors influencing the job satisfaction and career development of physiotherapists in private practice. *Manual Therapy, 25*, 56-61.

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, *27*, 29-46.

Driscoll, M. P. (2002). Psychological foundations of instructional design. In R. A. Reiser &J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (57-69). Upper Saddle River, NJ: Pearson Education Inc.

Driscoll, M. P. (2005). Psychology of learning for instruction (3rd ed.). Boston, MA: Allyn and Bacon.

Driscoll, M. P. (2007). Psychological foundations of instructional design. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and Technology* (2nd ed.), (36-44). Upper Saddle River, NJ: Pearson Education Inc.

Egi, T. (2010). Uptake, Modified Output, and Learner Perceptions of Recasts: Learner Responses as Language Awareness. *The Modern Language Journal, 94*(1), 1-21.

Ellis, R. (2008). A typology of written corrective feedback types. *ELT Journal*, 97-107.

Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal, 1*, 3-18.https://doi.org/10.5070/L2.V1I1.9054

Ellis, R., Basturkmen, H. & Loewen, S. (2001a). Preemptive focus on form in the ESL classroom. *TESOL Quarterly 35* (3), 407-432.

Ellis, R., Basturkmen, H. &. Loewen, S. (2001b). Learner uptake in communicative ESL lessons. *Language Learning 51*(2), 281–318.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*(2), 339-368.

Gagne, E. D. (1985). *The cognitive psychology of school learning.* Boston, MA: Little, Brown and Company.

Gass, S. M. (1997). *Input, interaction, and the second language learner.* Lawrence Erlbaum Associates Publishers.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition an introductory course*. New York: Routledge.

Gholami, J. & Bassirian, M. (2011). *Incidental focus on form and uptake: a new visage of learner- and teacher-initiated preemptive focus on form and uptake in EFL classes*. Berlin: Lambert Academic Publishing.

Gholami, L., & Gholami, J. (2018). Uptake in incidental focus-on-form episodes concerning formulaic language in advanced adult EFL classes. *Language Teaching Research, 24*(2), 189–219. https://doi.org/10.1177/ 1362168818817113

Gholami, L., Karimi, M. N., & Atai, M. R. (2017). Formulaic focus-on-form episodes in adult EFL communicative interactions. *System, 68*, 72–86. https://doi.org/10.1016/j.system.2017.06.015

Greher, M., R. (2000). *Measuring attention: An evaluation of the Search and Cancellation of Ascending Numbers (SCAN) and the short form of the Test of Attentional and Interpersonal Style* (TAIS) [Unpublished doctoral dissertation/master's thesis]. Denton, Texas. University of North Texas Libraries.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction. *Studies in Second Language Acquisition. 25* (3),1-36.

Le, T. (2018). Voice recording in second language outside the classroom: Process and product. *Journal of NELTA, 23* (1-2), 129-141.

Li, L. (2020). *Language Teacher Cognition: A Sociocultural Perspective*. London: Palgrave/Macmillan.

Llinares, A. & Lyster, R. (2014). The influence of context on patterns of corrective feedback and learner uptake: A comparison of CLIL and immersion classrooms. *The Language Learning Journal. 42*, 181-194. 10.1080/09571736.2014.889509.

Lockee, B. B., Larson, M. B., Burton, J. K., & Moore, D. M. (2007). Programmed technologies. In J. M. Spector, M. D. Merrill, J. V. Merrienboer., & M. P. Driscoll (Eds.), *Handbook of research for educational communications and technology* (187-197). New York, NY: Lawrence Erlbaum Associates.

Loewen, S. (2004a). Uptake in incidental focus on form in meaning focused ESL lessons. *Language Learning, 54*(1), 153-188.

Loewen, S. (2004b). The occurrence and characteristics of student initiated focus on form. In H. Reinders, H. Anderson, M. M. Hobbs, & J. Jones-Parry (Eds.), *Supporting independent learning in the 21st century. Proceedings of the inaugural*

*conference of the Independent Learning Association* (pp. 86-93). Auckland: Independent Learning Association Oceania. Retrieved from http://www.independentlearning.org /ila03/ila03_loewen%20.pdf

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27*(3), 361-386.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*(2), 126-141.

Long, M.H. (1991). Focus on form: a design feature in language teaching methodology. In *Foreign Language Research in Cross-cultural Perspective*, ed. K. de Bot, R. Ginsberg and C. Kramsch, 39–52. Amsterdam: John Benjamins.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In C. Ritchie, T. K. Bhatia (Eds.), *Handbook of language acquisition, vol. 2, second language acquisition* (pp. 413- 468). New York: Academic Press.

Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Laurence Erlbaum Associates.

Lyster, R. & Mori, H. (2006). Interactional Feedback and Instructional Counterbalance. *Studies in Second Language Acquisition, 28* (2), 269-300. 28. 10.1017/S0272263106060128.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37-66.

Lyster, R. (2004). Differential effects of prompts and recasts in form focused instruction. *Studies in Second Language Acquisition, 26*(3), 399-432.

Lyster, R. & Izquierdo, J. (2009). Prompts versus Recasts in Dyadic Interaction. *Language Learning*, *59*(2), 453-498.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates Publishers.

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22,* 471–497.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, *82*, 338–356.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL questions development. *Studies in Second Language Acquisition 27*(1), 79-103.

Mendez, H. & Cruz, R. (2012). Teachers' Perceptions About Oral Corrective Feedback and Their Practice in EFL Classrooms. Profile. *Issues in Teachers` Professional Development. 14*, 63-75.

Merrill, M.D. (1994). *Instructional Design Theory*. Englewood Cliffs, N.J. Educational Technology Publications.

Merrill, M.D. (2002). *First principles of instruction*. ETR&D 50, 43–59. https://doi.org/10.1007/BF02505024

Morris, F.A. & Tarone, E.E. (2003). Impact of classroom dynamics on the effectiveness of recasts in second language acquisition. *Language Learning. 53* (2), 325–368.

Nabei, T. & Swain. M. (2002). Learner awareness of recasts in classroom interaction: a case study of an adult EFL student's second language learning. *Language Awareness, 11*(1), 43-63.

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning, 59*(2), 411-452.

Nassaji, H. (2011a). Immediate learner repair and its relationship with learning targeted forms in dyadic interaction. *System, 39*(1), 17-29.

Nassaji, H. (2011b). Correcting students' written grammatical errors: The effects of negotiated versus non negotiated feedback. *Studies in Second Language Learning and Teaching*, *1*, 315–334.

Nassaji, H. (2016). Research timeline: Form-focused instruction and second language acquisition. *Language Teaching Research*, *49*, 35–62.

Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pair work. *Language Learning*, *50*(1), 119-151.

Panova, I. & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly 36* (4), 573–595.

Perkins, D. N. (1992). *Smart schools: Better thinking and learning for every child*. New York: Free Press.

Rahman, N., (2012). Challenges and opportunities for learning analytics when formal teaching meets social spaces. In *LAK 2012 - Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 54-58). (ACM International Conference Proceeding Series). https://doi.org/10.1145/2330601.2330619

Reigeluth, C. M. (1999). *Instructional design theories and models: A new paradigm of instructional theory* (Vol. 2). Mahwah, NJ: Lawrence Erlbaum Associates.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.

Sheen, Y. (2004). Corrective Feedback and Learner Uptake in Communicative Classrooms across Instructional Settings. *Language Teaching Research*, *8*, 263-300. https:// doi.org/10.1191/1362168804lr146oa

Skinner, B. F. (1958). Reinforcement today. *American Psychologist*, *13*(3), 94–99. https://doi.org/10.1037/h0049039

Summers, J. J., & Ford, S. K. (1990). The Test of Attentional and Interpersonal Style: An evaluation. *International Journal of Sport Psychology*, *21*, 102-111.

Ur, P. (1996). *A course in language teaching: practice and theory*. Cambridge University Press.

Wang, W., & Loewen, S. (2016). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Language Teaching Research*, *20*(4), 459–478. https://doi.org/10.1177/1362168815577239

Weiqing W. & Li, S. (2020): Corrective feedback and learner uptake in American ESL and Chinese EFL classrooms: A comparative study, Language, Culture and Curriculum, DOI: 10.1080/07908318.2020.1767124

## 10. Appendices

## Appendix A

Sample Uptake Sheet

| | Who said it? (check as many as you wish) | | | Was it NEW to you? | |
|---|---|---|---|---|---|
| What are you noticing about… | Teacher | Classmate | Me | Yes | No |
| Pronunciation | | | | | |
| Vocabulary | | | | | |
| Grammar | | | | | |

## Appendix B

**Test of Attentional and Interpersonal Style (Short Form)**

0 = Never; 1 = Rarely; 2 = Sometimes; 3 = Frequently; 4 = All the time

Please circle the number that describes you most.

1. I am good at quickly analyzing a complex situation such as how a play is developing in football or which of four or five kids started a fight. 0 1 2 3 4

2. It is easy for me to bring together ideas from a number of different areas. 0 1 2 3 4

3. In games I make mistakes because I am watching what one person does and I forget about the others. 0 1 2 3 4

5. It is easy for me to keep thoughts from interfering with something I am watching or listening to. 0 1 2 3 4

6. In a room filled with children or on a playing field I know what everyone is doing. 0 1 2 3 4

7. I have so many things on my mind that I become confused and forgetful. 0 1 2 3 4

8. I get confused trying to watch activities such as a football game or circus where many things are happening at the same time. 0 1 2 3 4

9. I have difficulty clearing my mind of a single thought or idea. 0 1 2 3 4 10.

10. When people talk to me, I find myself distracted by the sights and sounds around me. 0 1 2 3 4

11. It is easy for me to keep sights and sounds from interfering with my thoughts. 0 1 2 3 4

12. All I need is a little information and I can come up with a large number of ideas. 0 1 2 3 4

13. When people talk to me, I find myself distracted by my own thoughts and ideas. 0 1 2 3 4

## Appendix C

1. What would you do if you use a wrong word in your sentence formulation?
2. What would you do if teacher's FB does not suffice?
3. What would you do if your structure is ill-formed?
4. How would you rectify your errors concerning register of words?
5. How would you rectify your errors concerning cohesion and coherence?

**Research Paper**

University of Birjand

# Applied Linguistics Inquiry

*Spring 2023, Vol. 1, No. 1, pp. 142-170*

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2023.6314.1015

# Exploring Grit and Big Five Personality as Predictors of Foreign Language Achievement

Maryam Khodaverdian Dehkordi[1]

[1]Ph.D. in Applied Linguistics, Department of English language and literature, Yazd University, Yazd, Iran

**ARTICLE HISTORY**

Received: 14 June 2022
Revised: 14 October 2022
Accepted: 21 February 2023
Published: 01 March 2023

**CORRESPONDING AUTHOR**

E-mail: khodaverdianmaryam@stu.yazd.ac.ir

**ABSTRACT**

This study examines grit— trait-level perseverance and prolonged passion for primary goals— in a foreign language learning context. The participants of this study were 384 English as a Foreign Language (EFL) learners from different academic majors studying at Yazd University. They had enrolled in General English (GE) courses. A questionnaire and an inventory were administered to assess their grit and a set of five assumed predictors, respectively. This investigation sought two main objectives: to examine (a) the relationships between grit and big five personality traits, and (b) the roles of grit and big five personality traits as predictors of L2 achievement. Using structural equation modeling (SEM), we explored links among grit, five predictors of grit, and L2 achievement. The results indicated that three variables of big five personality traits (conscientiousness, extraversion, and openness to experience or intellect) are important predictors of grit in language learning. Furthermore, grit as a novel construct considerably predicted L2 achievement to a large extend and mediated the influence of the five predictive factors on L2 achievement. Grit, thus, constitutes a predictor of L2 learning and acts as an essential yet under-investigated role for success in language learning that is conceptually and empirically different from existing constructs.

**KEYWORDS:** Big five personality, Grit, L2 achievement, Predictors of grit

## 1. Introduction

Both literally and figuratively, people take journeys throughout their lifetimes. Education life may be regarded as one of these journeys. Nowadays, individuals may commence their education at a young age, either formally or informally, and endeavor to excel from that point onward. Success in any aspect of life necessitates a combination of skills, attributes, and personal qualities that an individual must possess. Nevertheless, irrespective of possessing a particular factor that impacts success, it is evident that such attributes or personal traits have a positive effect on academic results. Considering the fact that learning a language is like a marathon, not a sprint (Mercer, 2018), one of the personal traits that of vital importance is grit. Grit is as "a passion and perseverance to accomplish long-term goals whatever the obstacles and no matter how long it may take" (Duckworth et al., 2007, p. 541).

There has been growing interest in non-cognitive variables which influence students' academic success (Duckworth & Yeager, 2015). One of these variables, which has drawn much public attention in the educational and psychological literature, is grit. Grit is defined as "perseverance and passion for long-term goals" (Duckworth et al., 2007, p. 1087). Duckworth and Quinn (2009) asserted that grit can positively predict students' achievement and persistence, even beyond more traditional variables such as IQ. Research has shown that grittier students persist in doing valuable but unpleasant activities even in confrontation with setbacks and challenges (Duckworth et al., 2011).

**Cite this article:** Khodaverdian Dehkordi, M. (2023). Exploring grit and big five personality as predictors of foreign language achievement. *Applied Linguistics Inquiry*, *1*(1), 142-170. doi: 10.22077/ali.2023.6314.1015

Grit, which comprises perseverance of effort and consistency of interests, has been shown to predict life success and school achievement (Duckworth et al., 2007). It refers to extreme stamina and effort and it is a good predictor of life success such as scholastic achievement across the life span, graduation from high school, and job retention (Eskreis-Winkler et al., 2014). Grittier students often perceive school achievement to be like a marathon, and thus, persist longer, invest greater effort in their work to achieve their short and long-term goals (Fong & Kim, 2019). The combined evidence from the previous studies suggests that grit represents an efficient capacity that enables students to successfully navigate experiences of academic difficulty.

Similarly, according to Mihaela (2015), one of the personal traits that determines academic success is the personality factor. Therefore, this personality factor is a concern of education practitioners and researchers. One personality approach that can be used to identify aspects of personality is the big five personality approach (De Raad & Mlačić, 2015) and this approach is known as the structure of the personality model (De Raad & Mlačić, 2017). Neuroticism, extraversion, openness, agreeableness, and conscientiousness are 5 personality aspects of the big five personality approach model (McCrae & Costa Jr, 2008). Regarding language learning, Rosito (2018) found that these five aspects of big five personality traits have a significant impact on the students' academic achievement. Besides, Bhagat, Shetty, Husain, Mat, Simbak, Aung, and Oo, (2019) concluded that personality is the combination of individual character as an asset for academic achievement. In short, the big five personality traits can positively predict academic achievement.

In the field of educational research, the role that is played by intellectual and non-intellectual factors is widely acknowledged (Duckworth & Seligman, 2005). Grit, which is regarded as a non-intellectual factor and personal trait, can be defined as the ability to strive for a desirable goal with passion and without desisting in case of any obstacles. Therefore, as applies to every part of life, individuals who have higher levels of grit are more laborious in dealing with challenges and thus achieve more in the end. In general, some studies discover a relationship between grit and academic performance (Chang, 2014; Lee, 2017; Mason, 2018; Luthans, Luthans & Chaffin, 2019). However, the number of studies focusing on such issues in the field of English Language Learning is quite low and even scarce for the Iranian EFL context. Little research has been conducted to investigate whether students who belong to different types of Big Five personality traits might differ in their L2 achievement and it is the principal goal of this research study.

Therefore, this study explores Iranian EFL Learners' Grit in the language classroom and investigates the relationship between grit profiles, its big five personality traits predictors, and its L2 learning achievements. In the current study, it is hoped that understanding L2 learners' grit, sheds more light on how grit affects the L2 learners' success and failure and helps the teachers and teacher trainers achieve a fine-grained understanding of the ways through which they can help their students improve their state of grit.

## 2. Literature review

### 2.1. Conceptualizing grit

Duckworth et al. (2007, p. 1087) defined grit as "the tendency to sustain interest in and effort toward very long-term goals". Grit entails working strenuously toward challenges, maintaining effort and interest over years despite failure, adversity, and plateaus in progress. Duckworth et al. (2007) defined grit as trait-level perseverance and maintained passion for superordinate goals that foreground the sustained and intentional pursuit of a long-term outcome. Perseverance is more often studied as an outcome than as a predictor. For example, perseverance in difficult or impossible tasks has served as the dependent variable in studies of optimistic attribution style, self-efficacy, goal orientation, and depletion of self-control resources (see, e.g., Bandura, 1977; Baumeister et al., 1998; Elliott & Dweck, 1988; Muraven, Tice, & Baumeister, 1998; Seligman & Schulman, 1986). However, the study of perseverance as a predictor, in particular as a stable individual difference, was of keen interest to psychologists in the first half of the 20th century. In a review of the existing literature of his day, Ryan (1939, p. 737) concluded that "the existence of a general trait of persistence, which permeates all behavior of the organism, has not been established, though evidence both for and against such an assumption has been revealed". Very recently, positive psychology has renewed interest in the empirical study of character in general and in the trait of perseverance in particular (Peterson & Seligman, 2004).

In context, grit is primarily assessed through Angela Duckworth's twelve-item (Grit-O) and eight-item (Grit-S) self-report scales and these are the primary measures of grit applied across the majority of research studies on grit. Factor analysis procedures with these measures have identified that grit is comprised of two components, or factors (Duckworth et al., 2007). The Grit-S (i.e., the Short Grit Scale) retains this two-factor structure with four fewer items and improved psychometric properties (Duckworth & Quinn, 2009).

The first factor, consistency of interests, has been defined in the literature as a goal- and action-oriented long-term behavior—this is distinct from an individual or situational interest (Muenks et al., 2017). Consistency of interests is assessed via six items on the Grit-O, each measured using a five-point Likert scale: 1) I often set a goal, but later choose to pursue a different one; 2) New ideas and new projects sometimes distract me from previous ones; 3) I become interested in new pursuits

every few months; 4) My interests change from year to year; 5) I have been obsessed with a certain idea or project for a short time but later lost interest, and 6) I have difficulty maintaining my focus on projects that take more than a few months to complete.

The second factor, perseverance of effort, is conceptualized as the tendency to continue working toward a particular goal (Muenks et al., 2017). Perseverance of effort is also assessed via six items on the Grit-O, each measured using a five-point Likert scale: 1) I have achieved a goal that took years of work; 2) I have overcome setbacks to conquer an important challenge; 3) I finish whatever I begin; 4) Setbacks don't discourage me; 5) I am a hard worker, and 6) I am diligent.

Perseverance has been conceptualized as the capacity to remain focused and committed to both short- and long-term goals in the face of adversity (Constantin et al., 2011). Thus, perseverance reflects the perseverance of effort component of grit but does not necessarily address consistency of interests. Although constructs related to grit may have slight deviations in their operational definitions, they may provide important insight into relevant relationships within the current body of literature. Researchers have sought to disentangle grit from conceptually similar constructs and their findings provide support for investigating grit along with these related constructs. For example, Crede, Tynan, & Harms (2017) conducted a full meta-analysis to disentangle grit from other constructs, including conscientiousness, self-control, mental toughness, and cognitive ability. This research identified that grit is highly correlated with conscientiousness. In addition, this research identified that several of these related constructs were also strongly associated with academic performance.

The importance of intellectual talent to achievement in academic performance and all professional domains is well established, but less is known about other individual differences that predict success. During the last decade, there has been an increasing interest in identifying factors that make individuals excel and be more successful compared to others of equal intelligence, especially in the field of education (Duckworth et al., 2007). Apart from cognitive ability, talent and opportunity, high achievers demonstrate a variety of non-cognitive or motivational characteristics such as creativity, commitment, emotional intelligence, growth mindset, gratitude, self-confidence, and emotional stability (Duckworth, et al., 2007; Dweck, Walton, & Cohen, 2014). Such qualities have been found to positively impact academic outcomes, social relationships, as well as psychological and physical well-being (Duckworth & Gross, 2014; Duckworth & Yeager, 2015). More than 100 years before work on grit as one of the non-cognitive factors, Galton (1892) collected biographical information on eminent judges, statesmen, scientists, poets, musicians, painters, wrestlers, and others. Ability alone, he concluded, did not bring about success in any field. Rather, he believed high achievers to be triply blessed by "ability combined with zeal and with capacity for hard labour" (Galton, 1892, p. 33). Similar conclusions were reached by Cox (1926) in an analysis of the biographies of 301 eminent creators and leaders drawn from a larger sample compiled by J. M. Cattell (1903). Estimated IQ and Cattell's rank order of eminence were only moderately related (r = .16) when the reliability of data was controlled for. Rating geniuses on 67 character traits derived from Webb (1915), Cox concluded that holding constant estimated IQ, the following traits evident in childhood predicted lifetime achievement: "persistence of motive and effort, confidence in their abilities, and great strength or force of character" (Cox, 1926, p. 218). Following these studies, grit has become a central phenomenon in recent educational research (Gray & Mannahan, 2017) and as Duckworth (2016) states, it plays a much more significant role in students' success than talent. Both facets of grit are fundamental ingredients of success since perseverance of effort contributes to the achievement of mastery despite failure, and consistency of interest is essential in engaging in deliberate practice to reach mastery (Credé et al., 2016).

Duckworth et al. (2007) asserted that grit was an important predictor of success for a variety of individuals, including National Spelling Bee contestants, salespeople at a vacation-time-share company, students at West Point Military Academy, and Teach for America teachers. For instance, WSSDA Research Blast (2014) indicates that grittier students are more likely to pursue their graduate-level education. Therefore, defining grit can be considered essential to better understand its role in the educational context. Furthermore, Duckworth et al. (2007) introduced the construct of grit, defined as trait-level perseverance and passion for long-term goals, and showed that grit predicted achievement in challenging domains over and beyond measures of talent. For instance, at the U.S. Military Academy, West Point, cadets higher in grit were less likely to drop out than their less gritty peers, even when controlling for SAT scores, high school rank, and a measure of big five conscientiousness. In four separate samples, grit was found to be either orthogonal to or slightly inversely correlated with intelligence. According to Duckworth et al.'s (2007) review of the literature, studies showed that even if a person was intelligent, that intelligence was useful only if the person was persistent. Grit is also related to but distinct from the need for achievement (n Achievement: McClelland, 1961). Grit differed from achievement in that achievement was "short term intensity" and grit was "long-term stamina" (Duckworth et al's, 2007. p. 1089).

Additionally, Duckworth et al. (2007) proposed that grit is distinct from traditionally measured facets of big five conscientiousness in its emphasis on stamina. In particular, grit entails the capacity to sustain both effort and interest in projects that take months or even longer to complete. Individuals high in grit do not swerve from their goals, even in the absence of positive feedback. Grit and effort can highly motivate an individual to overcome setbacks that he or she encounters (Dweck, 2006; Robert, 2009). By framing the distinctions between these constructs as one related to the specificity of their respective focus, concepts in this domain of character strengths might be seen as complementing each other like pieces of a puzzle (e.g., Oxford, 2016).

This study reveals that the students with higher levels of grit tend to be more determined. Grittier students regarded learning as a kind of journey and valued every stage in their learning process. In sum, grit can be regarded as one of the most significant personality traits in an individual's successful education. Since one of the primary roles of education is to enable people to pursue their goals and acquire abilities (Horn, 2013), the concept of grit gains importance in the field of education.

## 2.2. Research on grit

In the past ten years, there have been many research studies on grit including Duckworth et al. (2007) that identified this non-cognitive construct as "perseverance and passion for long-term goals". Duckworth et al. (2007) started to investigate why some people are more successful among the others who have the same intelligence. Duckworth et al. (2007, p. 1087) looked for a specific characteristic that distinguishes some successful people from others and found out it is "perseverance and passion for long term goals", which they defined as 'grit'. Then, she put forwards that the role of grit in terms of success is more significant than talent and IQ (Duckworth, 2016). Grit entails about four percent of the variance in lifespan success that absorbed researchers' attention and then researchers found that grit is the main predictor of success. Moreover, researchers have found that grit is diverse. They found that grit is predictive of many outcomes in education, including adults 'educational attainment, Ivy League undergraduates' grade point average, and the National Spelling Bee ranking (Duckworth et al., 2007). In addition, they identified the role of grit in predicting the retention rate of cadets in the United States Military Academy West Point, the success of employers, and overall career decision-making self-efficacy. (Duckworth et al., 2007; Mooradian et al., 2016; Vela et al., 2018). Besides the impact of grit on educational or career success, it predicts a variety of non-cognitive abilities including the big five conscientiousness component, forgiveness, and positive happiness (Arya & Lal, 2018; Duckworth et al., 2007). Eskreis-Winkler et al., (2014) conducted a study in which they explored the "grit effect" in four settings including the military, the workplace, high school, and marriage. They found that in these settings, intelligence, physical fitness, the big five personality traits, and job tenure were traditional predictors of retention, but grit was a more effective predictor of retention.

Research lends support to the idea that the big five personality traits have been identified as predictors of grit. Big five personality traits are composed of extraversion, agreeableness, conscientiousness, openness to experience, and neuroticism (Rimfeld et al., 2016). In this study, the applicability of grit to L2 learners' big five personality traits, and L2 achievement is examined. So, the relationship between grit, its big five personality traits predictors, and its L2 learning achievements are investigated.

## 2.3. Big five personality traits and L2 achievement

Personality refers to "the psychological qualities that contribute to an individual's enduring and distinctive patterns of feeling, thinking, and behaviors" (Cervone & Pervin, 2015, p 7). Among various ways of classification for personality traits, the Big Five model is the most widely used and widely known. The Big Five model consists of five main personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Costa & McCrae, 1985, 2008). Although there are many lower-order personality traits, the model of the five high-order traits has received strong empirical support from many years of research (Dewaele, 2012).

Individuals have different characteristics, such as personality qualities related to specific behaviors, cognitive, and emotional patterns (Hogan et al., 1996). McCrae and Costa (1987) asserted that personality has a five-factor structure including extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience or intellect. The operationalization of these big-five personality traits includes grit tendencies. They are essential predictors of grit. McCrae and Costa (1987) used the big five personality traits as a model for exploring the relationship between personality and different academic behaviors (Poropat, 2009). Following the big five model, a descriptive framework has been pointed out for many empirical research studies on the characteristics of predicting success (Goldberg, 1993; John & Srivastava, 1999; McCrae & Costa, 1987; Tupes & Christal, 1992). Barrick and Mount (1991) in a meta-analysis showed that big five conscientiousness is more related to job performance than the other factors.

Existing research on the Big Five traits and L2 learning achievement has shown the relevance of these traits in the context of L2 learning achievement (Zafar & Meenakshi, 2012b), as discussed in the following sections.

### 2.3.1. Predictors of grit

#### *2.3.1.1. Conscientiousness*

Conscientious individuals tend to be responsible, well-organized, and self-disciplinary (Costa & McCrae, 1992; McCrae & Costa, 1987). Several studies have suggested that conscientiousness is positively related to L2 learning (Cao & Meng, 2020; Lee, 2001). As found in Fazeli (2011), in a sample of Iranian English language learners, conscientiousness was positively related to all the L2 learning strategy uses; in other words, learners with higher levels of conscientiousness tend to use more

strategies in language learning. Lee (2001) found that, among a group of variables, conscientiousness was the best predictor for English proficiency (i.e., reading and listening) in a sample of Korean university students. Cao and Meng (2020) showed similar findings in a sample of Chinese university English language learners. However, the findings in this area were not always consistent, as shown in Ghapanchi et al. (2011), where conscientiousness had a non-significant relationship with L2 learning proficiency.

Also, Tross et al. (2000) studied conscientiousness as a means to predict college performance and found that conscientiousness was a stronger predictor of college GPA than HSGPA. Conard (2006) examined the incremental predictive validity of Big Five personality traits for affecting college GPA while controlling for SAT scores and found that conscientiousness and SAT scores had a direct effect in predicting college GPA, but the other traits were not predictive. Consistent with this result, Noftle and Robins (2007) observed that conscientiousness was a significant and positive predictor of college GPA, even when controlling for gender, SAT scores, HSGPA, and the other four Big Five factors. Similarly, Noftle and Robins (2007) concluded that conscientiousness was a slightly stronger predictor of college GPA than SAT scores.

#### *2.3.1.2. Openness to Experience*

Openness to experience refers to a person's willingness of being curious, imaginative, investigative, and exploring (Costa & McCrae, 1992; McCrae & Costa, 1987). Learners with higher levels of openness tend to be more open to non-traditional experiences and challenges, and tend to have the curiosity and interest to explore new things and phenomena. Such personal tendencies could facilitate their active learning that would lead to higher levels of L2 learning achievement. Research in this area has suggested that openness to experience could be one of the more salient predictors for L2 learning achievement (Oz, 2014; Zhang et al., 2019). For instance, in a US college student sample, Jackson (2019) found that students with higher levels of openness outperformed those lower in openness. In a recent study (Zhang et al., 2019), it was shown that, in a large sample of Chinese secondary school students, openness was related to students' English language learning interests. Steele-Johnson and Leas (2013) examined whether race and gender jointly affect the influence of personality on college GPA. They found that openness accounted for unique variance in GPA for male students and interacted with race in their effects on GPA. Furthermore, openness was more strongly related to GPA for Black male students than for White male students.

#### *2.3.1.3. Neuroticism*

People with neuroticism tend to be anxious, worrying, nervous, and emotional (Costa & McCrae, 1992; McCrae & Costa, 1987). Research has suggested that neuroticism was related to L2 learning (Lee, 2001), mostly in a negative way (Cao & Meng, 2020; Zhang, 2013). For example, as proposed in Erfani and Mardan (2017)'s study, among the five personality traits in the Big Five model, neuroticism was the only personality trait that was negatively correlated with both language proficiency and academic success in a sample of Iranian students who studied at different English-speaking universities around the world. However, some research indicated that the relationship between neuroticism and L2 achievement was nonsignificant (Ghapanchi et al., 2011; Kirkagac & Oz, 2017).

#### *2.3.1.4. Extraversion*

Extraversion reflects one's characteristics of being sociable, talkative, and of love to actively participate in interpersonal and social activities (Costa & McCrae, 1992; McCrae & Costa, 1987). Research suggested that extraversion could be beneficial for L2 learning (Cao & Meng, 2020; Oz, 2014), particularly in the domain of oral communication. The reason for this may not be surprising, given that extroverted individuals tend to be more likely to take the risk in communication, and tend to have less regard for potential mistakes that they may make (Zhang, 2008). This outgoing personality usually leads to an increased amount of input and output through interpersonal communication and activities (Krashen, 1985; Zafar & Meenakshi, 2012a), hence it is beneficial for their language proficiency development by reducing their language anxiety.

The seemingly obvious relationship between extraversion and L2 language learning as described above, however, is not always supported, especially when L2 achievement is measured with methods other than speaking, such as listening and writing. For example, Nejad et al. (2012) compared introverted and extraverted L2 learners and found that there is no significant relationship between extraversion and writing ability. Similarly, Nosratinia and Kounani (2016) found that extraversion/ introversion did not play a significant role in L2 learners' writing ability. Furthermore, some researchers argued that being extraverted is not necessarily linked to better L2 learning. The results of Ehrman (2008) suggested that the best language learners were more likely to have an introverted personality. Alavinia and Sameei (2012) also found that introverted L2 learners performed better in L2 listening than extroverts in a group of intermediate-ability Iranian English language learners. As discussed in Dewaele (2012), quiet and hard-working introverts could also be successful in L2 through different ways, such as enjoying personal reading instead of social activities. Furthermore, the positive or negative impact of personality on L2 learning outcomes could vary in different pedagogical situations. For example, L2 classes with more activities that require deep

reflection and memorizing the rules and vocabulary may benefit introverted L2 learners, while in classes with more activities requiring speaking and communicating, extraverts may feel more engaged and may learn better (MacIntyre & Charos, 1996).

#### *2.3.1.5. Agreeableness*

Agreeable individuals tend to be soft-hearted, helpful, sympathetic, friendly, and caring for others (Costa & McCrae, 1992; McCrae & Costa, 1987). Generally, people who are kind and agreeable are more likely to have pleasant contact with the target L2 communities and groups (MacIntyre & Charos, 1996). Research suggested that a more positive attitude towards the L2 community and a higher degree of intercultural contact could enhance motivation in L2 learning (Al Qahtani, 2015; Au, 1988; Gardner, 1985). Shirdel and Naeini (2018) showed that there was a significant relationship between agreeableness and foreign language achievement among college English language students in Iran. Similar results were found in Oz (2014), which suggested that agreeableness was a significant contributor in promoting students' willingness to communicate in L2 in a sample of Turkish college student.

Generally, the theoretical justifications of conscientiousness, openness to experience, extraversion, and agreeableness connection to components of language learning and L2 achievement are more than the theoretical justifications of neuroticism.

Previous research studies have shown that L2 achievement was linked to higher levels of conscientiousness (Chamorro-Premuzic & Furnham, 2003), extroversion (Matthews, 1997), openness to experience (Busato et al., 1999), agreeableness (Vermetten et al., 2001) and lower level of neuroticism (Entwistle, 1988). Also, research studies that have correlated grit with the personality dimensions specified in the big five model show similar results, but there are some differences. In one of the research studies, Duckworth et al. (2007) found that grit and conscientiousness have a strong correlation ($r = 0.77$, $p < 0.001$) followed by neuroticism ($r = -0.38$, $p < 0.001$), agreeableness ($r = 0.24$, $p < 0.001$), extraversion ($r = 0.22$, $p < 0.001$) and openness to experience ($r = 0.14$, $p < 0.001$). In another study, Lin and Chang (2017) explored the connections between personality dimensions and grit in high school students. Similarly, the results of the latter study also showed that conscientiousness is the superior predictor of grit ($\beta = 0.44$, $p < 0.001$) followed by neuroticism ($\beta = -.17$, $p < 0.001$), openness to experience ($\beta = 0.13$, $p < 0.001$) and agreeableness ($\beta = 0.11$, $p < 0.001$). Furthermore, the results of these research studies indicated that extraversion did not have a predictive effect on grit. Based on these research studies, it is reasonable to assume that grit may be linked to increased levels of conscientiousness, extroversion, openness to experience, agreeableness, and decreased extent of neuroticism in the L2 learning context. In our model, we will test whether the big five personality factors predict grit in the L2 learning context.

Several research studies have been conducted on exploring the influences of grit on different educational settings and L2 achievement. However, little research has been conducted to investigate whether students who belong to different types of big five personality traits might differ in their L2 achievement and it is the principal goal of this research study. Studying the association between grit, big five personality and L2 achievement can provide valuable information about the types of grit profiles may be associated with adaptive L2 achievement. The results may have an impact on the design of grit interventions for students with different grit profiles.

## 3. Design of the study

This study adopted a quantitative approach to find answers to the questions. Questionnaire data from 384 EFL learners comprising students of various academic majors who attend General English (GE) courses, assessed their grit and a set of five hypothesized predictors. Using multivariate analysis of variance (ANOVA), we examined links among the five predictor variables, grit profiles, and L2 achievement (GPA).

Besides, consistent with our review of the theoretical and empirical work on grit, we examined predictors and tested links between constructs through structural equation modeling (SEM).

To avoid the type of confirmation bias common in SEM, we also tested several competing models. Specifically, we tested the fit of the model when each of the hypothesized predictors (e.g., conscientiousness, neuroticism, agreement, extraversion, and openness to experiencing) is successively used as the mediator instead of grit to examine how well each model accounts for the data.

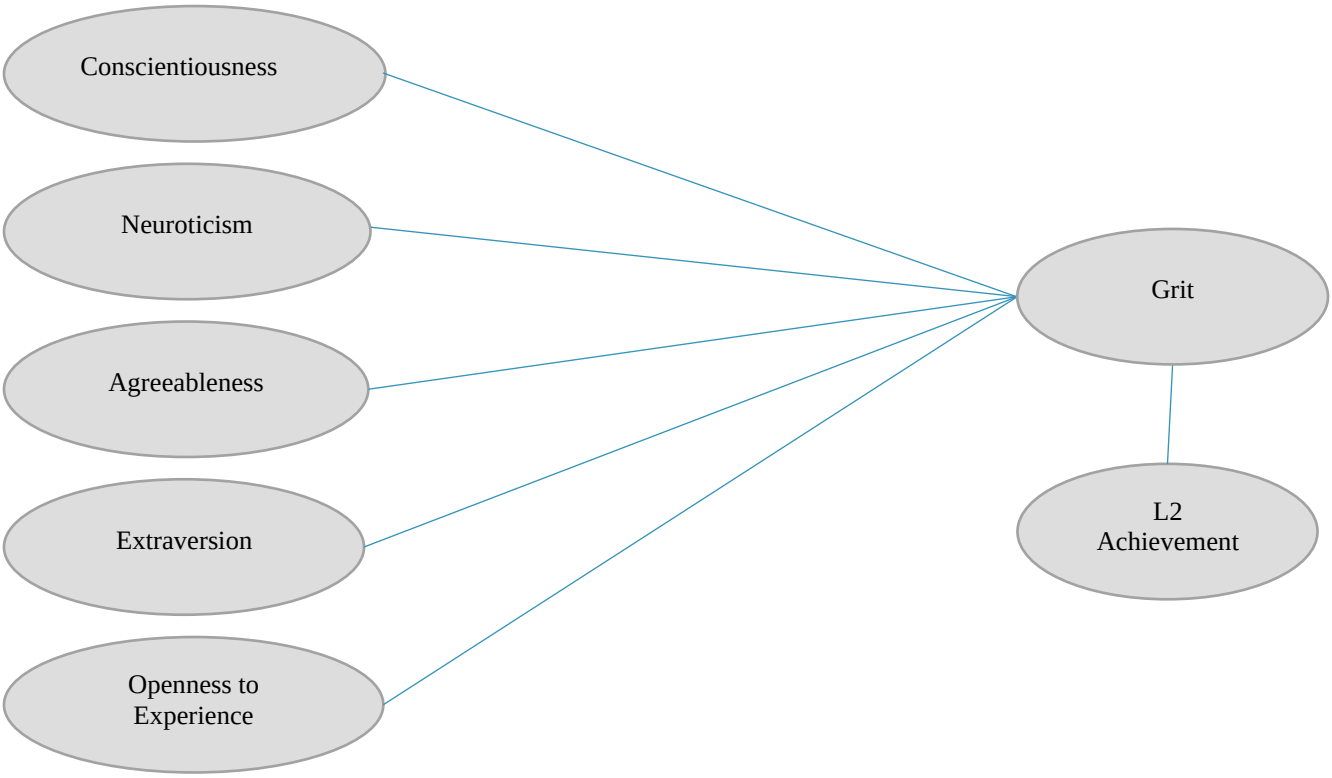



**Figure 1.** Hypothesized Conceptual framework of the study

In this study, to choose the standard methods and framework, one investigates whether grit exists as a researchable construct in the field of language learning and whether it influences L2 achievement. The following questions are investigated:

1. What is the relationship between the big five personality traits (extraversion, agreeableness, conscientiousness, openness to experience, and neuroticism) and grit?

2. How do grit and big five personality traits predict L2 achievement?

Structural equation modeling (SEM) is a variable-based method that allows us to test our questions.

## 4. Method

### 4.1. Participants

The current research was conducted at the Faculty of Humanities at Yazd University. Using quota sampling, the participants were 384 EFL students (60.4% male, 39.6% female) comprising students of various academic majors who attended General English (GE) courses that were designed to help students make rapid progress in English, and focus on the four key language skills – reading, writing, listening and speaking – with lots of additional work on vocabulary, grammar and pronunciation. The participants of this study comprised students of different ages from 19 to 24. As can be seen in figure 1, the highest frequency is related to the 20-year-old age group (40.6%) and the lowest frequency is related to the 22-year-old age group (2.9%). Language proficiency was identified by Students self-assessment as beginner (0.29%), lower-intermediate (0.53%), intermediate (2.3%), upper-intermediate (13.8%), and advanced (7%). All participants' mother tongue was Persian.

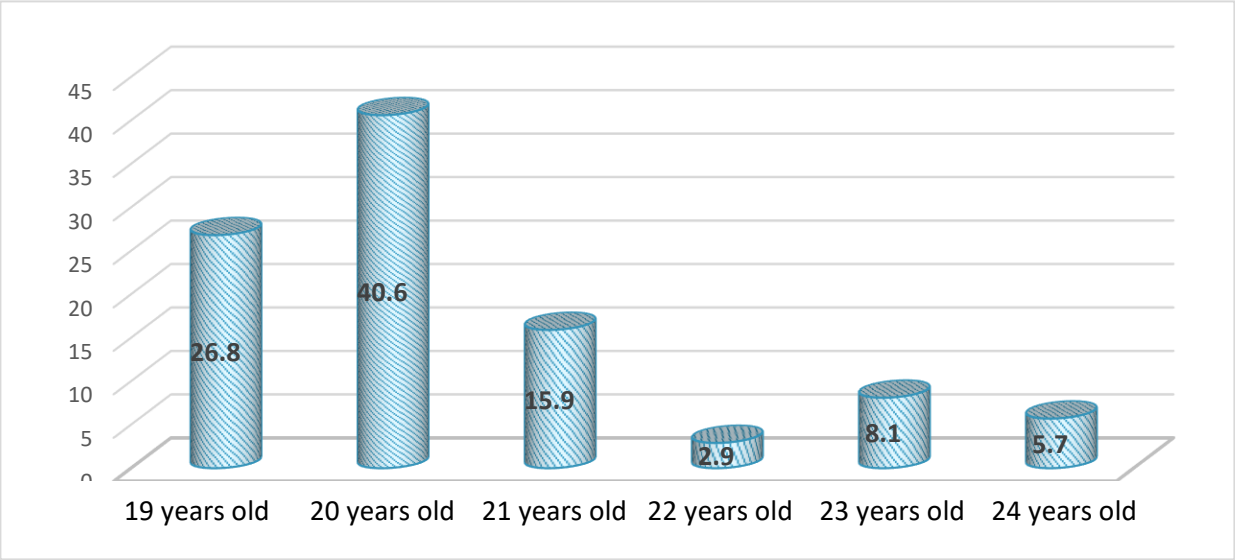



**Figure 2.** Distribution of the Participants According to Age Status

## 4.2. Instrumentation

The participants completed the grit questionnaire and big five inventory.

### 4.2.1.Grit questionnaire

The present study used the L2 short grit scale developed by Duckworth and Quinn (2009) to measure participants' passion and perseverance for long-term goals. The items were rated on a five-point Likert scale from 1 (Not like me at all) to 5 (Very much like me). The 8-item grit scale has two 4-item subscales, namely perseverance of effort and consistency of interest which is presented in Appendix I. The Cronbach's alpha reliability coefficients of the perseverance and consistency subscales were 0.827 and 0.909, respectively.

### 4.2.2.Big five inventory

According to John and Srivastava (1999) the BFI consists of 44 short items, rated on a 5-point scale from strongly disagree to strongly agree which is presented in Appendix II. The BFI items are assigned to five measurement scales: Extraversion (E; 8 items), Agreeableness (A; 9 items), Conscientiousness (C; 9 items), Neuroticism (N; 8 items), and Openness to experience (O; 10 items).

## 4.3. Data collection procedure

The data collection method was based on questionnaire. The short-term grit scale developed by Duckworth and Quinn (2009) was used to assess our participants' level of grit (i.e., perseverance and passion for long-term goals). Besides, all participants were asked to fill big five inventory (BFI; John & Srivastava, 1999), which is a 44-item questionnaire widely used to measure the big five personality factors of individuals. Participants rated items such as "I see myself as someone talkative" on a 5-point Likert scale with options ranging from (1= disagree strongly) to (5 = agree strongly).

It is worth mentioning that the whole data collection procedure happened in January 2020 before the Covid-19 pandemic crisis in Iran. Therefore, the paper and pencil version of all the instruments was used to collect the data from the students of Yazd University.

Since translation may jeopardize the reliability and validity of the instruments, a procedure called translation-back translation is performed to minimize this threat. Two non-affiliated researchers familiar with questionnaire construction translated the questionnaires into the participants' L1 and these questionnaires were back-translated for consistency. Cronbach's α was used to evaluate internal consistency. The Reliability of the scales is shown in Table 3.1. Following ethics approval, written consent forms from faculties in the universities and verbal participant assent were applied. They were told that their participation was entirely voluntary. The participants were assured about the confidentiality of their responses. The

participants were asked to answer the Grit Scale and big five inventory and at the end of the semester, their L2 test scores as measures of their L2 learning achievement were collected.

At the end of the semester, participants' final grades were obtained to assess their L2 achievement. They took a test that included reading comprehension, grammar, vocabulary, and pronunciation. In the Iranian educational system, scores range from 0 to 20 and the highest score is 20. If they get a minimum grade of 10, they will pass the course. Course grades are frequently used in L2 research (see Brown, Plonsky, & Teimouri, 2018 for a review).

### 4.4. Data analysis

ANOVA was conducted to determine whether or not these clusters significantly differed on the perseverance of effort and consistency of interests. The hypothesized predictors were selected for determining cluster membership (i.e., conscientiousness, neuroticism, agreement, extraversion, and openness to experiencing), while grit served as the criterion variable.

Consistent with our review of the theoretical and empirical work on grit, we examined predictors and tested links between constructs through structural equation modeling (SEM). Confirmatory factor analysis was used to indicate the validity of the questions related to the research variables. According to CFA, there was a significant correlation between the relevant latent variables and their corresponding indices. Specifically, we tested the fit of the model when each of the hypothesized predictors (e.g., conscientiousness, neuroticism, agreement, extraversion, and openness to experiencing) is successively used as the mediator instead of grit to examine how well each model accounts for the data.

More details on data analyses and the descriptive and inferential statistics are provided in the following chapter.

## 5. Results

### 5.1. Confirmatory factor analysis of research variables

In this part of the research, the validity of the questions related to the research variables has been evaluated using confirmatory factor analysis. The standardized coefficient measurement model (figure 2) shows that there is a significant correlation between the relevant latent variables and their corresponding indices.

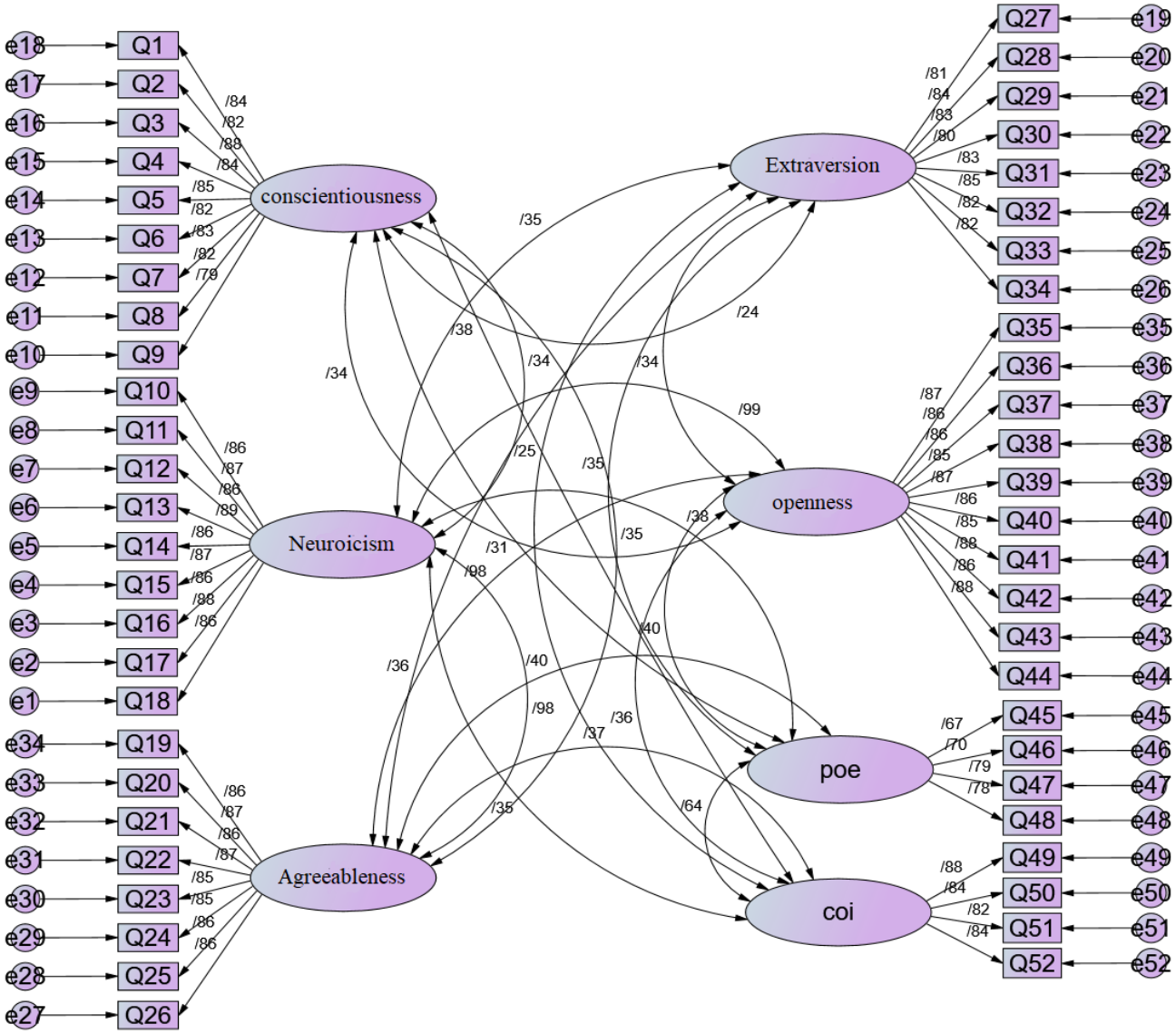



**Figure 3.** Standardized factor load coefficients of confirmatory factor analysis

## 5.2. Standardized factor loading

As can be seen in Table 1, the standardized factor loading for all questions is greater than 0.4. The construct reliability (CR) between items and corresponding latent variables is greater than 1.96. Also, their significance level is less than 0.05. Then, it can be said that the validity of the measurement structures of the relevant variables is confirmed at a significance level of 0.05. Therefore, there is no need to change or remove the question in the research model and questionnaire.

**Table 1.** Standardized factor loading between indices and latent variables

| | Questions | r | CR | sig | Cronbach's alpha |
|---|---|---|---|---|---|
| conscientiousness | Q1 | 0.836 | 18.695 | *** | 0.952 |
| | Q2 | 0.818 | 18.166 | *** | |
| | Q3 | 0.878 | 20.038 | *** | |
| | Q4 | 0.835 | 18.679 | *** | |

| | Questions | r | CR | sig | Cronbach's alpha |
|---|---|---|---|---|---|
| | Q5 | 0.854 | 19.254 | *** | |
| | Q6 | 0.819 | 18.183 | *** | |
| | Q7 | 0.827 | 18.445 | *** | |
| | Q8 | 0.818 | 18.174 | *** | |
| | Q9 | 0.790 | - | - | |
| neuroticism | Q10 | 0.681 | 23.153 | *** | 0.964 |
| | Q11 | 0.869 | 23.575 | *** | |
| | Q12 | 0.856 | 22.889 | *** | |
| | Q13 | 0.885 | 24.444 | *** | |
| | Q14 | 0.859 | 23.022 | *** | |
| | Q15 | 0.873 | 23.788 | *** | |
| | Q16 | 0.860 | 23.107 | *** | |
| | Q17 | 0.878 | 24.030 | *** | |
| | Q18 | 0.861 | - | - | |
| agreeableness | Q19 | 0.858 | 23.113 | *** | 0.958 |
| | Q20 | 0.867 | 23.555 | *** | |
| | Q21 | 0.859 | 23.137 | *** | |
| | Q22 | 0.869 | 23.681 | *** | |
| | Q23 | 0.853 | 22.845 | *** | |
| | Q24 | 0.853 | 22.840 | *** | |
| | Q25 | 0.857 | 23.047 | *** | |
| | Q26 | 0.863 | - | - | |
| extraversion | Q27 | 0.815 | - | - | 0.944 |
| | Q28 | 0.838 | 19.556 | *** | |
| | Q29 | 0.826 | 19.118 | *** | |
| | Q30 | 0.800 | 18.255 | *** | |
| | Q31 | 0.827 | 19.176 | *** | |
| | Q32 | 0.846 | 19.824 | *** | |
| | Q33 | 0.820 | 18.918 | *** | |
| | Q34 | 0.820 | 18.944 | *** | |
| openness to experience | Q35 | 0.871 | - | - | 0.967 |
| | Q36 | 0.858 | 23.597 | *** | |
| | Q37 | 0.862 | 23.822 | *** | |
| | Q38 | 0.854 | 23.381 | *** | |
| | Q39 | 0.871 | 24.305 | *** | |
| | Q40 | 0.862 | 23.795 | *** | |

| | Questions | r | CR | sig | Cronbach's alpha |
|---|---|---|---|---|---|
| | Q41 | 0.854 | 23.329 | *** | |
| | Q42 | 0.876 | 24.600 | *** | |
| | Q43 | 0.859 | 23.636 | *** | |
| | Q44 | 0.884 | 25.103 | *** | |
| POE | Q45 | 0.673 | - | - | 0.827 |
| | Q46 | 0.703 | 11.715 | *** | |
| | Q47 | 0.785 | 12.746 | *** | |
| | Q48 | 0.782 | 12.714 | *** | |
| COI | Q49 | 0.884 | - | - | 0.909 |
| | Q50 | 0.839 | 21.401 | *** | |
| | Q51 | 0.819 | 20.554 | *** | |
| | Q52 | 0.843 | 21.600 | *** | |

The measurement model showed good fit x2/ df = 2.795 that is less than 3, GFI= 0.81, RMSEA= 0.072, IFI= 0.91, CFI= 0.91. All factor loadings were significant and greater than .70. The skewness and Kurtosis coefficient of all variables are between +2 and -2, indicating that the observed covariance terms fit well with the estimated covariance terms. The construct reliability (CR) of variables in this model is more than 1.96. Also, Cronbach's alpha coefficient of all variables is more than 0.7. Table 2 also shows that the model had acceptable discriminant validity (see Hair, Black, Babin, & Anderson, 2010). The basic results of this model are shown in figure3. According to this model, it is clear that among the five predictors of conscientiousness, neuroticism, agreement, extraversion, and openness to experiencing; conscientiousness has the highest coefficient (0.26) and neuroticism has the lowest one (-0.10).

**Table 2.** Reliability of Constructs and Their Inter-correlations in the Measurement Model

| | Cronbach's α | CR | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|
| Conscientiousness | 0.952 | 5.130 | | | | | | |
| Neuroticism | 0.964 | -1.034 | 0.328 | | | | | |
| Agreeableness | 0.958 | 4.958 | 0.335 | 0.942 | | | | |
| Extroversion | 0.944 | 2.930 | 0.231 | 0.333 | 0.340 | | | |
| Openness to experience | 0.967 | 2.826 | 0.329 | 0.953 | 0.948 | 0.322 | | |
| Grit | 0.887 | 26.832 | 0.380 | 0.380 | 0.401 | 0.309 | 0.394 | |

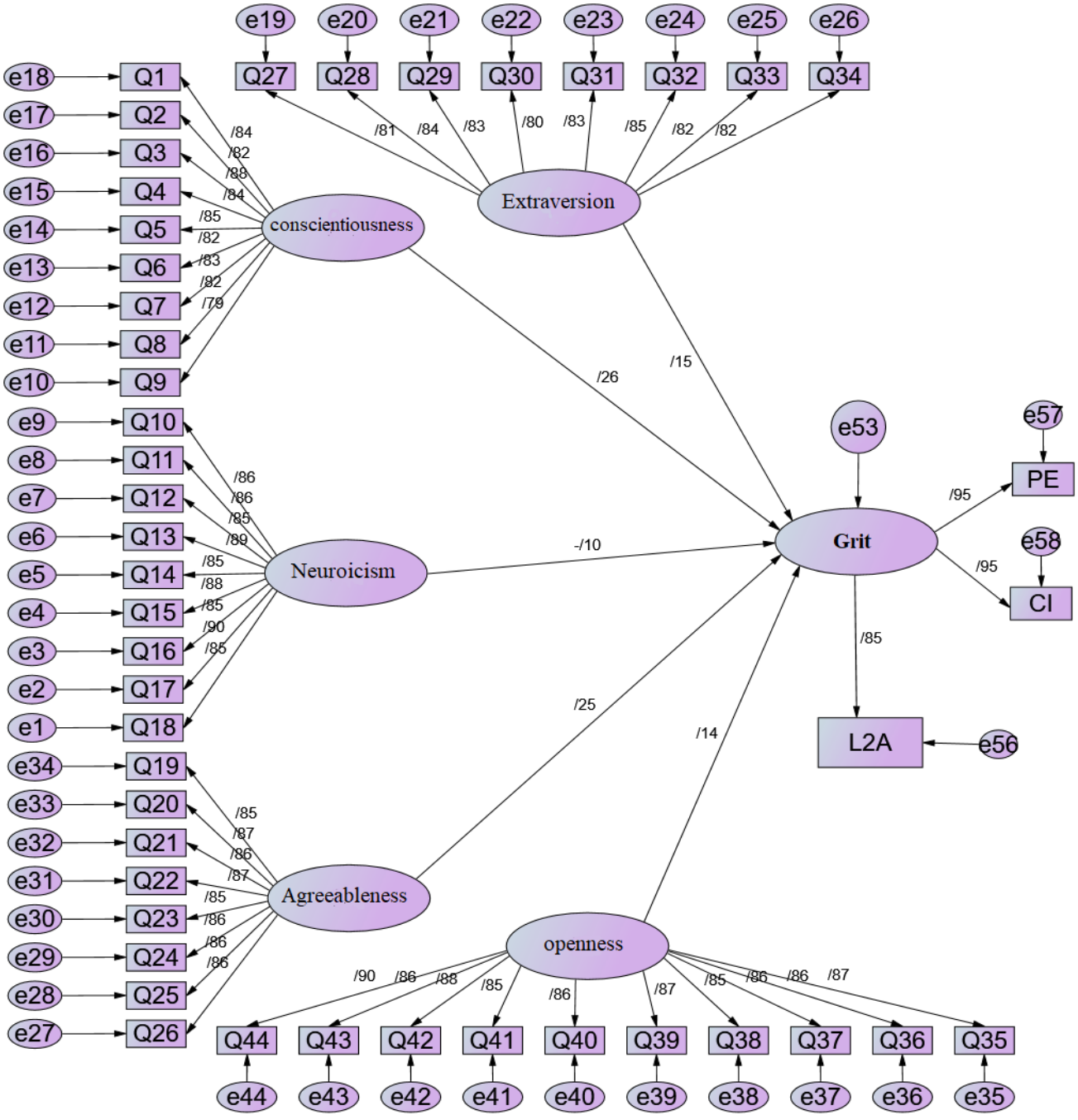



**Figure 4.** Research model in standardized path coefficient mode

### 5.3. SEM models

Six different SEM models were tested to investigate the relations among grit, big five personality traits, and L2 achievement. In the first model, the unique role of grit in L2 achievement was examined. To avoid the confirmation bias that is common in SEM, we examined five competing models (Hiver & Al-Hoorie, 2020). We examined the fit of the model when each of the assumed predictor variables (e.g., conscientiousness, agreeableness, extraversion, neuroticism, and openness to experiencing) was used as the mediator instead of grit to examine the way each model accounts for the data and then to indicate the best fit model.

### 5.3.1. Grit as a predictor of L2 achievement

The first model tested the role of grit as a predictor of L2 achievement (Model A). The Goodness of fit indices indicated that the model fitted the data adequately (Table 3). Model A in Figure 4 showed a good fit *x2/df* = 2.795, *CFI* = 0.91, *RMSEA* = 0.072.

**Model A**

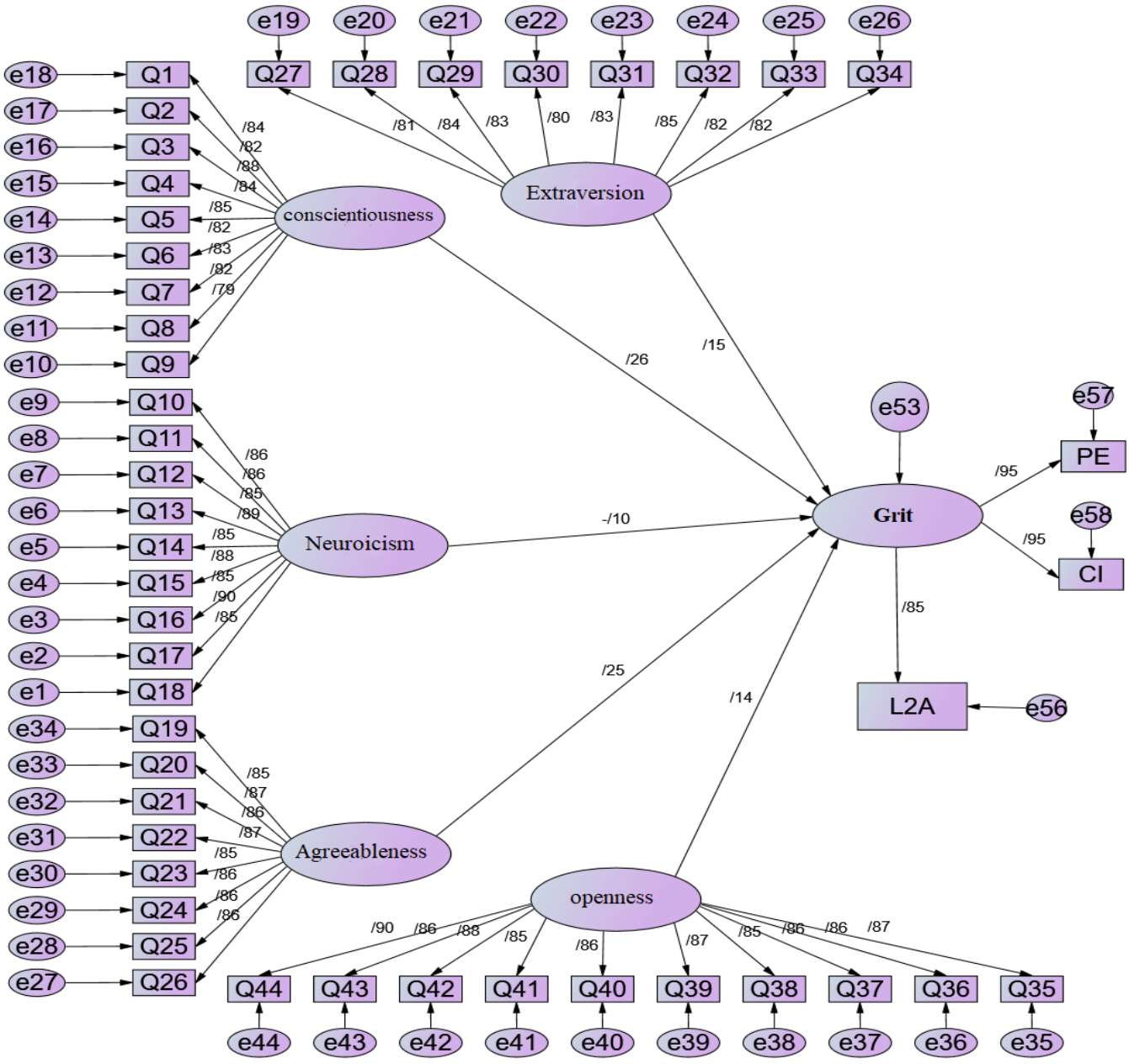



**Figure 5.** Grit as a Predictor of L2 Achievement with Standardized Coefficients

## 5.4. Competing models

Through five competing models, the simultaneous role of the big five personality factors and grit in predicting L2 achievement was examined. As Table 3 indicates, the goodness of fit indices for Model A is the better fit. The AIC and BIC values are basic

fit indices, in a way that lower values indicate better fit (Klein & Kotov, 2016). It can be seen from Table 3 that smaller values of AIC and BIC are identified for Model A. Both goodness-of-fit indices, AIC and BIC values show that Model A is the best model. Therefore, this model is reported to test the relationship between grit, big five personality traits, and L2 achievement. Competing models figures can be seen in Appendix B.

**Table 3.** Model Fit for Each of the Competing Models Tested

| | x2/df | CFI | GFI | IFI | RMSEA | AIC | BIC |
|---|---|---|---|---|---|---|---|
| Grit | 2.795 | 0.91 | 0.81 | 0.91 | 0.072 | 3279.888 | 3671.002 |
| Conscientiousness | 3.415 | 0.87 | 0.77 | 0.87 | 0.081 | 3712.200 | 4103.313 |
| Extroversion | 3.436 | 0.87 | 0.77 | 0.87 | 0.080 | 3734.056 | 4125.170 |
| Openness to Experience | 2.701 | 0.91 | 0.81 | 0.91 | 0.073 | 3771.608 | 4162.722 |
| Agreeableness | 2.864 | 0.90 | 0.82 | 0.90 | 0.074 | 3836.063 | 4227.177 |
| Neuroticism | 3.685 | 0.85 | 0.76 | 0.85 | 0.084 | 3996.573 | 4371.884 |

*Note:* AIC= Akaike Information Criterion, BIC= Bayesian Information Criterion, CFI= Bentler's Comparative Fit Index, RMSEA= Root Mean Square Error of Approximation.

The results showed that, except for neuroticism, all hypothesized predictors (i.e., conscientiousness, agreement, extraversion, and openness to experiencing) predicted grit highly significantly. Also, grit predicted L2 achievement.

## 6. Discussion

According to our SEM results, three variables are important predictors of grit: conscientiousness, extraversion, and openness to experience or intellect. This suggests that learners are mostly dependable, responsible, and able to set their goals and then persist to meet them. Highly intelligent people have a positive attitude toward challenges and setbacks in language learning (Barrick & Mount, 1991) as opposed to those who are mainly narrow-minded (McCrae & Costa, 1987). As we expected, our results illustrated that achievement striving, positive emotions, and wide interests have a positive impact on learners' grit, and this may be because these factors involve compliance and cooperativeness that enable individuals to be imperturbable in learning. (McCrae & Costa, 1987). Learners with relatively lower effort and interest in their L2 learning processes are less likely to be gritty individuals and then less accomplish their L2 learning goals. Besides, our SEM results confirm a part of Duckworth et al.'s (2007), Lin and Chang's (2017) research studies in which conscientiousness is the superior predictor of grit. It is followed by neuroticism, agreeableness, extroversion, and openness to experience in the former; and neuroticism, openness to experience, agreeableness, no predictive effect of extraversion on grit in the latter. In this study conscientiousness followed by extroversion, openness to experience, agreeableness, and no predictive effect of neuroticism. Also, according to this study grit is highly correlated with conscientiousness (Crede et al.,2017; Cao & Meng, 2020; Lee, 2001; Fazeli, 2011).

About tested competing models, the model where grit was a predictor illustrated a better fit than those where the other five variables were predictors: Grit either significantly predicts L2 achievement or completely mediates the impact of other big five factors on L2 achievement. It is aligned with the research on grit that found out the extreme stamina and effort as a good predictor of success in life such as academic achievement through life, graduation from high school, and job maintenance (Eskreis-Winkler et all., 2014; Akin and Arslan,2014). While not explicitly aligned with research on grit, Khajavy, MacIntyre, Hariri's (2020) study on the relation between grit components and L2 achievement found no significant relation either in correlations or in SEM analyses. Our results suggest that among the big-five factors, conscientiousness is more related to grit especially effort and persistence (Chamorro-Premuzic & Furnham, 2003). Also, our results confirm Geisler-Brenstein

et al.'s study (1996); asserted that there is a direct and positive relationship between conscientiousness and systematic learning and one way that enables learners to be successful in language learning is through the facilitation of grit.

Creating and maintaining conscientiousness can be accomplished by controlling, regulating, and directing the impulses which are not inherently bad. People see impulsive individuals as brilliant, fun-to-be-with, and natural. High conscientiousness has many benefits. Conscientious individuals keep away from trouble and accomplish high levels of success through definite planning and persistence. Also, people regard them as intelligent and reliable. (Kotchoubey, 2018)

As for neuroticism, contrary to previous research indicating its importance in motivation and effort (Norem and Cantor, 1986) it did not function as an essential positive predictor of grit. Although it is understood as a defensive force, has a negative effect rather than a positive (Matthews and Zeidner, 2004).

## 7. Conclusion

Conventional predictors of academic performance, including gender, race, ACT/SAT scores, and HSGPA, have been widely used to predict student academic performance (Komarraju et al., 2013). Yet research has begun to study specific personality traits, like grit, as predictors of academic performance in college in combination with traditional cognitive measures (Duckworth et al., 2007; Strayhorn, 2014). Unlike the extensive empirical research on the use of traditional cognitive ability measures to understand academic performance, personality factors, like grit, have not been empirically studied in the literature.

The primary purpose of this study was to examine whether personality traits and grit, predict EFL learners' academic performance. In addition, the moderating effect of grit in the relation of learners' big five personality traits was examined. These questions were explored using SEM. The results of this study suggest four important conclusions.

First, both POE and COI are essential elements of L2 achievement. This implies that, in the context of L2 learning, learners should be trained to improve in perseverance and consistency. Second, according to our expectation, the grit profile characterized by high perseverance and low consistency did show a robust buffering role against learners' success. This may indicate that the profile, highlighting low consistency only, might not have a strong effect on a learner's high academic performance. Third, neuroticism did not predict grit, and this might be because of the lack of effective cognitive skills (Eysenck, 1967) or surface learning instead of meaningful learning (Entwistle, 1988). Fourth, results suggest that one way for learners' achievement is through the facilitation of grit within the framework of the big five personality traits. It can be inferred that grit, even when big five traits were added may predict higher performance and grades.

Compared to the most positive factors for learners (conscientiousness, extraversion, openness to experience, agreeableness), neuroticism may not be as prominent here. Our research results delineate that L2 learners who possess these constructs have an elaborate high L2A despite a certain degree of neuroticism. Generally, the theoretical justifications of conscientiousness, openness to experience, extraversion, and agreeableness connection to components of language learning and L2 achievement are more than the theoretical justifications of neuroticism (McCrae & Costa, 1987). Despite the concentration on preventing or decreasing the negative constructs such as anxiety, hostility, depression, and impulsiveness that accompany L2 learning, it's better to build up positive indicators (Oxford, 2016). Repositioning success in this way will be more effective because these positive indicators have successfully strengthened the learners' L2 learning process and then help them successfully cope with and overcome daily stress and frustration. Given the point, that grit focuses on positive dealing with problems this may also make grit amenable to intervention at once. Our findings delineate that both POE and COI as two factors of grit have the potential to influence the key outcomes for instructed L2 learners.

Overall, this study shed light on L2 achievement by emphasizing the personality traits of grit. Furthermore, we focus on the individual level of grit, irrespective of variable level.

So, educational authorities should explore the ways through which they can incorporate elements of grit into students' educational environment. As this study shows, by exploring the characteristics of gritty students through the big five personality traits, educators can help their students enhance their grit and become better achievers. Cosequently, high achievers demonstrate higher L2 achievement (Duckworth, et al., 2007; Dweck, Wolton & Cohen, 2014). However, personality traits of grit have a key role to become a more successful language learner (Changlek & Palanukulworg, 2015). Given the point, if teachers employ care and control in the language classroom, grit is strongly associated with leranres' success. When teachers work with students to create grit, the biggest part of the process is helping students grapple with roadblocks. From teachers, students learn how to overcome hardship, thus learning grit (Hoerr, 2013).

## 8. Implications of the study

Some theoretical and pedagogical implications of the present study are presented in the following subsections.

### 8.1. Theoretical implications

The unique contribution of this study resides in a new proposal: By adopting a cluster analysis, the current study demonstrates when taking academic exams, high perseverance of effort and low consistency of interest were related to a higher level of neuroticism and lower level of consciousness. Also, all dimensions of the big five personality traits; openness to experience, conscientiousness, extraversion, agreeableness, and neurotisicm are significant predictors of academic achievement whether negatively or positively.

The results of this study provide evidence to support that the predictions of the previously mentioned models may not be realized because the role of the participants' characteristics was not considered as an influential factor.

Therefore, the studies in grit can be reexamined to see whether they had considered participants' characteristics or not and, if not, what are its possible effects.

This study provides basic new knowledge about learners' grit and their big five personality traits in L2 learning. In addition, it helps applied linguists to understand the general limitations and variations of grit in L2 achievement. It also awakens their enthusiasm to follow the paradigm, develop research, and make the subject more accessible to a large number of students in the field.

Regarding the theory, this study is related to the existing grit literature by elaborating the relationship between grit, big five personality traits, and L2A.

### 8.2. Pedagogical implications

The present study suggests that learners' grit will be increased in the light of their big five personality traits. Therefore, some implications for teaching and learning can be suggested: Language teachers, curriculum designers, and material developers should consider learners' characteristics in L2 learning.

Regarding the practice, our findings affirm the significance of developing and implementing the interventions of the big five personality traits in predicting grit.

## 9. Limitations

The current study, like any other study, bears several limitations. The first limitation of this study is the potential for social desirability bias from the use of self-report surveys, where respondents may answer questions in a way they think is more positive. It is the nature of self-report surveys that may result in social desirability bias.

Second, This study used data from a single institution, thus limiting the generalizability of the results. The institution in which the sample for this study was obtained is a highly selective institution with stringent admissions requirements. DeBerard, Spielmans, and Julka (2004) affirm that more selective universities in terms of the Scholastic Aptitude Test(SAT) should expect the greater achievement of their students. Therefore, a personality trait like grit may not be as meaningful in the current sample.

Third, The role of gender and the proficiency level of learners were not taken into consideration. Research shows that grit increases with age (Duckworth et al., 2007), and it would be valuable to understand how student grit levels may change through completing a college degree.

Forth, in the current study, students' L2 achievement is measured through their final-term grades. The course grade is usually used to assess L2 achievement, but it has been criticized for validity issues (see Brown et al., 2018 for a review).

## 10. Future research

Regarding the first limitation, researchers should continue to encourage respondents to answer truthfully on the Grit-S, as there would not be negative repercussions from their responses. To reduce the social desirability bias from the Grit-S, future research on grit should use informant-report versions of the grit survey along with the Grit-S to validate the self-reported grit scores. Duckworth and Quinn (2009) found that informants could reliably assess the grit of others ($\alpha = .83$). To ensure a more accurate grit score, faculty members or higher education administrators familiar with the students could serve as informants and complete the informant version of the Grit-S. The questions on the informant-report survey are identical to the self-report version, but the first-person pronouns are replaced with third-person pronouns.

Regarding the second limitation, further research on grit should include a more representative sample, so the results can be generalized to all college students.

Regarding the third limitation, future research can measure grit annually to see if it increases with experience and time in college.

Regarding the fourth limitation, researchers may have a more consistent measure of students' L2 achievement by using standardized foreign language achievement tests

## 11. References


Arya, B., & Lal, D. S. (2018). Grit and sense of coherence as predictors of well-being. *Indian Journal of Positive Psychology, 9*(1), 169-172.

Bhagat, Vidya, Charan Kishor Shetty, Rohayah Husain, Khairi Che Mat, Nordin Bin Simbak, Myat Moe Thwe Aung, & San San Oo. (2019). "The relationship between big five personality traits and academic performance in medical students." *Research Journal of Pharmacy and Technology, 12(9), 4189–96.*

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (1998). The relation between learning styles, the Big Five personality traits, and achievement motivation in higher education. *Personality and individual differences, 26*(1), 129-140.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74*, 1252–1265.

Brown, A. V., Plonsky, L., & Teimouri, Y. (2018). The use of course grades as metrics in L2 research: A systematic review. *Foreign Language Annals, 51*(4), 763-778.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of research in personality,* 37(4), 319-338.

Chang, W. (2014). Grit and academic performance: Is being grittier better? (Doctoral dissertation, University of Miami).

Constantin, T., Holman, A., & Hojbotă, M. A. (2011). Development and validation of a motivational persistence scale. *Psihologija, 45*(2), 99-120.

Cox, C. M. (1926). Genetic studies of genius. II. The early mental traits of three hundred geniuses.

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. Journal of personality and social psychology, 113(3), 492-511.

De Raad, B., & Mlacic, B. (2015). Big five factor model, theory and structure. *International Encyclopedia of the Social & Behavioral Sciences, 2*, 559-566.

De Raad, B., & Mlačić, B. (2017). The lexical foundation of the big five factor model. *The Oxford Handbook of the Five Factor Model*, 191-216.

Dornyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Duckworth, A., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current directions in psychological science, 23*(5), 319-325.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology, 92*(6), 1087-1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment, 91*(2), 166-174.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. Psychological Science, 16(12), 939–944.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells success: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science*, 2, 174-181.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*,44,237251. http://dx.doi.org/10.3102/0013189X1558437

Duckworth, A. L. (2016). Grit: *The power of passion and perseverance*. New York: Scribner.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House Digital, Inc.*Dweck, C.*,

Walton, G. M., & Cohen, G. L. (2014). Academic Tenacity: Mindsets and Skills That Promote Long-Term Learning. Seattle, WA: Bill & Melinda Gates Foundation.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*, 5–12.

Entwistle, N. (1988). Motivational factors in students' approaches to learning. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 21−49). New York: Plenum Press.

Eskreis-Winkler, L., Duckworth, A. L., Shulman, E. P., & Beal, S. (2014). The grit effect: Predicting retention in the military, the workplace, school, and marriage. *Frontiers in psychology*, *5*, 36-42.

Eysenck, H. J. (1967).*The biological basis of personality. Springfield, IL: Charles C. Thomas.*

Fong, C. J., & Kim, Y. W. (2019). A clash of constructs? Re-examining grit in light of academic buoyancy and future time perspective. *Current Psychology*, 1-14.

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*(1), 47-64.

Galton, S. F. (1892). *Hereditary genius: An inquiry into laws and consequences*: MacMillan.

Geisler-Brenstein, E., Schmeck, R., & Hetherington, J. (1996). An individual difference perspective on student diversity. *Higher education, 31*(1), 73-96.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American psychologist, 48*(1), 26-34.

Gray, J. P., & Mannahan, K. K. (2017). How Well Do Trait Measures of Achievement Predict Students' Perceptions of the Link between Personal Effort and Academic Performance? *Journal of Effective Teaching, 17*(1), 16-27.

Gregersen, T. (2016). The positive broadening power of a focus on well-being in the language classroom. In *Positive psychology perspectives on foreign language learning and teaching*, 59-73.

Hair, J., Black, W., Babin, B. Y. A., Anderson, R., & Tatham, R. (2010). Multivariate Data Analysis, 7, New Jersey: Pearson Prentice Hall.

Hiver, P., & Al-Hoorie, A. H. (2020). Reexamining the role of vision in second language motivation: A preregistered conceptual replication of You, Dörnyei, and Csizér (2016). *Language Learning, 70*(1), 48-102.

Hoerr, T. R. (2013). *Fostering Grit: How do I prepare my students for the real world? (ASCD Arias)*: ASCD.

Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American psychologist, 51*(5), 469-477.

http://wssda.org/Portals/0/News/Research%20Blast/WSSDA%20Research%20Blast%20May%202014.pdf

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102-138.

Kaplan, A., Katz, I., & Flum, H. (2012). Motivation theory in educational practice: Knowledge claims, challenges, and future directions. *APA educational psychology handbook*, *2*, 165-194.

Klein, D. N., & Kotov, R. (2016). The Course of depression in a 10-year prospective study: Evidence for qualitatively distinct subgroups. *Journal of abnormal psychology*, *125*(3), 337-348.

Kotchoubey, B. (2018). Human consciousness: where is it from and what is it for. *Frontiers in psychology*, *9*, 567-582.

Khajavy, G. H., MacIntyre, P. D., & Hariri, J. (2020). A closer look at grit and language mindset as predictors of foreign language achievement. *Studies in Second Language Acquisition*, 1-24.

Lee, S., & Sohn, Y. W. (2017). Effects of grit on academic achievement and career-related attitudes of college students in Korea. Social Behavior and Personality: An International Journal, 45(10), 1629-1642.

Lin, C.-L. S., & Chang, C.-Y. (2017). Personality and family context in explaining grit of Taiwanese high school students. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(6), 2197-2213.

Luthans, K. W., Luthans, B. C., & Chaffin, T. D. (2019). Refining grit in academic performance: The mediational role of psychological capital. Journal of Management Education, 43(1), 35-61.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, *52*(1), 81-90.

McCrae, Robert R., and Paul T. Costa Jr. (2008).“The five-factor theory of personality.”

Matthews, G. (1997). Extraversion, emotion, and performance: A cognitive-adaptive model.

In *Advances in Psychology* (Vol. 124, pp. 399-442): Elsevier.

Matthews, G., & Zeidner, M. (2004). Traits, states and the trilogy of mind: An adaptive perspective on intellectual functioning. *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development*, 143-174.

McClelland, D. C. (1961). The achieving society. *Oxford, England*: Van Nostrand.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, *52*(1), 81-90.

Mihaela, Păişi Lăzărescu. (2015). “Psychological factors of academic success.” *Procedia-Social and Behavioral Sciences*, *180*, 1632–37.

Mooradian, T., Matzler, K., Uzelac, B., & Bauer, F. (2016). Perspiration and inspiration: Grit and innovativeness as antecedents of entrepreneurial success. *Journal of Economic Psychology*, *56*, 232-243.

Mason, H. D. (2018). Grit and academic performance among first-year university students: A brief report. *Journal of Psychology in Africa*, *28*(1), 66-68.

Mercer, S. (2018). Psychology for language learning: Spare a thought for the teacher. Language Teaching, 51(4), 504-525.

Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students’ personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, *109*(5), 599-620.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*, 774–789.

Norem, J. K., & Cantor, N. (1986). Defensive pessimism: Harnessing anxiety as motivation. *Journal of personality and social psychology*, *51*(6), 1208.

Oxford, R. L. (2016). Toward a psychology of well-being for language learners: The "EMPATHICS" vision. *Positive psychology in SLA*, (10), 10-87.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322-338.

Peterson, C. & Seligman, M.E., 2004, Character strengths and virtues: A handbook and classification, Oxford University Press, Oxford.

Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of personality and social psychology*, *111*(5), 780.

Roberts, Y. (2009). Grit. The Skills for Success and How They Are Grown. The Young Foundation. *London*.

*Rosito, Asina Christina. (2018). "Eksplorasi tipe kepribadian big five personality trait dan pengaruhnya terhadap prestasi akademik." Jurnal Psikologi Pendidikan Dan Konseling: Jurnal Kajian Psikologi Pendidikan Dan Bimbingan Konseling 4(1), 6–13.*

Ryans, D. G. (1939). The measurement of persistence: An historical review. *Psychological Bulletin, 36*, 715–739.

Seligman, M. E., & Schulman, P. (1986). Explanatory style as a predictor of productivity and quitting among life insurance sales agents. *Journal of Personality and Social Psychology, 50*, 832–838.

Slaats, A., Van der Sanden, J., & Lodewijks, J. (1997). *Relating personality characteristics and learning style factors to grades in vocational education.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, USA.

Taspinar, H. K., & Kulekci, G. (2018). *International Journal of Languages' Education and Teaching* 6(3), 208-226.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of personality*, *60*(2), 225-251.

Ushioda, E. (2008). Motivation and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 19–34). Cambridge, UK: Cambridge University Press.

Vela, J. C., Sparrow, G. S., Whittenberg, J. F., & Rodriguez, B. (2018). The role of character strengths and importance of family on Mexican American college students' career decision self-efficacy. *Journal of Employment Counseling*, *55*(1), 16-26.

Vermetten, Y. J., Lodewijks, H. G., & Vermunt, J. D. (2001). The role of personality traits and goal orientations in strategy use. *Contemporary educational psychology*, *26*(2), 149-170.

Waninge, F., Dörnyei, Z., & De Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, *98*(3), 704-723.

Webb, E. (1915). Character and Intelligence. *British Journal of Psychology*. *1*(3), 99-120.

Yun, S., Hiver, P., & Al-Hoorie, A. H. (2018). Academic buoyancy: exploring learners' everyday resilience in the language classroom. *Studies in Second Language Acquisition*, *40*(4), 805-830.

## 12. Appetencies

### 12.1. Appendix I

**Short Grit Scale**

*Directions for taking the Grit Scale: Please respond to the following 8 items. Be honest – there are no right or wrong answers!*

1. New ideas and projects sometimes distract me from previous ones. *

- ❑ Very much like me
- ❑ Mostly like me
- ❑ Somewhat like me
- ❑ Not much like me
- ❑ Not like me at all

2. Setbacks don't discourage me.
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

3. I have been obsessed with a certain idea or project for a short time but later lost interest.*
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

4. I am a hard worker.
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

5. I often set a goal but later choose to pursue a different one. *
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

6. I have difficulty maintaining my focus on projects that take more than a few months to complete. *
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me

- ❑ Not like me at all

7. I finish whatever I begin.
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

8. I am diligent.
   - ❑ Very much like me
   - ❑ Mostly like me
   - ❑ Somewhat like me
   - ❑ Not much like me
   - ❑ Not like me at all

## 12.2. Appendix II

**The Big Five Inventory (BFI)**

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

| Disagree strongly | Disagree a little | Neither agree nor disagree | Agree a little | Agree Strongly |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

I see Myself as Someone Who...

____1. Is talkative

____2. Tends to find fault with others

____3. Does a thorough job

____4. Is depressed, blue

____5. Is original, comes up with new ideas

____6. Is reserved

____7. Is helpful and unselfish with others

____8. Can be somewhat careless

____9. Is relaxed, handles stress well

____10. Is curious about many different things

____23. Tends to be lazy

____24. Is emotionally stable, not easily upset

____25. Is inventive

____26. Has an assertive personality

____27. Can be cold and aloof

____28. Perseveres until the task is finished

____29. Can be moody

____30. Values artistic, aesthetic experiences

____31. Is sometimes shy, inhibited

____32. Is considerate and kind to almost everyone

____11. Is full of energy

____12. Starts quarrels with others

____13. Is a reliable worker

____14. Can be tense

____15. Is ingenious, a deep thinker

____16. Generates a lot of enthusiasm

____17. Has a forgiving nature

____18. Tends to be disorganized

____19. Worries a lot

____20. Has an active imagination

____21. Tends to be quiet

____22. Is generally trusting

____33. Does things efficiently

____34. Remains calm in tense situations

____35. Prefers work that is routine

____36. Is outgoing, sociable

____37. Is sometimes rude to others

____38. Makes plans and follows through with them

____39. Gets nervous easily

____40. Likes to reflect, play with ideas

____41. Has few artistic interests

____42. Likes to cooperate with others

____43. Is easily distracted

____44. Is sophisticated in art, music, or literature

### 12.3. Model B

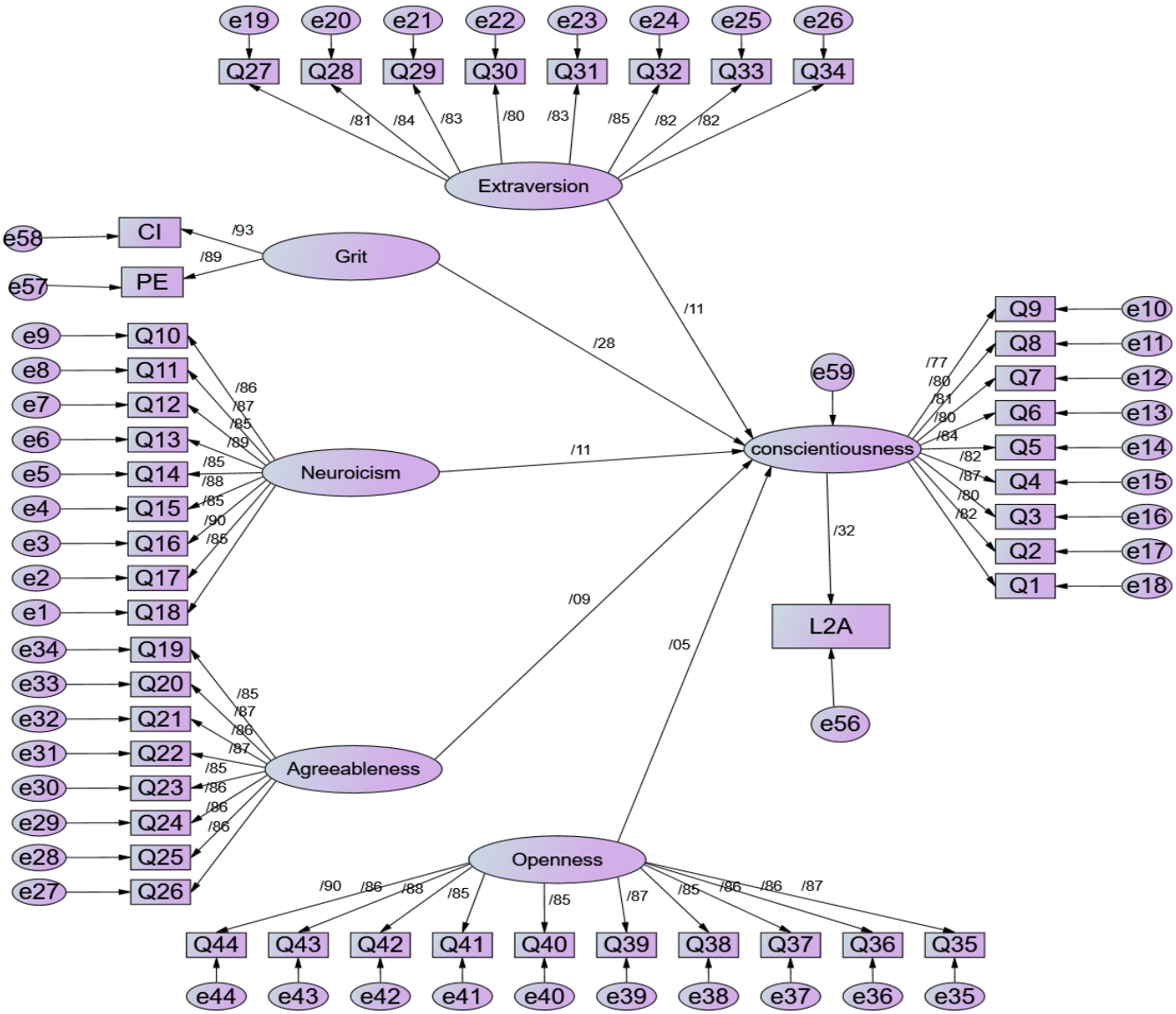



**Figure 6.** Research Model in the Case of Standardized Path Coefficients Mediated by Conscientiousness

## 12.4. Model C

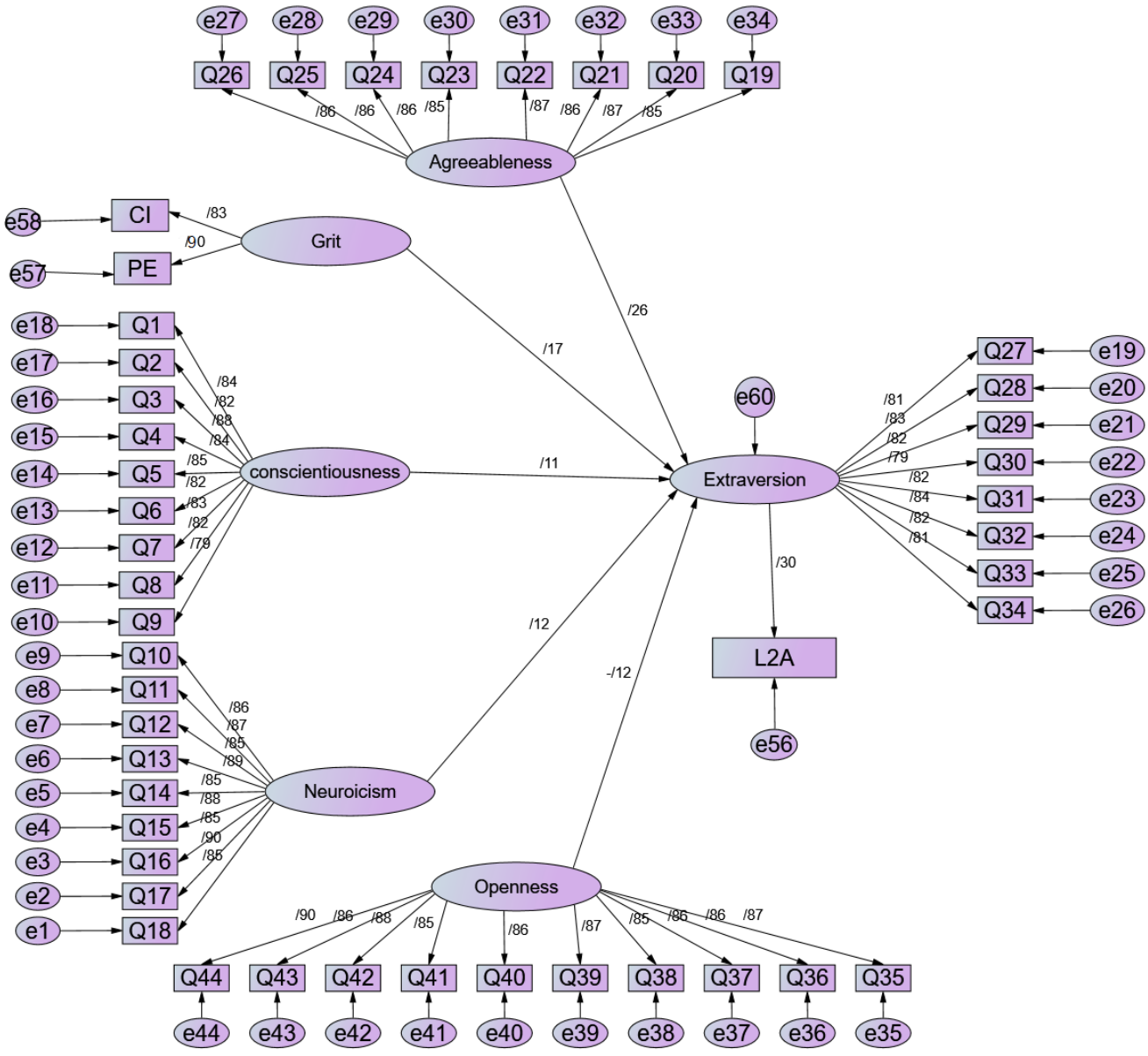



**Figure 7.** Research Model in the Case of Standardized Path Coefficients Mediated by Extraversion

### 12.5. Model D

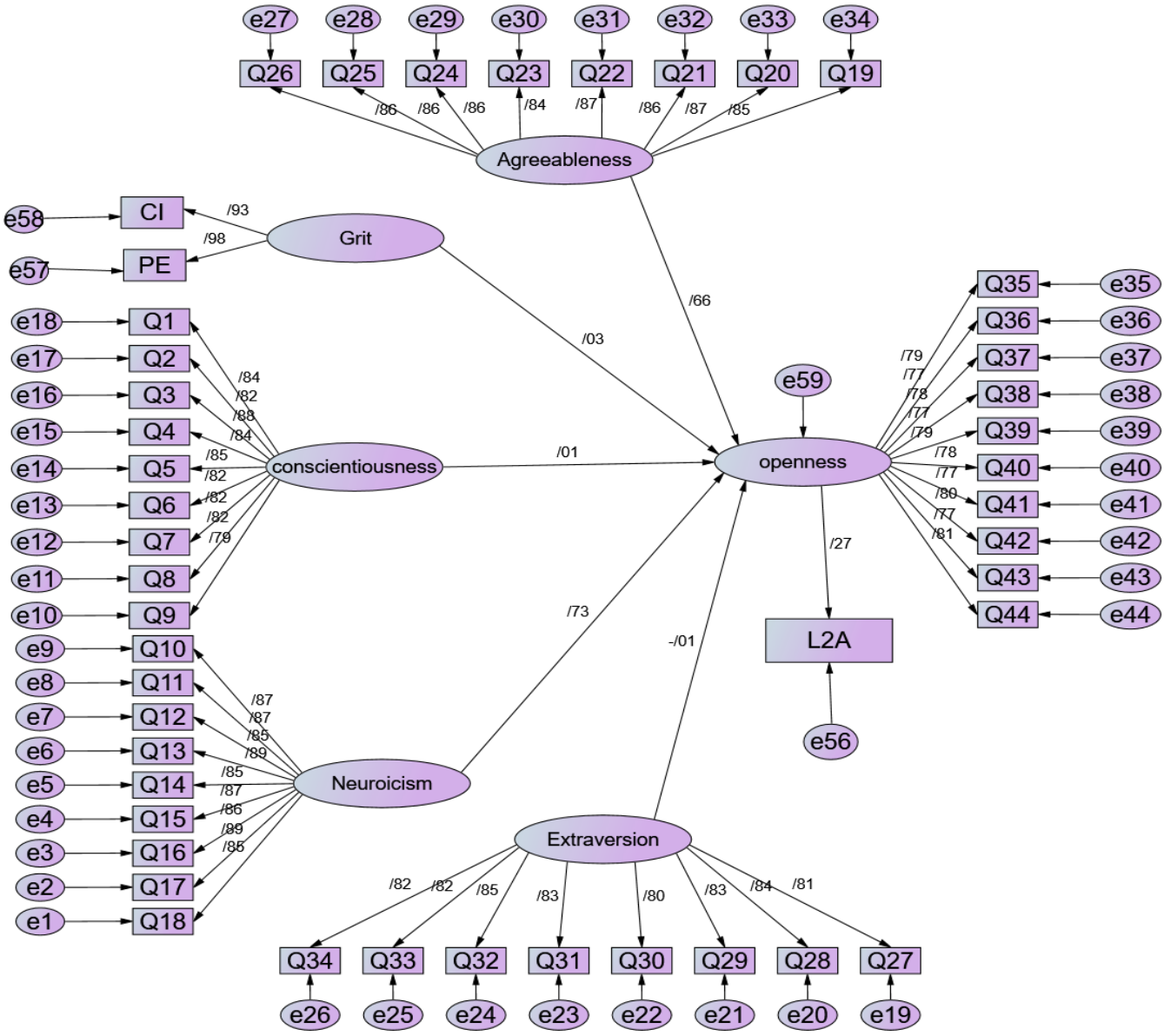



**Figure 8.** Research Model in the Case of Standardized Path Coefficients Mediated by Openness to Experience

## 12.6. Model E

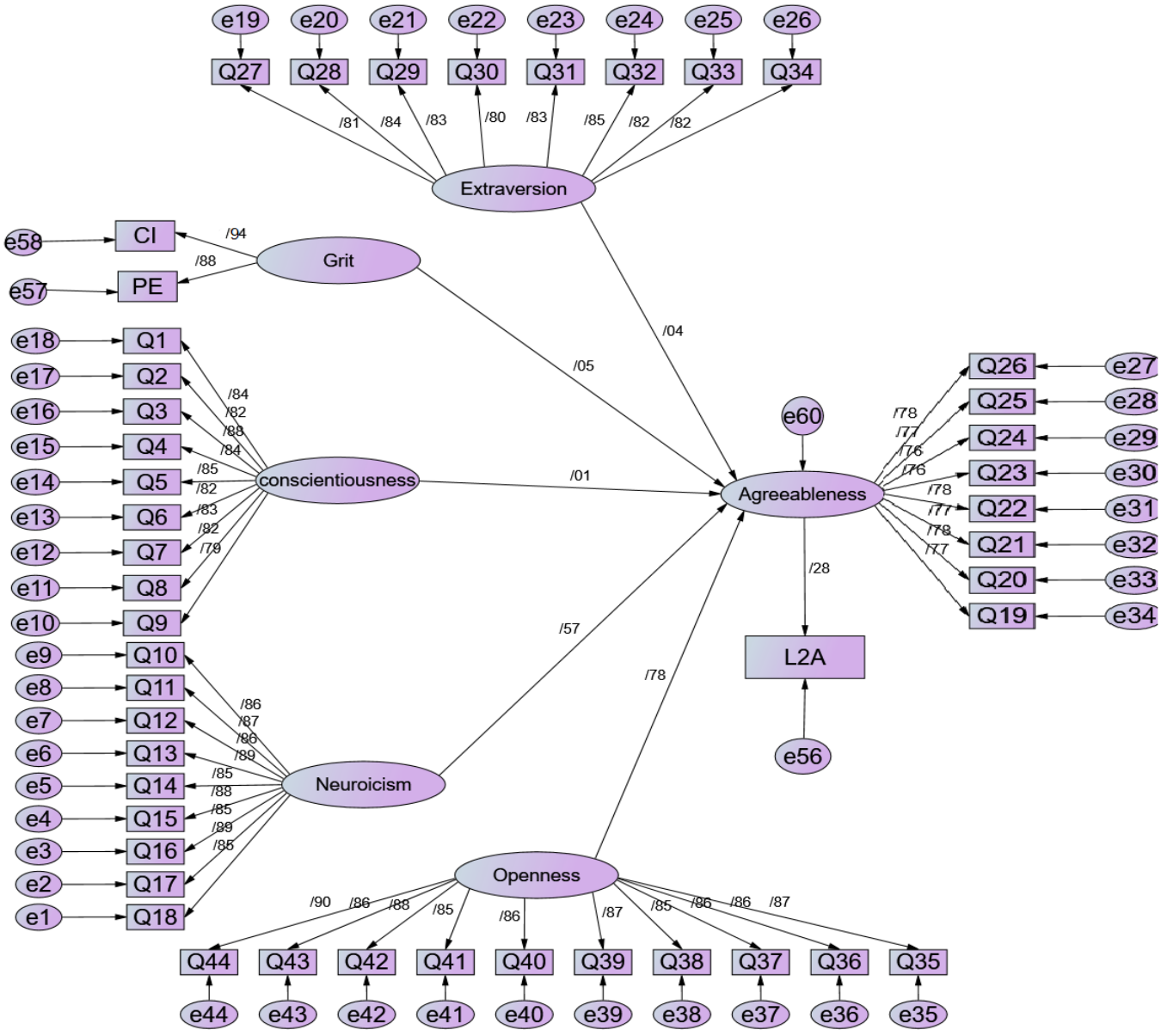



**Figure 9.** Research Model in the Case of Standardized Path Coefficients Mediated by Agreeableness

### 12.7. Model F

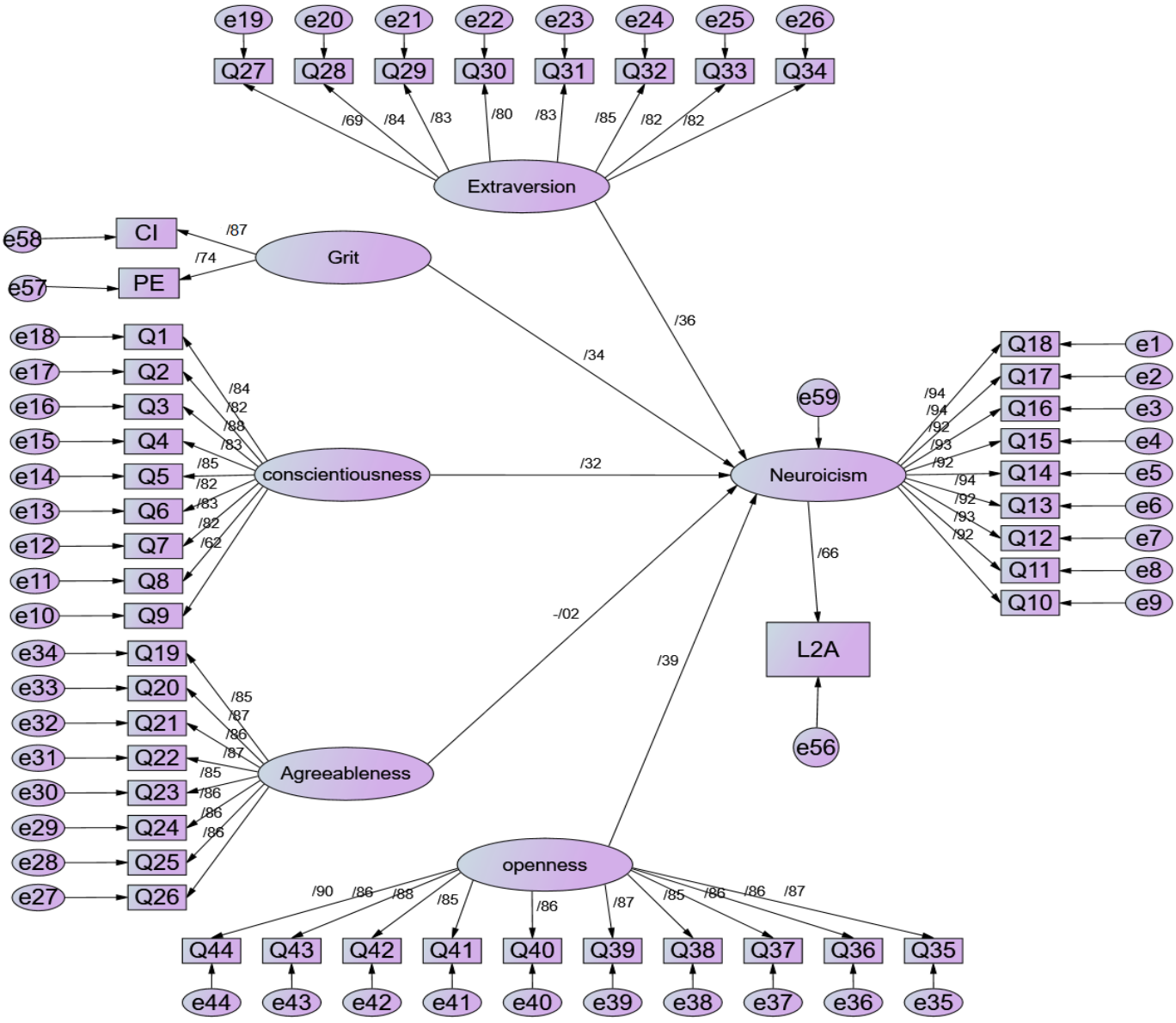



**Figure 10.** Research Model in the Case of Standardized Path Coefficients Mediated by Neuroticism