

University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

2021

The Washback Effect of Self and Peer Assessment in Second Language Writing

Musbah Abdussayed

Follow this and additional works at: <https://digitalcommons.memphis.edu/etd>

Recommended Citation

Abdussayed, Musbah, "The Washback Effect of Self and Peer Assessment in Second Language Writing" (2021). *Electronic Theses and Dissertations*. 2389.

<https://digitalcommons.memphis.edu/etd/2389>

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khhgerty@memphis.edu.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT IN SECOND
LANGUAGE WRITING

by

Musbah Bashir Abdussayed

A dissertation

Submitted in partial fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: English

Concentration: Applied linguistics

The University of Memphis

August 2021

Approved by:

Teresa Dalle, Ph.D., Chair

Emily Thrush, Ph.D.

Angela Thevenot, Ph.D.

Evelyn Wright Fogle, Ph.D.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

DEDICATION

I dedicate this work to my parents and my wife, whose patience and support had made me keep going until I finished this degree.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

ACKNOWLEDGMENTS

This research would not have been possible without the following persons:

Dr. Teresa Dalle, you were more than a committee chair; you were always there for me to sweep out obstacles and uncertainties. I cannot thank you enough.

My committee members: Dr. Emily Thrush, Dr. Angela Thevenot, Dr. Evelyn Wright Fogle, thank you for your time and supportive feedback throughout the dissertation process.

Saad Alamri, thank you for your endless support and your faithful friendship throughout the years of our graduate school.

Dr. Hatim Al Masoodi, thank you for the help you offered during your time at the University of Memphis.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

ABSTRACT

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT IN SECOND LANGUAGE WRITING

MUSABH BASHIR ABDUSSAYED

The influence of tests and assessments on learning and teaching is known as the washback effect. The purpose of this study is to investigate the washback effect of self and peer assessment in Arabic writing as a second language. The study was conducted on high school advanced Arabic learners at an Islamic school in the USA. The study examines the learners' attitudes and perceived capability towards self and peer assessment, their perceived importance of using self and peer assessment in revision, and their perceived effectiveness of taking the two assessments together. Moreover, the study seeks to reveal whether self and peer assessments have washback effects on revision, and whether considering peer writer's self-assessment in peer assessment has washback effects on learning writing. Throughout a mixed-methods approach, two-questionnaires, that is, pre-and post- and learners' interviews, were employed to elicit the data. The findings demonstrated that learners held more positive attitudes and perceptions towards the two assessments after taking the assessments. The study also concluded that self and peer assessment had mostly positive washback effects on revising writing. Some of the washback effects that appeared to be jointly caused by the two assessments included helping to meet the writing genre needs, filling writing gaps, and creating metacognitive awareness. Other effects that were found to be caused by either self or peer assessment included editing through self-assessment but learning new lexical items and becoming creative through peer assessment. It was also found that the use of a peer's self-assessment in peer assessment made peer assessors

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

reinforce and expand on peers' self-critiques, but that restricted their decision making.

Conversely, the lack of a peer's self-assessment in peer assessment resulted in independent and fresh peer assessments. Finally, the whole assessment process made learners obtain multiple feedback perspectives, getting more familiarized with their writing gaps, and having a more effortless experience doing a peer assessment with the consideration of the peer writer's self-assessment. Finally, the study suggests a constant examination of washback of alternative assessments (e.g., self and peer assessment) as those findings can detect the effective and ineffective sides of the assessment.

TABLE OF CONTENTS

| | |
|--|-------------|
| CHAPTERS..... | Page |
| LIST OF TABLES..... | xi |
| LIST OF FIGURES..... | xii |
| INTRODUCTION..... | 1 |
| My teaching and the current study..... | 1 |
| Assessment and Testing..... | 2 |
| Self and Peer assessment..... | 3 |
| Washback Effect..... | 5 |
| The rationale and the purpose of the study..... | 6 |
| Context of the study..... | 8 |
| Summary..... | 9 |
| LITERATURE VIEW..... | 12 |
| Overview..... | 12 |
| Definitions and concepts of washback effect | 12 |
| Washback contextuality and Dimensions..... | 14 |
| Washback 15 hypotheses of Alderson and Wall (1993)..... | 17 |
| The scope of the study..... | 19 |
| Mechanisms of Washback | 20 |
| Hughes's (1993) & Bailey (1996) Washback Models..... | 20 |
| Bailey's (1996) basic Model of Washback..... | 21 |
| Green's (2007) model of the direction, intensity, and length washback..... | 23 |
| Shih's (2007) Washback Model..... | 24 |
| Pan's (2008) Washback Model..... | 25 |
| Tzagari's (2009) Washback Model..... | 26 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | |
|---|-----------|
| Previous studies on the washback effect..... | 27 |
| Overview..... | 27 |
| Washback effect on learning..... | 29 |
| Attitudes and perceptions towards the test’s washback effect on learning..... | 31 |
| Washback effect of Alternative assessments and criterion tests..... | 34 |
| Self and Peer assessment in L2 Writing..... | 37 |
| Self-assessment..... | 37 |
| Peer assessment..... | 38 |
| Attitudes and perceptions towards Self and Peer Assessment..... | 39 |
| Self and Peer Assessment together in Second Language writing..... | 40 |
| Summary..... | 41 |
| Conceptualization of the Current study | 41 |
| METHODOLOGY..... | 43 |
| Overview..... | 43 |
| Purpose and Significance..... | 43 |
| Research questions..... | 44 |
| Research Design..... | 45 |
| Overview..... | 45 |
| Qualitative and quantitative research..... | 45 |
| Mixed method..... | 45 |
| Data collection..... | 46 |
| Self and peer assessment forms..... | 46 |
| Pre & post questionnaires..... | 47 |
| Semi-structured interviews..... | 48 |
| The writing task..... | 49 |
| Setting..... | 50 |
| Procedure..... | 52 |
| Data analysis..... | 55 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | |
|---|----|
| Theoretical underpinning..... | 55 |
| Quantitative Data Analysis..... | 56 |
| Qualitative data Analysis..... | 56 |
| RESULTS..... | 57 |
| Overview..... | 57 |
| The washback effect on learners’ attitudes..... | 58 |
| Learners’ feelings about the assessments before taking the assessments..... | 58 |
| Pre-perceptions of learners’ capability to self and peer assess..... | 60 |
| Learners’ feelings about assessment after taking the assessments..... | 61 |
| Post perceptions of learners’ capability to self and peer assess..... | 62 |
| Attitude changes from pre- to post-questionnaire..... | 63 |
| Change of learners’ feelings about self and peer assessment..... | 64 |
| Feeling more confident..... | 65 |
| Feeling less nervous..... | 66 |
| Feeling entertained..... | 67 |
| Changes in learners’ perceived capability to self and peer assess..... | 67 |
| Difficult to peer rate..... | 69 |
| Learners’ perceived importance of self and peer assessment, and the washback effects on revision..... | 70 |
| Pre-perceived importance of self and peer assessment to revise writing..... | 71 |
| Post perceived importance of self and peer assessment to revise writing..... | 72 |
| Changes in the Perceived importance of self and peer assessment..... | 74 |
| Difficult to objectively self-assess writing..... | 76 |
| The washback effects of self and peer assessments on learners to revise their writing..... | 77 |
| Washback effects of self and peer assessment on revision..... | 77 |
| Helping to meet the writing genre requirements..... | 78 |
| Bridging writing gaps through reading..... | 79 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | |
|--|-----|
| Editing through self-assessing and expanding lexicon reservoir through peer assessment..... | 81 |
| Becoming more creative through peer assessment..... | 83 |
| Building metacognitive awareness..... | 84 |
| Learners' perceived effectiveness of using self-assessment with peer assessment and its washback effects on learning..... | 85 |
| Pre-perceived effectiveness of considering and not considering the results of a peer writer's 'self-assessment in peer assessment..... | 86 |
| Post-perceived effectiveness of considering or not considering the results of peer self-assessment in peer assessment..... | 88 |
| Change in the perceived effectiveness of the results of a peer's self-assessment in peer assessment..... | 90 |
| Washback effects of considering the results self-assessment in peer assessment..... | 93 |
| Reinforcing peer's self-critique..... | 93 |
| Expanding on peer's self-critique..... | 94 |
| Restricting peer's decision making..... | 95 |
| Washback effects of not considering the results self-assessment in peer assessment | 96 |
| Independent peer's assessment..... | 96 |
| Fresh peer's perspective..... | 97 |
| The washback effects of self and peer assessment process..... | 98 |
| Getting multiple feedback perspectives..... | 98 |
| Getting more familiarized with self-writing gaps..... | 99 |
| Making peer assessing easier..... | 100 |
| Summary..... | 101 |
| DISCUSSION | 103 |
| Overview..... | 103 |
| RQ1: What are the participants' pre and post attitudes towards self and peer-assessment?..... | 103 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | |
|--|-----|
| RQ 2: The importance of self and peer assessment in revising second language writing: | |
| a: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing? | |
| b: Are there any washback back effects of self and peer assessment on the participants to revise their writing..... | 106 |
| R3: The effectiveness of considering and not considering the peer's self-assessment results in peer assessment: | |
| a: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment? | |
| b: Are there any washback effects of considering and not considering peer's self-assessment results in peer assessment? If so, what are they?..... | 109 |
| CONCLUSION..... | 113 |
| Overview..... | 113 |
| Purpose..... | 113 |
| Methodology..... | 114 |
| Findings..... | 115 |
| Significance | 117 |
| Limitations..... | 118 |
| Recommendations for future research..... | 118 |
| Ending..... | 119 |
| REFERENCES..... | 121 |
| APPENDICES..... | 135 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

LIST OF TABLES

| | |
|---------------|----|
| Table 1..... | 52 |
| Table 2..... | 59 |
| Table 3..... | 60 |
| Table 4..... | 62 |
| Table 5..... | 63 |
| Table 6..... | 64 |
| Table 7..... | 68 |
| Table 8..... | 71 |
| Table 9..... | 74 |
| Table 10..... | 75 |
| Table 11..... | 87 |
| Table 12..... | 90 |
| Table 13..... | 92 |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

LIST OF FIGURES

| | |
|---------------|----|
| Figure 1..... | 22 |
| Figure 2..... | 23 |
| Figure 3..... | 25 |
| Figure 4..... | 26 |
| Figure 5..... | 27 |

The Washback Effect of Self and Peer assessment is Second Language Writing

Chapter 1: Introduction

My teaching and the current study

Teaching second language writing is my preferred language teaching area. I am always eager to update myself with the teaching techniques that make my writing classes effective and purposeful. One of the effective writing methods I usually adopt is making my students develop a writing process rather than following one-time writing. At one stage throughout the writing process-could happen more than once- learners are situated to take the teacher's role to peer-assess writing products; learners may also be assisted to self-assess their writing (Larson, 2000). From my humble teaching experience, I would notice my students improve their writing when they had their writing either self or peer-assessed; this pushed me to dig deeper in self and peer assessments in second language writing.

The research portrays self and peer assessments as transferable tools that can be applied in multiple learning areas. Advocates of communicative language teaching have highly recommended utilizing self and peer assessment practices in teaching writing. For example, it is argued that self-assessment promotes self-efficacy in writing (McCarthy et al., 1985) and learning autonomy that enables learners to bridge their writing gaps (Bardine & Fulton, 2008). As for peer assessment, for instance, it promotes classroom communication and collaboration (Hu & Lam, 2010) and strengthens audience awareness (Liu and Hansen, 2002). The benefits list of self and peer assessment in teaching writing is long. I am only bringing up only a sample for the sake of the context. Thus, after realizing the significance of these two-teaching tools, I heavily implemented them in my writing classes. Over time, I realized that my second language

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

learners, in some cases, requested classifications or expressed some explanations about some points (e.g., instructions, assessment process, etc.).

Moreover, my learners would propose ideas about the assessment practices. Their ideas made me develop a research idea to investigate how the assessment influences learners and how an assessment designer can improve the assessment's effectiveness by considering learners' viewpoints. This scenario led me to consider the washback effect phenomenon, which is simply the assessment's influence on learning and teaching (Alderson & Wall, 1993).

Assessment and Testing

The terms “assessment and testing” are sometimes used interchangeably to refer to examinations. For the purpose of this study, it is worth highlighting the difference between the two terms. To put it simply, the test means to measure learners on a specific amount of knowledge to judge their performance eventually. On the other hand, assessment refers to documenting and monitoring the learning process to enhance the learning performance. In second and foreign language teaching, tests are associated with many evaluation forms (e.g., standardized and high-stake type of evaluation). In contrast, assessments are associated with alternative types of evaluation, known as alternative assessments).

Language assessment is undoubtedly a vital part of language learning and teaching in the educational setting in general. Assessments contribute vastly to the teaching process (James, McInnis, and Devline, 2002). They are the “engine” that drives learning Cowan (1999). A new language assessment movement has come to replace the traditional testing school; it has provided various approaches that focus on examining what learners can do with what has been taught to them. The purpose behind designing, adopting, and adapting assessment materials is to

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

promote learning through communicative assessment tools. The theory of communicative language teaching and the approaches it has introduced to the field of second language teaching promotes the concept that assessment should contribute to teaching and learning. In other words, assessments should be introduced to learners as part of a communicative learning experience. Alternative assessments are a good choice for those who advocate teaching language communicatively; this is because alternative assessments promote communication, interaction, negotiation, and task-based teaching. Unlike traditional testing, assessments are not geared towards comparing learners' performance. Instead, those tasks document learner' progress and growth over time (O' Malley and Valdez Pierce, 1996), focus on the learners' strengths rather than their weaknesses, consider the learners' learning styles and cultural background (Tannebaum (1996), and reflect real-life practices (Stiggins, 1987). Alternative assessments are communicative because they can be incorporated into the school and classroom daily activities (Hamayan 1995) and be utilized to collect data about how learners are advancing and completing authentic tasks rather than comparing students' performance at a particular point in the learning process.

Self and Peer assessment

Self-assessment is also called self-evaluation, self-reflection, and self-critiquing (Matsuno 2009), and all these terms tend to be used interchangeably (Andrade & Du (2007). For the purpose of this study, I use the term “self-assessment.”. Moreover, I use other terms related to self-assessment, such as self-rating, to refer to the action of rating. Peer assessment is also called peer evaluation, peer feedback, peer review (Zou et al., 2018), and peer response (Caulk,

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

1994). However, for the purpose of this study, I use the term “peer assessment,” but I also use other terms like “peer feedback” to refer to the action of giving feedback and comments.

Self and peer assessments are two types of alternative assessments that are often used in second language teaching, but peer assessment is more prevalent in second language writing. Both assessments are valuable assessment tools for those who advocate communicative and active learning as they provide learners with opportunities to be actively engaged in the assessment process. In second language writing, self-assessment has been proven to enhance the writing performance of learners: it activates skills such as self-regulatory learning (Bing, 2016; Roose, 2006) and self-awareness skills (Ashraf & Mahdinezhad, 2015) about what the learner can /cannot do in writing. Likewise, peer assessment, being a popular alternative assessment tool in second language writing, provides active learning opportunities to learners: learners take the teacher’s role to assess their peer's performance, and learners have chances to engage in active learning through negotiations with their peers.

Engaging learners in the assessment process through self and peer assessment practices leads to the best possible learning gains by teaching learners how to be successfully and effectively involved in the assessments (Anderson, 2001). Hence, after designing, adopting, and adapting self and peer assessments, teachers need to make sure that learners understand the purpose behind using them and implement them. This way ensures that the learner self and peer assess writing performance to provide rich data on how learning and teaching occur.

However, even though teachers design, select, adopt, and adapt the “well-designed” assessment tools, there is always a need to hear from learners to be aware of what they think of the assessment; this enables the test designer to keep and enhance the good elements of the assessment (questions items, instructions, process, time, language), and heal the weak sides of

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

the assessment. Learners' perspective is critical to consider because their attitudes and perception can determine their engagement degree. Therefore, incorporating the washback effect (the assessment's influence) can reveal those perceptions, attitudes, and experiences.

Washback Effect

The washback effect refers to the test's influence on teaching and learning (Alderson & Wall, 1993). The phenomenon has many terms used in various educational settings, but in the context of language assessment and testing, the two terms washback and backwash are interchangeably coexistent to refer to the influence of the test/ assessment on learning and teaching (Alderson & Wall, 1993). Nevertheless, for this study's purpose, the term washback effect is the term to be used throughout the study. The attention towards studying the washback effect of the tests resulted from the fact that high-stakes standardized tests determine and shape what and how teaching and learning may occur. In applied linguistics, the first pioneering research-based data on the washback effect of tests and assessments on teaching and learning was in the early 1990s (Bailey, 1996). The study was carried out by Alderson and Wall (1993), in which they introduced the first systematic work to explain the washback phenomenon; this study is now one of the most cited references to many washback studies. Alderson and Wall (1993) looked into the effect of the newly introduced high-stakes language tests in Sri Lanka on the teaching practices, and they concluded that a test would have a washback if the test is designed to have "important consequence"; conversely, a test is not expected to have washback effect if the test is not aimed to have "important consequences." This means that when stakeholders do not value the assessment, it is not expected to influence the participants.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Since large-scale and high stakes tests heavily affect learning and teaching, most washback studies focus on them (e.g., entrance exams, standardized tests, etc.). It is argued that those high-stakes tests influence teaching and learning at two levels: macro and micro. For example, high-stakes tests influence institutional policies, educational systems (macro effect), curriculum practices, students' learning strategies (micro effect). Moreover, many washback studies have concluded that tests can influence learning and teaching in many areas: students' learning behaviors (Pan & Newfields, 2012) or teaching practices (Qi, 2005). In short, "Tests are held to be powerful determiners of what happens in the classroom" (Alderson and Wall, 1993).

The rationale and the purpose of the study

As mentioned in the previous paragraph, the washback effect in applied linguistics is widely studied through large-scale and high-stakes tests (e.g., standardized, University examinations, etc.). This is because those tests have both micro effects (e.g., practices, learning behaviors, materials, etc.) and macro effects (e.g., policies) on teaching and learning. However, alternative assessments such as self and peer assessment can have effects on learning. The phenomenon of the washback effect of high-stake tests can also exist in alternative assessments and small-scale tests. In the body of applied linguistics, there have been few attempts to examine how alternative assessments and small-scale tests affect teaching and learning. There are some studies, such as by Hung (2012) in which he examined the effect of an e-portfolio project on content learning, and by Safa (2014), in which she investigated how task-based language assessment influences learning English grammar (TBLA). These studies and a few more demonstrated that the washback effect exists in these types of assessments.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Self and peer assessment in second language writing, as discussed above, are two practical evaluation tools to monitor learners' abilities to accomplish language tasks (Cheng, and Curtis as cited in Hung 2012). Thus, they influence learning and determine how learning may take place. Therefore, examining the washback effect of self and peer assessments is not any less beneficial than doing that with standardized and large-scale tests. It is important to examine what effects the assessments may have on learners to validate how they work. We, as teachers, do not certainly know the mechanisms of self and peer assessments in writing; we may just know they help, but we do not know for sure which items in the assessment work better or which items fail to work. Since it is essential to actively engage learners in self and peer assessment, it is fundamental to figure out what learners think, like, dislike, and prefer about the assessment; consequently, providing the teacher with a detailed description of how the assessment operates the learning process. The final result of this process is developing more effective assessments in the future. Hence, there is a need to examine the connection between the assessment and learning; here is where the assessment mechanism should be investigated to detect how the assessment practically works for the learners. A popular way to do this is by looking at the washback effect of the assessment.

When studying the washback effect of the test or the assessment, it is indispensable to examine the learners' perceptions, attitudes, learning practices, and learning behaviors. Many studies on self and peer assessments addressed learners' attitudes and perceptions towards the assessment to reveal what learners think of the assessment and why they think the assessment works or does not work. However, those studies did not elaborate on how the assessment affects learning (Green 2013) or what changes (good or bad) peer and self-assessment have made on learning second language writing. Moreover, most of the studies on self and peer assessments in

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

second language writing have examined the two assessments separately. A few studies have considered examining both self and peer assessment together. For example, Iraj et al. (2016) concluded that self and peer assessment harmoniously coexisted and contributed to learning. Black et al. (2004) stated that Self and Peer assessment can both be complementary assessment tools and that when self-assessment is administered before peer assessment, learners are enabled to determine their gaps and weaknesses. Then, they transfer those pitfalls to the peer assessment so that their peers are aware of what to focus on (Lee 2017).

With that all being said, this study seeks to investigate the washback effect of self and peer assessment on second language writing to reveal the influences on learners and learning.

Context of the study

This study investigates the washback effect of self and peer assessment on learning second language writing; it specifically examines Arabic heritage speakers in the USA and Muslims who learn Arabic as a second language in the USA. However, most learners are heritage speakers.

Heritage languages, broadly, are ethnic minority languages spoken around the world. There are at least two types of minority languages in the United States. The other type includes the languages spoken by immigrants who move to the country where another language denominates and represents the majority (Mark 2011). So, in the United States, Arabic speakers are categorized as heritage speakers. Heritage Arabic speakers from different Arabic-speaking countries speak varieties of colloquial Arabic at their homes and restricted local communities. Children acquire their parents' colloquial Arabic, but when they are sent to the Islamic schools, they start learning the Modern Standard Arabic, the version of Arabic taught in educational

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

institutions in all Arabic speaking countries; the reason behind this is the distinctive differences between Arabic versions in different Arabic speaking countries. Hence, learning the modern standard Arabic means the ability to communicate across borders.

Moreover, learning Arabic has a religious and cultural significance. Arabic language to Arabs is ethnic identity and cultural print; for Muslims (both Arabs and non-Arabs), Arabic is also a religiously sacred language. Muslims must perform their Islamic practices in Arabic. All prayers should be uttered in the original language (Arabic), and Quran (Islamic holy book) should also be recited in Arabic (the original language by which the book was revealed from God, as muslims believe). So, many Muslim families (Arabs and non-Arabs), who can afford to teach their children in private Islamic schools, make sure to enroll their children in those schools to acquire and learn Arabic as a second language and the Islamic practices. Those learners study Arabic in Islamic schools from kindergarten to K12, and Arabic is a mandatory subject that is studied every day at school; it is taught as a second language for 50 minutes a day, five days a week. (majority of Arabic learners in Islamic schools in the USA).

Summary

Self and peer assessment scaffold learners to autonomously discover their writing gaps and, consequently, strengthen their writing skills. Also, they help teachers and learners monitor the learning progress and inform what learners can do with the language tasks (Cheng and Curtis as cited in Huung 2012). In terms of second language writing, self and peer assessment have been proven to be effective learning tools; yet, there is sometimes a need to detect how assessments work to make sure assessments operate as they should. That is, there might be a need-to-know what effects self and peer assessments may have on learners. Thus, the study

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

considers the washback effect phenomenon to reveal what effects the assessments may have on learning and learners. As demonstrated in the introduction, the washback effect phenomenon is often associated with large-scale and high-stakes tests as those tests are known to have tremendous effects on teaching and learning at the micro-level (classroom setting) and the macro-level (social and institutional level). However, the effects of alternative assessments (e.g., self and peer assessment) in writing are also strongly existent at the micro-level (classroom level). Therefore, the study considers looking into the washback effect of self and peer assessments on learning Arabic writing as a second language in the USA. The study investigates whether self and peer assessments can influence learners and learning (what changes the assessments may create in learning). To do that, the study examines the learner's attitudes, perceptions, and learning experiences before and after taking the self and peer assessments. It also seeks to reveal the washback effect of using self and peer assessment versus only peer assessment. The research questions framing this study are as follows:

1. What are the participants' pre- and post-attitudes towards self and peer assessment?

2. The importance of self and peer assessment in revising second language writing:

A: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing?

B: What washback back effects do self and peer assessments have on the participants to revise their writing?

3. The Effectiveness of considering and not considering the peer's self-assessment results in peer assessment:

A: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

B: What are the washback effects of considering and not considering peer's self-assessment results in peer assessment?

Chapter 2: Literature View

Overview

This chapter is divided into two main parts. Part one is dedicated to the washback effect, and part two is devoted to self and peer assessment. Part one of this chapter clarifies the concept of washback effect of language testing; it also presents definitions of the phenomenon of washback effect. Section two of part one of this chapter demonstrates the idea of washback contextuality and dimensions. Section three introduces the theoretical underpinning of the phenomenon of washback effect in language testing as well as the scope of the study; this section (section three) also includes reporting several models that attempted to describe mechanisms of washback effect of testing on language learning and teaching. Section four reports several previous studies of the Washback effect of tests and assessments on language teaching and learning in second and foreign language contexts. The last section of this chapter lays out conceptual and theoretical justifications for the current study.

Part two of this chapter sheds light on utilizing self and peer assessment in second language writing: it starts with providing definitions, concepts, and some theoretical underpinnings to use self and peer assessment in second language writing. Then, this part (part two) of this chapter also reports, in general, what the research says about the perceptions and attitudes of participants concerning self and peer assessment in second and foreign language writing.

Definitions and concepts of washback effect

How the learning process occurs and its implications for learning and teaching processes can result in two critical questions in any educational setting. As tests and assessments are

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

central teaching and learning tools, they affect how teaching and learning occur; therefore, they have implications on learning. Alderson and Wall (1993) stated that “tests are held to be powerful determiners of what happens in the classroom.” This means that tests and assessments can structure what to teach, how to teach, and probably when to teach; subsequently, they structure what to learn, how to learn, and when to learn. This is what is known as the “washback effect.”

The definitions of washback effect in language testing and assessment are many. For example, in its simplest definition, washback effect is viewed as the influence of testing on teaching and learning (Alderson and Wall, 1993, p. 115) and /or an impact on teaching and learning (Buck, 1988, Alderson & Wall, 1993). The washback effect is also seen as the influence of the test on teaching and learning “done in preparation for it” (Green, 2013), a connection between teaching and learning (Hamp-Lyons, 1997), and interaction between teaching and learning (Green, 2013), “the extent to which the introduction and use of a test influences language teachers and learners to do things they would not otherwise do that to promote or inhibit language learning (Messick, 1996), and an educational reforming tool that leads to making necessary changes on learning and teaching (Shoamy, 1992).

The terms: “test impact” (Baker, 1991), “system validity” (Frederiksen & Collins, 1989), “consequential validity” (Messick, 1989, 1996), “measurement-driven instruction” (Poham, 1987), or “curricular alignment” (Madaus 1988; Smith, 1991) are associated with the influence of the test in educational settings. As for the context of language testing and assessment, the terms “washback” and “backwash” are interchangeably used to indicate the influence of the tests and assessments on teaching and learning (Hughes, 1989; Alderson & Wall, 1993; Bailey, 1996; Saville et al. 2004; Alderson, 2004; Cheng & Curtis, 2004).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Two types of washback effects can result in positive or negative effects (Bachman & Palmer, 1996; Brown Hudson, 1998). The positive washback effect of tests and assessments occurs when the test promotes “desired changes” on the curriculum, teaching practices, and learning outcomes. Conversely, the negative washback effect occurs when those “desired changes” do not appear (Hung, 2012, p. 27). However, evaluating the relationship between the test or the assessment and the washback effect is not as simple as it sounds (Brown & Hudson, 1998, p. 667); it entails considering many relevant factors such as features determined by the test or the assessment maker, the characteristics of the test or assessment taker, and the decision points resulting from the test or the assessment (Cheng, 2004, p. 25). The way the test, the assessment, or the language task is designed is significantly influential concerning the quality of the resulted washback effect, whether good or bad (Bailey, 1996, Hughes, 1989).

Whether the test's washback effect is “good” or “bad,” they are sometimes unintended. Messick (1989) stated that tests result in both “unintended and intended” washback effects. Intended effects are usually associated with the “positive effect,” while the unintended effects can be either positive or negative. What determines “good” and “bad” requires considering the educational and social setting where the test is taken. Messick (1989) wrote: “Judging validity in terms of whether a test does the job it is employed to do . . . requires evaluation of the intended or unintended social consequences of test interpretation and use.”

Washback contextuality and Dimensions

The dimensions of the washback effect of language tests and assessments in second language teaching have always been various and distinctive from one setting to another. This is simply because the washback effect is contextual in nature (Green, 2103) and can be investigated

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

through different scopes as the intentions and implications of tests and assessments are usually geared by the stakeholders' interpretations. Tests and assessments are contextually perceived to have different roles as stakeholders may have different beliefs towards those assessments and tests. Those roles and effects can also be interpreted differently at different dimensions.

Educationally, there are different views over the terms “washback” and “impact” concerning the dimension of the tests' influence. Wall (1997), for instance, views the test's washback to be merely tied to the test's effects on teachers and learners' behavior in the classroom. In contrast, the test's impact refers to broader effects, such as their influence on curriculum practices, school policies, and educational systems in the society (p. 100). Nevertheless, for language testers and assessors, washback and impact are one dimension; the first refers to the test's micro effect, while the latter refers to the test's macro effect (Tsagari, 2009; Bachman and Palmer, 1996).

As mentioned in the above paragraph, washback effect can be interpreted variously and contextually; likewise, so can the dimensions of the washback effect. In an attempt to form a common platform on what to look at when analyzing the dimensions of washback effect, Cheng & Curtis (2004) stated five dimensions of washback effect of tests and assessment:

1. The specificity of the washback: this refers to the test's effect, whether it is general (effects of tests in general) or specific (effects of a particular test).
2. The intensity of the washback: the effect can be strong or weak. The strong effect may mean that the test heavily dominates what and how learners learn and what and how teachers teach. Whereas weak effect means that the test partially influences what takes place in classrooms and partially influences what, how, and when materials are taught, and thus, it may affect some students and teachers, but not others.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

3.The length of the washback: this is concerned with how long the effect lasts on learners.

4.The value of the washback: the value is associated with the negativity or passivity of the test's effects on learning and teaching. The intended washback effect is usually attributed to the positive washback effect, while the unintended can be a positive or negative washback effect.

5.The intentionality of the washback: this indicates the targeted effects, as well as the accidental, untargeted washback effects.

The contextuality of tests and assessments is also dependent on other areas to consider. Green (2013) suggested looking into the roles of the tests and assessments in the light of some “issues of particular relevance”:

1.Setting: Who are the key participants, and where the test or assessment is used? What are the benefits associated with the test to the participants?

2.Test use: Do the participants equally or variously value the test? What kind of stakes the test the assessment has, and how are they perceived to provide for the participants? How hard is the test taught to be?

3.Knowledge participants: how much knowledge the participants have about the test or assessment. What misconceived ideas do the participants have about the test or assessments?

4.Resources: what resources do the participants have to have for the test or assessment? What resources are existent? Do the instructors have the necessary training in “requisite language skills” and pedagogical skills?

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

5. Beliefs about testing: what is familiar to the participants about assessment and test?

What is the contribution of the test or assessment in the participants' lives? How participants respond to the utilization of the test or the assessment.

6. Interactions between participants: how do participants get educated on the test or the assessment? What knowledge do participants transmit to the other participants? How do other participants push them to get prepared for the test or assessment? (Green 2013, p. 46-47).

Washback 15 hypotheses of Alderson and Wall (1993)

When examining the test's washback effects, it is essential to distinguish between the overlapping potentially affected areas of teaching and learning. Alderson and Wall (1993) proposed 15 systematic hypotheses in order to draw lines between those prospective affected areas (see table 2.1). These proposed 15 hypotheses succeeded in explaining the complexity of the phenomenon by differentiating between the effects of the test on stakeholders' attitudes and the effects on the content of learning and teaching. On the other hand, they differentiated between the impacts on teaching methods, teachers, learning, learners, and impacts on the process of teaching and learning (Green, 2013, p. 42).

The hypotheses came out as a result of a washback study entitled "Does washback Exist?" (Alderson and Wall 1993), in which the researchers investigated how a newly introduced large scale test in Sri Lanka affected the areas of teaching (e.g., the methods of teaching, and curriculum). The hypotheses found by Alderson and Wall (1993) are as following:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- A test will influence teaching.
- A test will influence learning.
- A test will influence what teachers teach.
- A test will influence how teachers teach.
- A test will influence what learners learn.
- A test will influence how learners learn.
- A test will influence the rate and sequence of teaching.
- A test will influence the rate and sequence of learning.
- A test will influence the degree and depth of teaching.
- A test will influence the degree and depth of learning.
- A test will influence attitudes to content, method, etc., of teaching/learning.
- Tests that have important consequences will have washback
- Tests that do not have important consequences will have no washback.
- Tests will have a washback on all learners and teachers.
- Tests will have washback effects for some teachers and some learners, but not for others.

Alderson & Wall (1993), p.p. 120-121

The dependent variables these 15 hypotheses introduced, including methods, rate, content, sequence, degree of teaching, and learning, have been a guideline for washback studies of second and foreign language tests (Green 2007a). Those variables can help make predictions about the following: content (what), methods (how), rate, sequence, degree of depth of learning, and teaching (Green 2013, p.47).

Alderson and Hamp-Lyons (1996) revisited the theory and expanded on the 15 hypotheses suggesting that washback effect may also vary based on these:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

1. The stake of the test.
2. How much the test is in harmony with the current practices.
3. The beliefs held by the teachers and textbook designers' beliefs concerning the suitable methods to prepare for the test.
4. The creativity of teachers and textbook designers.

The scope of the study

In language testing, the terms “washback” and “backwash” are interchangeably used to indicate the influence of the assessment on teaching and learning (Hughes, 1989; Alderson & Wall, 1993; Bailey, 1996; Saville, 2004; Alderson, 2004; Cheng & Curtis, 2004), but for the purpose of this study, the term “washback” is used to refer only to the effect of self and peer assessment. The study only seeks to investigate the washback effect of self and peer assessment on learning and learners in second language Arabic writing; it does not consider the effect on teaching and teachers. Another prominent feature of this study is the adaptation of examining the washback effect of large-scale and high-stakes tests to examining the washback effect on alternative assessments, namely self and peer assessment. As previously introduced in the introductory chapter, most washback studies in the literature body are associated with large-scale and high-stake tests. Only a few works shed light on alternative assessments' washback effect on learning (Hung, 2012; Burksaitiene & Tereseviciene, 2008).

The dimension of washback effect in language testing is one dimension with two levels: micro and macro effect (Tzagari, 2009; Bachman & Palmer, 1996). Since peer and self-assessment are criterion assessments (not standardized), the current study merely looks into the micro effect of the assessment on learning and learners.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The theory of washback effect proposed by Wall & Anderson (1993) introduced 15 hypotheses that simplify the four areas of learning and teaching that the test may affect. Those potential four areas to be affected are the content, method, attitudes, and teaching and learning process. As this study is only dedicated to examining the washback effect on learning and learners, only five variables from the 15 hypotheses are relevant to the purpose of this study:

- A test will influence learning.
- A test will influence how learners learn.
- A test will influence the degree and depth of learning.
- Tests that have important consequences will have a washback.

(Alderson & Wall,1993, pp.120-121)

Mechanisms of Washback

This section views several popular washback models, namely Hughes' (1993) model, Bailey's (1996) basic model, Green's (2007) model of the direction, intensity, and length of washback, Shih's (2007) Washback Model, Pan's (2008) Washback Model, Tsagari's (2009) Washback Model. These selected models have proposed ideas to clarify and simplify the complex mechanism of washback effect.

Hughes's (1993) & Bailey's (1996) Washback Models. Hughes (1993) categorized the areas affected by the test into three areas:

1. Participants, such as learners, instructors, and policymakers.
- 2.Processes, such as learning & teaching, learning & teaching practices, and curriculum-making procedures taken by instructors, learners, curriculum designers, policymakers.
- 3.Products, such as test and assessment scores, learning and teaching materials.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

This model suggests having higher chances of washback effect. Hughes (1993) stated that the following conditions should surround the test to have positive washback effects:

- 1.Learners should have an incentive to succeed on the test.
- 2.Teachers should have an incentive to have their learners succeed on the test.
- 3.Participants should comprehend the nature and the content of the test and what implication they may cause.
- 4.Participants should be knowledgeable about the test's demands (e.g., teaching methods, syllabus design, etc.).

(Hughes, 1993, p.2-3).

Bailey's (1996) basic Model of Washback. Bailey (1996) based her model (Basic Model of Washback) on Hughes (1993) and Alderson and Wall's 15 hypotheses; she demonstrated, as shown in figure (1), the relationship between the three components (participants, process, and product).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

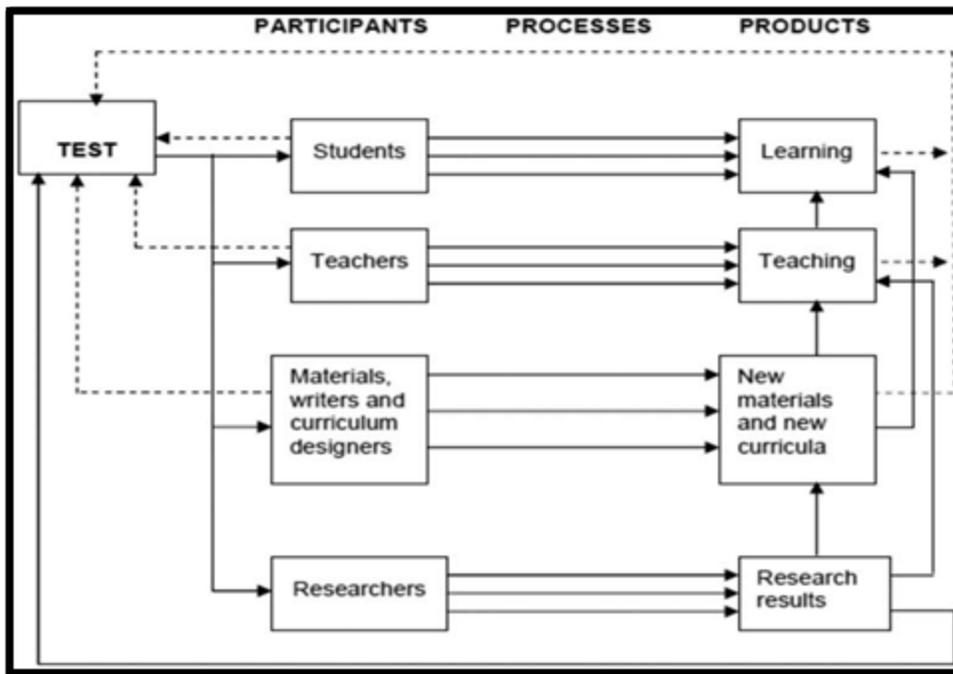


Figure 1 Basic Model of Washback (Bailey, 1996, p. 264).

Bailey (1996) made a distinction between the participants as:

1. Students and teachers.
2. Material writers and curriculum writers or designers.
3. Researchers.

The model shows how, for instance, the test influences teachers' and students' attitudes and perceptions, and eventually, will have a behavioral effect towards the test. Those behavioral effects could mean, for example, the way participants prepare for the test. However, one shortcoming that can be pointed out from the model is that the “process” is not given much description as the “participants” and the “product” are.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Green's (2007) model of the direction, intensity, and length of washback. The direction of Green's (2007) model refers to the passivity and negativity of washback determined by the overlap between the test characteristics and the social construct and determined by how those test characteristics and social constructs are conceptualized by participants. As for the availability, the model claims that it is ruled by the participants' beliefs, values, and knowledge; thus, the washback effect may appear to be different according to the social and cultural contexts. Finally, the intensity of the washback, as shown in this model, is taught to be dependent on how a test's importance and difficulty is perceived by the participants. If the test is perceived to be important and challenging, then the test will result in a positive washback effect and vice versa. While the model directly relates to both participants' perception of the test with the test's difficulty and importance.

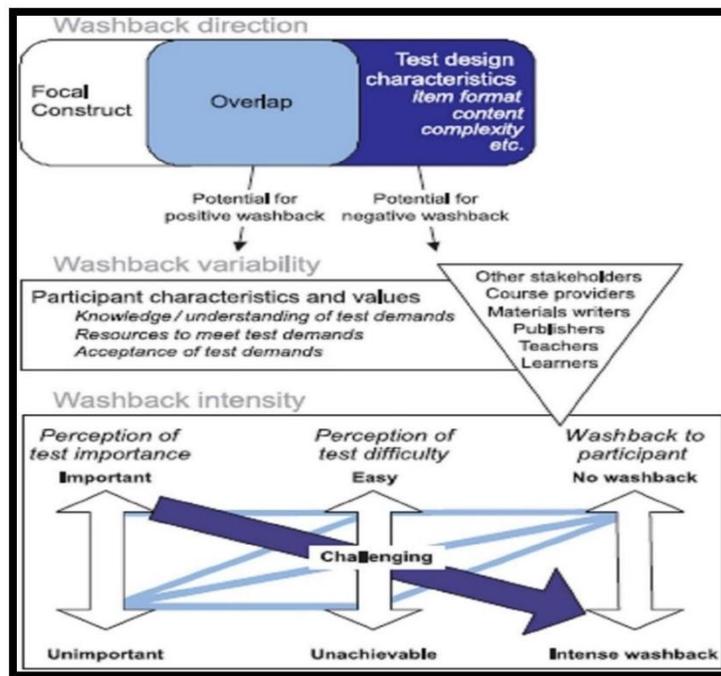


Figure 2 Model of Washback direction, variability, and intensity (Green, 2007, p.24).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Green's (2007) model came out as a result of a study that aimed to look into the role of test preparation to enhance students' IELTS scores. Green (2007) investigated the participants and the variables attributed to the "process," such as the motivation, teaching/learning resources, and beliefs; it was concluded that that in order for students to improve scores on a test, covering the materials on the test should be a part of the normal teaching practices rather than relying on the instructions of the test. Moreover, participants, namely, students and teachers, should be aware of the purposes of the test. Figure (2) outlines the findings of this study and shows what components claimed to be influential on the learning products.

Shih's (2007) Washback Model. Shih's (2007) model classifies several contributing factors to the washback effect. Shih (2007) labeled the socioeconomic and educational factors as "extrinsic factors"; while factors such as individual differences and personal perceptions are referred to as "intrinsic factors." This model attempted to show and explain how these various factors are connected to produce a washback effect.

The model clearly describes how extrinsic, intrinsic, and test factors impact the test. As can be seen in figure (3), the solid line arrows refer to the influence expected to take place. The dotted line arrows indicate possible influences. Finally, the "Time axis "indicates the importance of the time factor as washback can evolve over time. Yet, one may argue against the "test factors" as there are some items in this category of the same nature (e.g., test content and test structure) that may lead to similar impacts.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

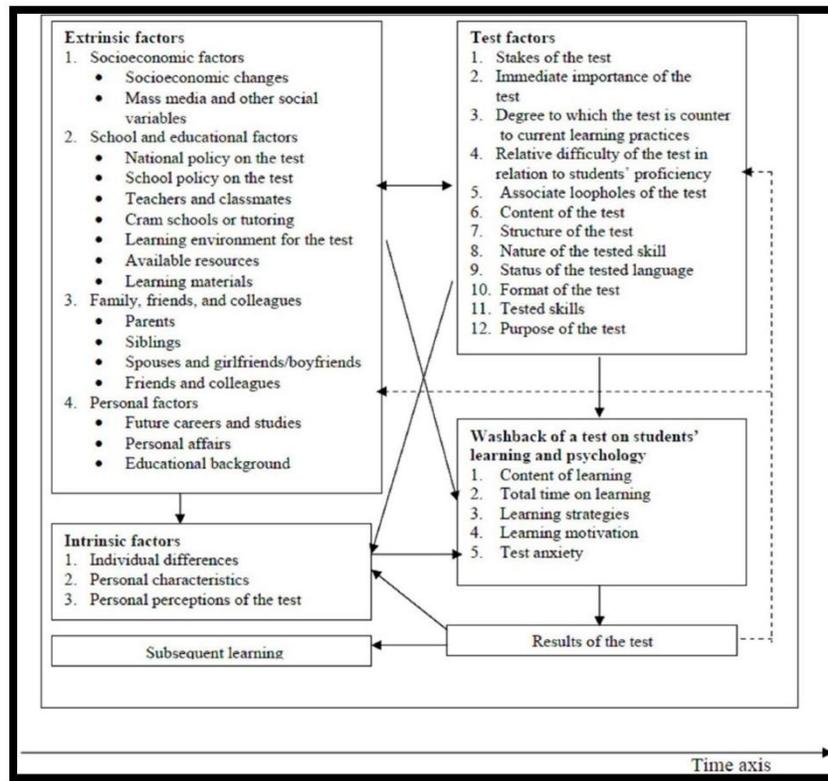


Figure 3 Washback Model (Shih, 2007, p.151)

Pan's (2008) Washback Model. Pan (2008) aimed to elaborate on the three test components (participants, process, and product) proposed by Hughes's (1993). The model also considers Alderson and Wall's (1993) 15 proposed hypotheses. Pan's (2008) model categorizes the washback effect as "macro" and "micro" levels; the first refers to the wide social and educational dimension while the latter indicates the classroom learning and teaching dimension. The model emphasizes that the test does not only affect the participants' perceptions, attitudes but also teaching and learning practices. Therefore, participants can properly address the test's demands.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

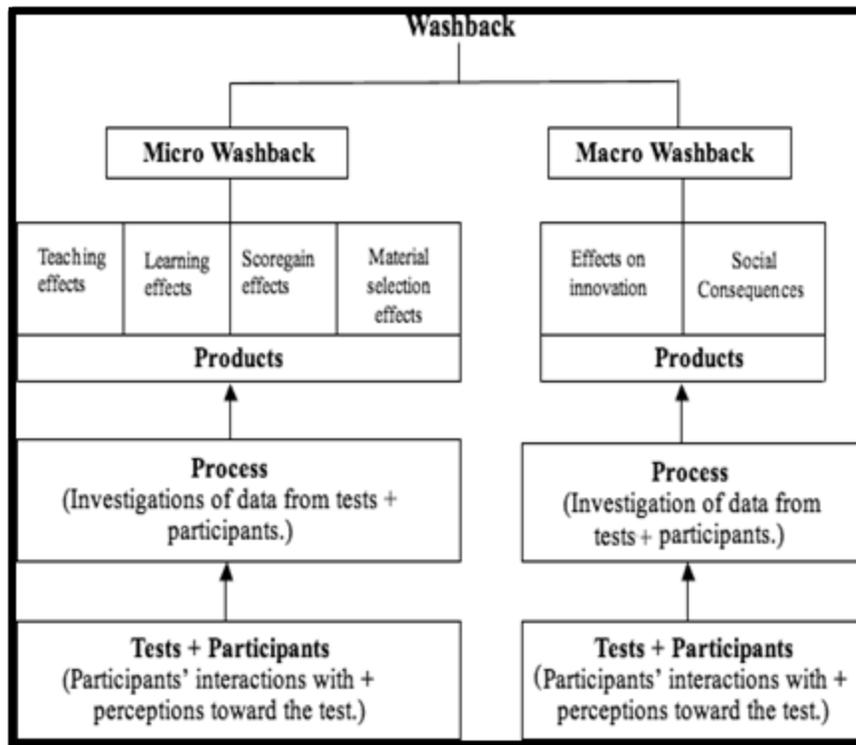


Figure 4 Pan's (2008) Washback Model

The model, as can be noticed from figure (4), combines the two levels of washback effect (macro and micro), aiming to simplify how “participants, process, and products” work interactively on those two levels.

Tsagari's (2009) Washback Model. Tsagari's Washback Model (2009) shows how the washback effect works in a circulated process in which the testing process goes through the different social and contextual settings; it also shows the relationship between the test participants in those ecological settings (see figure 5).

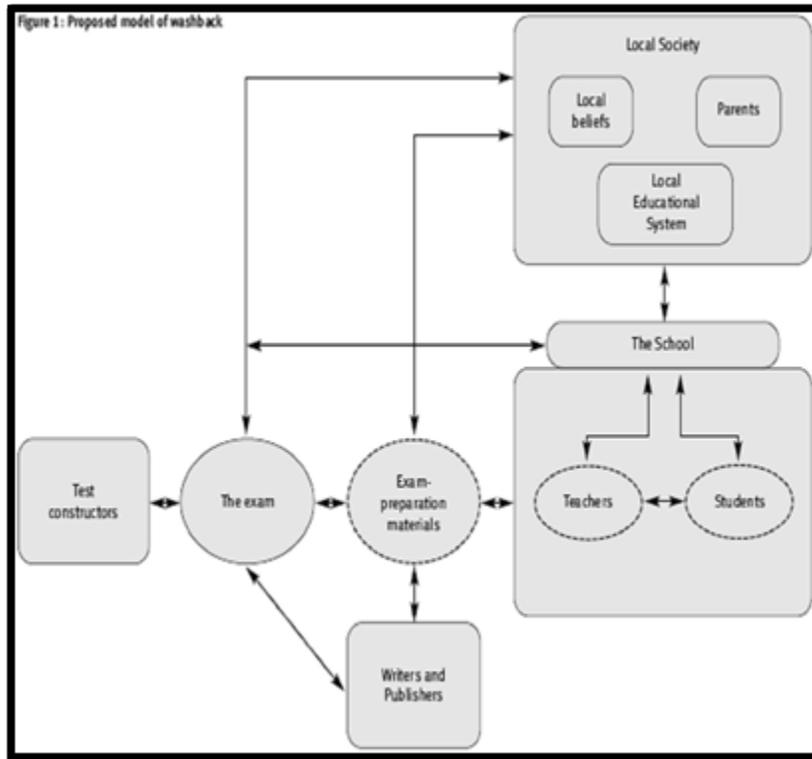


Figure 5 Tsagari’s (2009) Washback Model

The model indicates the impact of the participants’ values, beliefs, and test needs, as well as the test preparation materials. In this model, Tsagari’(2009) indicated that teachers' role is to facilitate and transform the exam preparation materials to their students, and the school’s role in the washback process is to adopt the values and needs of the local or national society.

Previous studies on Washback effect

Overview. The washback effect of tests and assessments has been investigated for different purposes. Some washback studies looked into the phenomenon at macro levels to see how large-scale tests affected the educational curriculum and the effect of the test on the

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

materials. Other studies, most of which are on large-scale tests, shed light on the perceptions and attitudes of learners and teachers towards the test and how these perceptions and attitudes affected learning and teaching. As the purpose of this study is to examine the washback effect on learners and learning, this part of the chapter focuses on reporting several studies that reported only the results of washback effects on learners and learning or studies that reported the washback effect on learning and teaching together.

Research on the washback effect of language testing started to take place in the early 1990s, with Alderson and Wall's 1993 literature review entitled "Does washback effect exist." That view laid out a number of "possible" washback hypotheses aimed to simplify the areas of teaching and learning to be affected by the test. Later, Alderson and Wall (1993) conducted a study entitled "Examining washback: The Sri Lankan Impact study," which examined whether washback of the test existed or not. The study sought to reveal the washback effect of a large-scale national second language test on teaching in Sri Lanka. The study concluded that the test had washback effect on the content of teaching as well as on the way tests were designed by teachers and local educational boards. The study suggested that there is a need to investigate the nature of washback effect of the test, not just looking for whether washback effect exists or not. The study has ever since been considered foundational for all washback research agenda (Green, 2013, p. 42) and is frequently cited in many washback studies.

Alderson and Hamps-Lyons (1996) compared TOEFL preparation courses with other ESL courses in order to look at different types of washback effects on teaching. This study was a revisit to Alderson and Wall's (1993) 15 hypotheses. Their conclusion resulted in an expansion on the 15 hypotheses and argued that the washback effect might vary based on how well the test is designed to be aligned with the current practices, how much the test maker considers the

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

suitable strategies for the test preparations, and how much a teacher and materials designer is capable of being creative.

Washback effect on learning. Hoa (2020) conducted a washback study on the English proficiency test at National University in Vietnam. Hoa examined how the test may influence classroom activities and the content of learning. The test is what is called "the Vietnam six-levels of foreign language proficiency framework (VNFLPF)," on which students must obtain a level three score upon graduation. The study aimed to explore whether the test had a washback effect on the teaching process and whether the changes in the process of teaching lead to changes in the learning styles. The participants were nine teachers and 679 non-English major students. Data collected from classroom observations, questionnaires and interviews demonstrated that the test resulted in both positive and negative washback effects on teaching and learning. The test affected how teachers designed their teaching methodologies in a way to make it correspond to the test; teachers expressed that the test influenced the content and the materials of teaching. On the other hand, the students reported that the content of their learning did not focus on the improvement of the four skills, whereas the teachers expressed otherwise.

Saglam (2018) examined the washback effect of a local integrated theme-based high-stakes test, an English language proficiency test, on what and how Turkish learners learn. The "integrated theme-based test" was assumed to cause a positive washback effect since it was presumably designed to imitate authentic assessments. Data analyzed from pre-post questionnaires and interviews revealed that the test had both positive and negative effects. The positive effects were observed on what and how learners learn. Learners expressed that using notes and sources-based information was helpful to enhance the listening skill. However, the

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

negative effect of the test was that learners valued the learning activities only because they were test oriented.

Damankesh & Babai (2015) looked into the effect of a high school final exam on students' test-taking and preparation strategies. In order to reveal what test-taking strategies students employed, 80 Iranian participants were videotaped as they were involved in a think-aloud protocol during the exam. The results revealed that learning behaviors were affected by the exam; students were noticed to use some types of test-taking strategies: successful guessing, lexical cohesion, reasoning, world knowledge technique, consider orthographical clues information, gathering, visual aids, and semantic and grammatical clues.

Similarly, Allen (2016 a) investigated the washback effect of academic IELTS preparation courses on the learners' score gain and test preparation strategies. Two IELTS tests were administered in 11 months to compare the score gains. There were 190 undergraduate Japanese participants who took a survey, and then 19 of them were interviewed. The study showed that the score gains occurred on speaking and listening skills. The results also revealed a positive attitude from participants towards the test preparation strategies as participants focused more on the productive skills than on the receptive skills. This may be due to the fact that the design of the test could have been deliberately oriented to target some specific skills; yet, this could also mean a poor design as the test did not cover all language skills.

The poor design of a test or assessment can lead to a limited washback effect. This means only some areas of learning are affected, but not other areas. Shohamy et al. (1996) reported that the poor design of an EFL test in Israel's educational system only contributed positively to the oral skills; they stated that the reading section of the exam did not enhance the reading skills in the classroom. On the other hand, a good design of a test may target a broader range of language

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

skills. Hughes (1988) reported that the student's language skills were enhanced as a result of the new Standardized university exam in Turkey. The participants' language proficiency advancement was evaluated by the Michigan test, which made the evaluation even more valid. Hughes (1988) argued that the university exam contributed positively to the learners' progress because the exam was aligned with the student's linguistic needs, not only what the students needed to study to pass.

Allen (2016 b) investigated the washback effect of a medium language scale University entrance exam at Tokyo University in Japan. The study targeted examining the learning behaviors and learning experiences of 133 undergraduate learners. The researcher conducted a survey and interviews to investigate learners' aspects of engagement (what skills and activities etc.), students' motivation towards studying some specific skills, and their perceptions towards measured language development. Results from the survey showed that learners were more engaged in the reading skill, grammatical structures, and test-taking techniques, whereas the speaking skill was left out from their focus. Allen concluded that the learning behavior was clearly geared by the design of the test, which centrally focused on multiple-choice grammar and vocabulary exercises. The motivation of learners was also oriented by the test design as the learner's goal was to succeed on the test. Likewise, the learners' perceptions towards their language skills development were aligned with the learner's focus on the "needed" skills to prepare for the test.

Attitudes and perceptions towards the test's washback effect on learning. Dong (2020) did a washback study on a national high-stake English test in china (NMET) to examine the effect of the test on learning. Dong looked into the connections between the perceptions, learning practices, and learning outcomes. Dong (2020) studied learners' perceptions towards the

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

test's validity and how those perceptions influenced the learner's learning practices and, consequently, how those learning practices influence learning outcomes. There were three thousand and one hundred and five (3105) Chinese high school seniors involved in the study. The findings from surveying the participants revealed that the learners' perceptions of NMET's validity influenced the learners' learning practices in a way that learners were numerous engaged in the test preparatory learning practices. Dong concluded that the more learners are involved in the learning practices, the better learning outcomes are.

Abbasian et al. (2017) conducted a washback study to examine the effect of a General English Language test; the test was a university entrance requirement for Ph.D. applicants of humanities and science school. Participants were grouped into a science group and a humanities group. The study examined how the test affected the student's perceptions and practices. Then, the study sought to examine the student's perspective regarding the alignments between students' second language needs and language test needs to pass the test. Data collected from surveying 560 Ph.D. students and then interviewing 16 of them demonstrated no significant difference in washback effect on the two groups'-science and humanities- perceptions and learning practices. Students' attitudes towards their perceived second language need and the language test needs signaled a positive washback effect of the test on the reading skill; this was because the test reading content included relevant readings to their majors. However, students expressed a negative washback effect as the test was not aligned with their motivation to improve their listening, speaking, and writing skills; the test did not target those skills. Regarding the washback effect on practices, science and humanities students showed some differences at the test preferred learning activities; for instance, humanities students valued self-learning more than science group.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Baksh et al. (2016) did a washback effect study to examine the effects of a newly proposed school-based assessment (SBA) in Malaysia. The study's goal was to measure the washback effect of SBA on the student's perceptions towards learning English as a second language. Moreover, the study sought to reveal their perceptions on the idea of implementing an "external examination (school-based assessment)." Participants were 34 students (17 females, 17 males) at the lower secondary school level. The researcher surveyed the participants and utilized official document analysis (e.g., booklets, press releases) to draw a picture of the intended washback effect. It was concluded that students had ambivalent attitudes; they equally expressed their uncertainty of which assessment (school-based assessment versus any other external examination) was more helpful for learning. Moreover, students stated that there were some barriers to effectively taking the school-based assessment (e.g., lack of adequate class time for practicing the assessment tasks, lack of assessment materials, teacher's knowledge).

Kennedy & Lui (2013) aimed to uncover the washback effect of the role of English classes in the final secondary school at preparing for the Beijing Matriculation English Test (BMET) through learners' and teachers' perceptions. Seventy secondary school seniors in China took questionnaire items, and 3 of the students and their teachers were later interviewed. Findings showed that the teachers and students thought that the only role of the final year English preparatory classes was to prepare students for the BMET, whereas enhancing the language skills was not as important. Kennedy and Lui stated that BMET may have "strongly" impacted the teaching activities and that those MBET oriented activities may have formed those negative perceptions expressed by learners.

Lewthwaite (2007) aimed to explore the washback effect of IELTS on writing tasks and the preparation activities for those IELTS writing tasks. Seventeen teachers and 36 students from

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

different majors at the college level in the United Arab Emirates took a survey that aimed at eliciting responses about the perceived usefulness of the IELTS writing tasks and the preparatory stage for the IELTS. The findings demonstrated that the writing classes were perceived as helpful since the exam writing task pertained to writing skills needed in that educational context. Only students from business and law majors expressed a negative perception towards the test writing tasks. Nevertheless, in general, participants stated that there was a substantial overlap between, on the one side, what carries out the assessment tasks, and the course IELTS writing tasks required, and, on the other side, what students and staff thought was needed in a writing course.

Washback effect of Alternative assessments and criterion tests. Investigating the washback effect of alternative and criterion tests is less popular. Few efforts have been made to examine the washback effect of alternative assessment (Hung 2012). A more significant number of washback studies has been dedicated to large-scale language tests, standardized tests, University entrance exams, etc. The reason behind that lies in the conviction that large-scale tests have a strong effect on teachers' and learners' behaviors in the classroom and a strong effect at the macro level, in school policies, and in local and national educational systems (Wall, 1997). Yet, there has now been a direction towards considering the washback effect of alternative assessments and criterion tests as they have a washback effect on learners at the micro-level (Hung 2012). One of the most critical effects of alternative assessments is that they primarily affect the learning progress (Eckstein and Noah 1993). This section reports several washback studies that attempted to prove the washback effect of criterion tests and alternative assessments. Shirzad & Amerian (2020) investigated the washback effect of a grammatical-focused writing task on writing performance progress. There were 120 ESL upper-intermediate learners (females

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

and males) in Iran who participated in the study. The researchers created three grammatical questions (cloze, multiple-choice and metalinguistic) to test learners on grammatical points such as present perfect and perfect continuous. Those questions were given after the participants had had practicing sessions on those grammatical points. Participants were categorized into three experimental groups (multiple choice, cloze, and metalinguistic) versus one control group. Afterward, the four groups were required to compose two grammatical-focused writing tasks. Post-tests' findings showed that the grammatical tests caused a positive washback effect on the experimental groups' writing performance. Participants showed higher levels of accuracy and syntactic complexity. The metalinguistics group achieved the highest, followed by the cloze and then the multiple-choice group.

Burksaitiene and Tereseviciene (2008) created a portfolio assessment to study the washback effect of the assessment on learning writing. Students were granted the freedom to form their learning groups as well as the topics they want to discuss and write about. Learners were also involved with the teacher to build the assessment criteria. Students were given a questionnaire to fill out to reveal the students' attitudes and satisfaction towards the writing portfolio assessment. They were asked to express their opinions towards the assessment. The study concluded that the integration of alternative assessment was appealing to the students and helped them be more motivated towards learning second language writing.

Safa (2014) studied the washback effect on task-based language assessment (TBLA) in the context of learning English as a foreign language. The study looked into the washback effect of TBLA on learning English grammar compared to the traditional assessment. The participants were 74 pre-intermediate EFL Iranian learners who were placed in two groups, a control, an experimental group. They were taught grammatical points the same way for 90 minutes, twice a

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

week, for ten sessions. After every three sessions, the two groups were given a quiz: the control group took a traditional grammar quiz (multiple-choice, fill in the blanks, true-false, and matching), whereas the experimental group took a task-based quiz. After that, the two groups received the proper treatment, and then they took a post-test. The post-test findings demonstrated that the TBLA group enhanced their English grammar knowledge significantly more than the traditional language assessment group.

Hung (2012) examined the washback effect of an e-portfolio project on content learning. The study was an e-portfolio assessment project during a language teacher preparation content course (a graduate-level content course). The study was conducted on 18 EFL graduate students who were also EFL teachers in Taiwan. The content course introduced the topic of language assessment that is intended to enhance EFL teachers' language assessment skills through language assessment theories and practices. The course lasted 18 weeks, three-hour meetings a week. Data from interviews, observations, reflective journals, and document analysis demonstrated a positive washback effect of the assessment on content learning. Participants formed a community of practices, enhanced their critical thinking skills, and created a peer learning and professional development environment. However, the findings also detected a negative washback effect: the participants said that they felt a little anxious as the other students saw their written works. Moreover, another negative washback was noticed on the participants' reluctance to use technology because of some technological inconvenience that occurred during the course.

In conclusion, the washback effect is a complex phenomenon (Alderson & Wall 1993), and so Alderson and Wall (1993) proposed 15 hypotheses to simplify that washback complexity. They attempted to categorize the areas of learning and teaching that can affect the test. Then,

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

washback models came to propose ways to detect the mechanisms of the washback effect.

Moreover, those models stressed the fact that the washback effect has a contextual nature.

Washback studies of different goals were dedicated to studying various effects.

Similarly, the results from washback studies were contextually interpreted due to many involving factors, such as the stakeholders' attention, social and educational values, etc. Most washback studies have dealt with standardized and large-scale tests. Yet, criterion tests and alternative assessments have recently been However, it has been found that tests can have both positive and negative washback effects. Many studies have shown that learners and teachers can adjust their behaviors to align with the demands of the test (Green 2013, p.44). Studies also demonstrated that the demands of negativity of the washback effect of the tests and assessments depend heavily on the design of the test and the tests' and assessments' alignment with test's design needs.

Self and Peer assessment in L2 Writing.

Self-assessment. There is no one standard definition for self-assessment, but there are common self-assessment features for multiple definitions (Andrade & Du 2007). For instance, one of the most referenced definitions is by Falchion & Boud (1989), who defined self-assessment as the learners' involvement to "make judgments about aspects of their own performance." Another widely used definition is that self-assessment is a process of a formative assessment during which learners are allowed to make their evaluations on their learning and make judgments on their learning according to a stated criterion to identify their weaknesses and strengths (Goodrich, 1996; Gregory et al., 2000; Hanrahan & Isaacs, 2001; Paris & Paris, 2001; Andrade & Boulay, 2003 as cited in Andrade & Du 2007).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Self-assessment in learning is an effective learning tool because it contributes to self-learning (Roos 2006). Learners get the chance to practice and enhance their learning autonomy, regulate their learning skills (Honsa 2013), and improve their self-awareness and critical thinking (Ashraf, & Mahdinezhad, 2015, p. 111). This all applies to the skill of writing as learners, through self-assessment, are pushed to read their writings critically (Singh and Terry 2008), and eventually, they improve their writing skills (Bing, 2016, p.99). Thus, self-assessment should consistently be considered in second language writing to offer learners opportunities to analyze their writings and explicitly think about and observe their learning progress. Consequently, students develop mental awareness (Schendel and O'Neill, 1999), enabling them to figure out their weaknesses and strengths (Ferris & Hedgcock 2013, p. 262).

Peer assessment. Peer assessment, in the educational setting, is defined as an arrangement in which learners consider and specify "the amount, level, value, quality or success of the products or outcomes of learning from peers of similar status" (Topping,1998, p. 250). Peer assessment has been consistently reported to help both learners who assess and for those who are peer assessed (Berg 1999; Paulus 1999; Patchan et al., 2011), resulting in improving performance in later assessments (Jhangiani 2016). It forces learners to take the teacher's role; subsequently, learners are prompted to use the language naturally. Peer assessment brings students of different areas of competence together, and consequently, students benefit from each other (Min, 2006). During the assessment process, the peer's criticism, especially in writing, is expected to be accepted with less tension than the teachers' criticism (Black et al. 2004). Similarly, learners are believed to interact and negotiate more with one another than with the teacher. One teacher quoted a statement as a result of an observation of peer assessment: "When students do not understand an explanation, they are likely to interrupt a fellow student when they

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

would not interrupt a teacher" (Black et al. 2004). While some claim that peer assessment is a time-consuming activity," as the teacher provides the learners with feedback, the benefits of peer assessment have been strongly depicted in the body of literature (Lee 2017). The effects of peer assessment on learners' writing performances were proved to be positive in new studies and many well-known early studies (e.g., Hughes & Large, 1993; Freeman, 1995; Dyer, 1996; Brown, 2001).

Attitudes and perceptions towards Self and Peer Assessment. Attitudes and perceptions towards peer assessment in second language writing are not consistent. Some concluded that learners had positive attitudes towards peer assessment (Gatfield 1999; Roskams 1999; Liu and Yuan 2003; Collimore, Paré, and Joordens 2015; Schunn, Godley, and DeMartino 2016). Other studies, on the other hand, reported that learners expressed different degrees of negative attitudes towards the implementation of peer assessment (Roskams 1999; Lin, Liu, and Yuan 2002; Wen and Tsai 2006; Kaufman and Schunn 2011; Praver, Rouault, and Eidswick 2011; McGarr and Clifford 2013). Other studies reported the causes why some negative attitudes occurred: a result of the learner's perceived reliability of the peer's rating (Liu and Carless 2006), perceived expertise of peers (Cheng and Warren 2005; Liu and Carless 2006; Kaufman and Schunn 2011), power relations between the assessed and assessor (Brown and Knight 1994; Liu and Carless 2006), time dedicated for the peer assessment (Liu and Carless 2006), and the lack of accuracy of peer evaluation generated from peer assessment (Wen and Tsai 2006; Kaufman and Schunn 2011), students' low L2 proficiency level, and learners' concerns about their interpersonal relations (Wang 2014).

Less attention has been paid to study the attitudes towards self-assessment in L2 writing. A few examples that can be reported in this regard include Vasu et al. (2018). They investigated

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

how valuable and effective self-assessment checklists are (adopted from Nimehchisalem et al.2014) in undergraduate ESL writing class. The assessment was implemented in five weeks. Data collected from interviews showed that the self-assessment checklist from motivated learners to write helped them be more aware of their writing skills improved their comprehension of the writing task, promoted self-learning autonomy, and scaffolded them to write systematically. Another example is by Oscarson, A. D. (2009), who sought to reveal how self-assessment can contribute to EFL writing and skills at an upper secondary level in Sweden. The assessment form was used within the school writing tasks (timed writing task). The assessment was a self-questionnaire in which learners had to grade several parts of their writings (5 Likert scales) according to the writing task requirements. Data from interviewing both teachers and learners revealed that both teachers and learners positively perceived self-assessment as a transferable learning skill that can be effectively applied to other areas of learning.

Self and Peer Assessment together in Second Language writing. Self and peer assessment are complementary (Black et al., 2004). Peer assessment practices enhance the skills of self-regulated learning and autonomous learning (Liu and Hansen 2002). Thus, peer assessment might be incorporated with self-assessment to enlighten the students on how to do self-evaluation. Another way is to implement self-assessment before peer assessment where learners reflect on their writing, diagnose the gaps and strengths in their writings, then take those gaps to peers to focus on them in the feedback session (Lee 2017, p.97). One of the rare studies that considered implementing self and peer assessment together in L2 writing was by Iraj (2016). It was concluded that EFL Iranian learners improved their argumentative writing skills through self and peer assessment (p.720).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Summary

The first section of this chapter presented clarifications on the concept of the washback effect of testing and provided definitions of the washback effect phenomenon; it also determined the terms to be used for the current study. Next, the section viewed and discussed the big issue of washback mechanisms and how scholars, variously, attempted to simplify it. The section also viewed the most popular washback theoretical underpinning proposed by Alderson and Wall (1993), which classifies the learning areas affected by tests. The section also reported several washback studies of tests and assessments in second and foreign language contexts. The section focused only on the washback effect of tests and assessments on learning.

The second section of this chapter provided definitions and concepts of self and peer assessment in second and foreign language writing; it also shed light, briefly, on the findings on the participant's attitudes and perceptions towards the use of self and peer assessments in second and foreign language writing. Then, the section concluded with the significance of utilizing self and peer assessment together in second language writing. Finally, the section viewed the importance of self and peer assessment in second language writing backed up by several theoretical underpinnings.

Conceptualization of the Current study

As clearly demonstrated in previous sections, the Washback effect is a complex phenomenon as the intentions, and the washback effect results are always interpreted contextually (Green 2013). The models and the studies presented in this paper stress the fact that Washback is heavily based on the social contextual factor and the participant's perceptions and attitudes. Therefore, there is a need to consider looking into the perceptions and attitudes of the

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

participants. Participants' perceptions and attitudes play an essential role in determining the type of washback that might be resulted (Dong 2020, p.10).

Anderson and Wall (1993) suggested 15 systematic hypotheses that differentiated between effects on attitudes and effects on the content of learning and the between the effects on methods and effects on the process of learning (Green, 2013, p. 42). Those variables can help make predictions about: content (what), methods (how), rate, sequence, and the depth of learning (Green 2013, p.47). Alderson and Hamp-Lyons (1996) later on suggested that there is also a need to consider several factors: the stake of the test, how much the test is in harmony with the current practices.

Most of the studies on the washback effect have been dedicated to large-scale national and tests (Standardized, entrance exams, etc.) Even though fewer studies have targeted criterion tests and alternative assessments (Green, 2013, p.43), those studies proved that alternative assessments (e.g., self and peer assessment) could have a washback effect on learning. Since peer and self-assessment, similar to any other alternative assessment, provide practical evaluation tools of the learner's abilities to accomplish language tasks (Cheng, and Curtis as cited in Huung 2012), it is very informative to study how the assessment affects the learning. In second language writing, peer and self-assessments are used to help learners improve their writing skills. However, detecting what learners think of the assessments can help detect gaps and improve the assessment.

Given that, the study examines how self and peer assessment affect learning second language writing. The study will take into account the attitude changes of participants towards the assessments.

Chapter 3: Methodology

Overview

The methodology chapter starts by describing the purpose and the significance of the study to pave the way for re-introducing the research question afterward. Then, the chapter goes on to describe the research design and materials. Also, the setting, participants, and the procedure of the data collection are explained. Finally, the chapter demonstrates the theoretical base that underpinned the study and the data analysis instruments.

Purpose and Significance

The study aims to investigate the washback effects of self and peer assessment in second language writing. Determining those washback effects will provide a comprehensive overview of how the assessments operate on learners so that the pros and cons of the assessments are spotted and dealt with for future assessment implementations. When conducting a washback study, it is indispensable to consider the participants' perceptions and attitudes about the assessment (Dong 2020). Thus, as the washback on learning is the target in this study, learners are asked about their opinions about their experiences taking the self and peer assessments. Considering the learners' opinions when examining the assessment effects would make learners feel heard, considerate, and motivated.

The idea of implementing self-assessment in writing is not as popular as peer assessment. More than that, the idea of considering both self and peer assessment in the same process is even less popular. The study combines self and peer assessment in order to increase the intensity of critique in writing. The study was designed to have learners complete self-assessment first and

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

then peer assessment. To make the assessment more purposeful and personal, learners may take self-assessment before peer assessment; then they have the chance to reflect on their writing, diagnose the gaps and strengths in their writings, and finally take those gaps to peer assessment to focus on them (Lee 2017, p.97). Adding the element investigating the washback effect, as explained in the above paragraph, would help measure the effectiveness of self and peer assessment. Given the purpose and the significance of using both and self and peer assessment in second language writing, the study seeks to answer the following three questions:

Research Questions

1. What are the participants' pre-and post-attitudes towards self and peer assessment?
2. The importance of self and peer assessment in revising second language writing:

A: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing?

B: Are there any washback effects of self and peer assessment on the participants to revise their writing? If so, what are they?

1. The effectiveness of considering and not considering the peer's self-assessment results in peer assessment:

A: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

B: Are there any washback effects of considering and not considering peer's self-assessment results in peer assessment? If so, what are they?

Research design

Overview. Washback studies have been investigated through the mixed-method approach (Green 2013, p.42). This study employed a mixed-method for data triangulation. The quantitative research tools were pre-and post-questionnaires administered to uncover the attitudes and perceptions of learners. The post questionnaire also provided qualitative data through the learners' responses to several open-ended questions. Additionally, the qualitative research was provided through interviews to gain more insightful data about their attitudes and perceptions and reveal the washback effects on learning.

Qualitative and quantitative research. The nature of washback effect is complex and contextual, and so the need to adopt a qualitative approach is essential to draw a comprehensive picture of the learners' attitudes and perceptions about the assessments. Sowden (1992) suggested that qualitative research reveals and breaks down complexities associated with complex phenomena such as washback. Moreover, the need for a quantitative approach is also vital to draw unbiased conclusions, restrict the alternative interpretations, and simplify complex phenomenon explanations (Creswell, 2014).

Mixed method. Integrating the qualitative and quantitative data leads to a "more complete understanding of research problems than either approach alone" (Creswell (2014, p. 4). The mixed-method also helps enhance the washback validity (Tasgari 2006). Therefore, for this study, I wanted to get rich data and have it rigorously validated as much as possible. The study's nature entails examining the attitudes and perceptions towards the assessments quantitatively and then looking for more in-depth relevance. However, revealing the washback effects of the assessment on participants is best searched through a magnifying qualitative scope because there

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

is always a need to hear from participants about their experiences and behaviors to interpret the phenomenon accurately.

Data collection

Self and peer assessment forms. The self and peer assessment forms used for the study were analytically designed to be aligned with the teacher's rubric. The reason for that was to make the assessment forms accurately target the various writing areas learners had been taught to fulfill in the writing task. The design of the peer assessment forms was partially adapted from Lee (2107). The form included the ten major writing components of the story writing task that existed in the teacher's rubric; each writing component had scaled prompting descriptors (a scale from 1–4). Thus, learners were prompted to peer-assess and rate the writing performance within a performance scale of 1-4 with the scaled descriptors' help on each writing component. Learners had to give a score on each writing component the total score on the ten components.

Additionally, the peer form contained a summative assessment section where a student can summarize the peer's writing performance.

The self-assessment form was more a checklist design adapted from Nimehchisalem (2014); the form contains the ten writing components about the story writing tasks (as in the peer assessment form) but with a single prompting descriptor for each writing component that had to be rated on a scale of 1-4. Finally, the students would give a total score on the ten writing components. Also, the self-assessment form had a summative self-assessment section adapted from Hung (2012) in which the self-assessor was asked to provide comments and examples about self-writing weaknesses and strengths.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Pre & post questionnaires. The questionnaire's goal was to reveal the learners' attitudes and perceptions towards self and peer assessment in writing. The questionnaire was administered in English (the learners' first language), and it was administered twice-- before the assessments and then after the assessments. The reason behind administering two questionnaires was to examine possible changes in perceptions and attitudes. Learners' perceptions and attitudes towards self and peer-assessment can be good indicators and contributors to explain the washback effects generated from the assessments (Bachman and Palmer (1996). The post-questionnaire, moreover, contained several open-ended questions for some qualitative data. The pre-questionnaire (see the Appendix) consisted of two main sections. The first section included demographic question items about gender, speaking, and writing in Arabic outside of the class and the prior experience with self and peer assessment in writing. The second section aimed to address attitudes and perceptions, but the section was divided into subsections to target various attitudes and perceptions stimulated by the research questions. Those subsections started with nine items addressing attitudes: 5 items about the feelings and four items on the perceptions of learners' capability to self and peer assess. These nine items addressed the study's first question: What are the participants' pre- and post-attitudes towards self and peer assessment? The second subsection contained six items to address the perceived importance of self and peer assessments; these six items were to address the first part of the second question of the study: What washback effects do self and peer assessment have on the participants to revise their writing? Finally, the third subsection had six items to reveal the learners' perceptions about the effectiveness of considering and not considering self-assessment in peer assessment; these six items were to target the first part of the third question of the study: What are the participants' pre- and post-perceptions about considering the peer's self-assessment results in peer assessment?

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The post questionnaire was identical to the pre-questionnaire; yet, it included several open-ended questions. Those questions were adapted from Planas et al. (2014) and McGarr et al. (2013) and aimed to elicit more in-depth explanations about the preconceptions, attitudes, and experiences about taking self-and peer-assessment. Also, there were some open-ended questions to reveal the potential washback effects of the assessments on learning.

A five Likert scale was used to measure the responses to the items. Measuring attitudes and perceptions is famously and effectively done through the Likert scale (Cohen et al., 2007; Sullivan et al. (, 2013). The Likert scaled items were predominant and were based on a 5-points scale: 1= strongly disagree, 2= disagree, 3= neutral, 4= agree, 5= strongly agree.

Semi-structured interviews. Semi-structured interviews are beneficial in mixed-method research to be used as "an adjunct to supplement" to survey questionnaire items and also to explore relevance that can emerge (Newcomer et al., 2015). Therefore, one aim behind the interviews was to collect more detailed qualitative data about the learners' perceptions and attitudes drawn from their experience in taking self and peer assessments. Moreover, the interviews attempted to reveal the washback effects on learning generated from peer and self-assessment; those washback effects are detailed in the second part of questions 2 and 3 of the study: 2. B: What washback back effects do self, and peer assessment have on the participants to revise their writing? 3. B: What are the washback effects of considering and not considering the peer's self-assessment results in peer assessment?

Interviews were conducted individually with ten learners via Zoom application, and each interview took around 10 –12 minutes. The interviews were conducted virtually because of the spread of the COVID-19 pandemic at the time of data collection in November 2020. The benefit that I found from virtual interviews was the free-tension environment. I noticed relaxation signs

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

from the interviewees, unlike other interviews I had conducted previously in other studies. I chose Zoom because of its simplicity and user-friendliness (Archibald et al., 2019).

Identifying and selecting the right participants for the semi-structured interview was based on how informative those participants were for the research (Raworth et al., 2012). I selected the ten participants for the semi-structured interviews based on their contributions to the self and peer assessments; I looked through the assessment forms, and from there, I selected them. At the beginning of each interview, participants were told to express their opinions freely and that nobody would be judged, and that their names would be anonymous. At the beginning of each interview, this statement aimed to make the participants feel free and relaxed to express ideas. Questions were predetermined to know as much as possible about the learners' perceptions, attitudes, and experiences about taking self and peer assessing writing. I also implemented "thought-provoking interjections" to indicate to the interviewees that I wanted to hear more about their expressions (Pathak, 2016).

The writing task. The writing task I integrated into this study was genuinely a part of the course curriculum. It was a story writing task that the teachers had previously assigned as a final writing project at the end of the semester. The teacher had already decided to have learners go through a writing process to develop the topic; she also shared with learners an analytical rubric (see the appendix) for the story that she would eventually use for grading the learners' writing performance. Learners were permitted to choose a story topic; yet, they were given a prompt (see table 3.1) for paper requirements and structure. The writing process took place on google docs; the teacher shared separate docs with each learner along with the rubric. The story prompt and the rubric were written in the learner's first language (English) and the target language (Arabic). The teacher shared both versions during class time, but the teacher only shared the Arabic

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

version on google docs. According to the teacher, the reason for that was to let the learners grasp the meanings of terms in the rubric during the class so that they would be ready to deal with those terms in the target language during the process of writing. The story prompt and teacher's rubric to be shared in this section are only the English versions:

Story writing g prompt used by the teacher for the participants

Write a story to share with the class later.

- You may use an actual event that happened to you or a friend, or you can simply make up a good story. You will have 6 days to complete the first draft and then 5 days for the final draft, so I expect it to be a good one. The story must be typed. I will not accept any computer problems as excuses, and so plan. In the end, the class will vote on the best story.

Requirements:

- **500- 700 words**
- Write the word count at the bottom of the page
- Title the story
- Use descriptive words.
- The setting must be fully developed.
- The plot must be fully developed.
- The story must be suspenseful.
- The ending should be suspenseful or surprising.
- Follow the plotline (background, conflict, rising action, climax, falling action, resolution)

Setting. The city of Memphis, located in the state of Tennessee that sits in a region in the United States known as Mid-South, is a home of a relatively large Arab and Muslim community. As I am a member of this community, I can confidently say that Arabic is frequently spoken and heard in all public places in and around the city; you can even notice Arabian restaurants and shops with Arabic signs and writings. What is more relevant to this study is the existence of a big Islamic school where teaching Arabic is considered a backbone of the school mission. The reason for that is the holy status that Arabic has in the Muslim faith since, according to religious tradition, all primary Islamic practices should be delivered through the sacred language (Arabic).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Therefore, all Muslims (Arab & non-Arab) need to use Arabic in their faith rituals. The cultural and social variables are also involved. Arabic in the eyes of Arabic-speaking people is an identity; thus, those Arabs who live in non-Arabic speaking countries (like in the USA) seek to teach their children the Arabic language. The Islamic school in the city of Memphis teaches all subjects in English, yet Arabic is introduced as a second language from KG to 12th grade. The Arabic program stresses teaching four skills communicatively. One teacher is in charge of a particular grade level. The Arabic class lasts fifty minutes five times a week. The skill of writing is usually delivered through the process of writing with some occasional timed short essays. With the school's relatively high enrollment rate, I was able to maintain the number of participants I needed to conduct this study.

Participants. The study involved twenty learners whose first language is English, and they study Arabic as a second language. Most of the participants were born in the United States, but few newly immigrated to the country. All participants are from either Arabic-speaking countries or Muslim majority countries such as Pakistan, Bangladesh, Ethiopia, and Turkey. The participants were high schoolers who were enrolled in advanced Arabic levels. The participants' usage and exposure to Arabic out of the school is dependent on one major factor--whether there is an Arabic-speaking parent in the household. Basically, if the family is Arab, learners speak Arabic in the household, but if the family is from a Muslim country where Arabic is not spoken, then the learners' use of Arabic is restricted.

The learners participated voluntarily in the study, and they were mostly females (6 males vs. 14 females). Their ages ranged from fourteen to sixteen years old. The learners came from two different grades (8th & ninth), but both grades were at the advanced Arabic level. Considering advanced Arabic learners was due to the larger amount of writing advanced learners at the school

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

had to do. This, consequently, led to having richer data for the study. Also, advanced learners can express themselves better in interviews; thus, advanced Arabic learners would provide more clearly expressed responses than lower-level second language learners would do. Above that, the study considered using self and peer assessment where learners needed to write comments and feedback, so the selection of advanced learners provided high-quality data in this regard.

Demographic information about the participants' experience with Arabic and their experience with self and peer assessment was collected through a questionnaire. As indicated in table (1), most participants were females, with 70% of the class number. Percentages also show that most participants speak and write (e.g., text messages) outside the classroom. Table (1) also indicates that most learners have experienced taking self and peer assessment in writing.

Table (1)

Demographic information of the learners, N=20

| Variables | | Sum | Percentage |
|--|--------|-----|------------|
| Gender | Female | 14 | 70% |
| | Male | 6 | 30% |
| Speaking Arabic outside the classroom | Yes | 12 | 60% |
| | No | 8 | 40% |
| Writing in Arabic outside of the class | Yes | 11 | 55% |
| | No | 9 | 45% |
| Past experience with self-assessment | Yes | 15 | 75% |
| | No | 5 | 25% |
| Past experience with peer assessment | Yes | 16 | 80% |
| | No | 4 | 20% |

Procedure. Once the participants had completed the first draft of the writing task (story writing) on google docs, it was time to administer the pre-questionnaire items. Due to the spread of the COVID-19 pandemic at the time of the data collection in November 2020, the class was

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

delivered virtually; thus, the researcher contacted the teacher to virtually visit the classroom on the zoom application in order to explain the purpose and the content of the questionnaire. The researcher sent the questionnaire items that had been digitally created through google survey format to the participants' school emails. In the meanwhile, the researcher was present to respond to the learners' questions about the items. All learners' responses to the questionnaire items were electronically stored in google survey format in the researcher's google account.

The next day, the researcher visited the class virtually and introduced the self-and peer-assessment forms to the learners. The teacher also intervened to elaborate on some points as well. Learners were asked to carefully read the content of the forms to conceptualize the criteria as learners necessarily need to understand what makes a piece of work good or bad (Brown, Rust, & Gibbs, 1994). The assessment forms were introduced in English (participants' first language) to avoid any misunderstanding of the criteria. The criterion was based on the teacher's English version of the rubric that she had introduced to learners at some point. The researcher asked for a digital copy of the learners' papers on google, and, therefore, the teacher shared the whole google folder with the researcher for the researcher to observe and edit. To start the self-assessment, the researcher shared the self-assessment forms with learners on google docs by posting the self-assessment form under each learner's first draft. Then the learners were asked to self-assess their writing using the self-assessment form. Learners were given 35 minutes to finish the self-assessment.

After finishing with the self-assessment, the researcher removed the names on the first draft and from the self-assessment and replaced the real names with numbers. Afterward, each learner's first draft was doubled, and the two versions were placed in two files: one draft for each learner along with the filled out self-assessment form on google docs, and the other version for

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

each learner would be without the self-assessment results. Now the researcher had prepared the action for the peer assessment. So, each student's first draft on google docs had two versions (one with self-assessment, another without), and each version would be sent to a different peer to peer- assess it. The next class time, the researcher asked the teacher permission to share with each participant the two classmate's papers they would peer assess. So, firstly, each learner received a classmate's story writing draft along with the self-assessment results on their google docs account, and they were asked to peer-assess it by considering their peer's self-assessment results had been written down in the self-assessment form. After all peer assessors completed peer assessing the first classmate's draft, they received another classmate's writing draft on their google docs account, but this time, the draft has no self-assessment form. Now the peer assessors were asked to peer-assess another classmate's draft without considering the classmate's self-assessment results. Learners were given 30 minutes on each peer assessment, with a total of 60 minutes on both sessions. Afterward, learners were given four days to consider the peer comments and feedback given on their two versions and turn in the final draft.

So, the process of the assessments started with a self-assessment and then two peer assessments. Self-assessment sessions preceded peer assessments so that learners were given a chance to diagnose the gaps and then share those gaps with one of the peer assessors- the one who had to consider the self-assessment results-to focus on them (Lee 2017, p.97), and so the assessment is now more purposeful and personal. Also, this way allowed me to examine how self-assessment contributes to peer assessment--one of the questions this study seeks to answer.

Also, during the peer assessments, as explained above, participants received anonymous written products; this was meant to increase learners' involvement, produce a sense of security,

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

and reduce the level of pressure (Vickerman 2009; van Gennip et al. 2010; Raeset et al. 2015). Consequently, learners would produce more critical evaluations (Lu et al., 2007).

The post questionnaire items were administered two days after the learners turned in the final draft. Like the pre-questionnaire, the post questionnaire was formulated in a google survey format and sent to the participants' school emails. The questionnaire was administered during class time with the presence of the researcher and the teacher to respond to the participants' questions. All answers were received and stored in the researcher's google survey account. Finally, interviews were conducted with ten learners on five days (2 participants a day). The researcher scheduled the meetings during class time. The researcher sent the selected participants a sheet with time slots with the meeting zoom link. The interviews took ten days, an average of 2 meetings a day. The ten interviews were recorded on zoom applications well as the researcher's cellphone. Finally, the interview data were transcribed and categorized for the data analysis.

Data analysis

Theoretical underpinning. The theoretical base for the study is Alderson and Wall's (1993) fifteen hypotheses. The hypotheses deal with the washback effects of large-scale standardized tests. This study investigated the washback effect of assessments, not tests, so the hypothesis guidelines were adopted to be used for the sake of the self and peer assessment. Also, the study considers only four principles of the fifteen that Alderson and Wall (1993) suggested:

2.A test will influence learning

6.A test will influence how learners learn

10.A test will influence the degree and depth of learning

12.Tests that have essential consequences will have washback

(Alderson and Wall 1993, pp.120-121)

Quantitative Data Analysis. Data collected from the closed-ended responses in the pre- and post-questionnaire were analyzed by the Statistical Package for Social (SPSS) version 26.0. The data from the questionnaires were input and analyzed using descriptive and inferential statistics. The analysis included each item's frequency distributions (Mean, Standardized median, Skewness, and Kurtosis).

Qualitative data analysis. Data collected from the open-ended responses in the post questionnaire and the interview transcripts were thematically categorized using NVivo application version 12.0. The thematic analysis method is one effective way to interpret the individual experiences and narratives (Pavlenko 2008). The analysis process started by reading the data multiple times to conceptualize the content, coding chunks of data, then merging the chunks into more extensive categories, and finally, all categories were labeled. The themes were premeditated to address the study questions, and, therefore, the codes and categories were revised and moved around multiple times to make the themes, as much as possible, rigorously arranged according to the study questions. Some of the data were used to support and explain the quantitative data, and some were used to answer parts of the questions that the quantitative data did not address. Moreover, emerging findings were also considered in coding and categorizing, and so those themes were separately themed.

CHAPTER 4: Results and Analysis

Overview

This chapter shows the findings resulted from administering pre- and post-questionnaire on twenty participants and from interviewing 10 participants (learners) to answer the three questions of this study:

1. *What are the participants' pre- and post-attitudes and perceptions towards self and peer assessment?*

2. *The importance of self and peer assessment in revising second language writing:*

A: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing?

B: Are there any washback effects of self and peer assessment on the participants to revise their writing? If so, what are they?

1. *The effectiveness of considering and not considering the peer's self-assessment results in peer assessment:*

A: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

B: What are the washback effects of considering and not considering peer's self-assessment results in peer assessment?

All research questions were aimed to be answered both quantitatively and qualitatively through questionnaire items and interviews. The questionnaire items were to show quantitative data about the participants' pre-and post- attitudes and perceptions towards self and peer assessments. The post questionnaire also included open-ended questions for more in-depth qualitative data. The

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

interviews were conducted to elaborate on the questionnaire's findings and reveal any possible washback effects of the assessments on learning and learners.

The washback effect on learners' attitudes

This section demonstrates findings to address the first question: *What are the participants' pre- and post-attitudes towards self and peer assessment?* Descriptive statistics are reported to show an overview of the data distributions of learners' attitudes and perceptions towards self and peer assessment in Arabic writing as a second language. Moreover, descriptive statistics are also used to show the attitude change- if there is any- after taking the assessments. Qualitative data from open-ended questions in the post questionnaire and interviews show more insightful data about the attitude changes towards self and peer assessments.

This part includes two subsections: feelings about self and peer assessment and the perceived capability to self and peer assess writing as a second language to reveal the learners' attitudes towards self and peer assessment.

Learners' feelings about the assessments before taking the assessments. Table 1 presents the descriptive statistics of the learners'-feelings about self and peer- assessment before taking the assessments. The examination of variable distributions showed that all variables had skewness and kurtosis values within or close to the satisfactory range (+/- 1). Even though some items were beyond +/- 1, the skewness and kurtosis values were both considered within the guideline frame by Kline (2011): no skewness values are beyond the absolute value of three, and no kurtosis is beyond the absolute value of ten. That is, the observed variables of the learners' feelings were seen within the normal distribution range.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

This section in the questionnaire- feelings- aimed to capture the learners' levels of excitement, confidence, embracement, and perceived fairness of the assessments. As can be seen in table 2, most learners expressed their excitement towards the assessments: variable 1 (M=3.30); and SD value shows relatively clustered data around the mean (SD= 0.87). As for the learner's confidence, statistics of variables 2&3 show that learners were more confident to take peer assessment (M=3.35) as opposed to self-assessment (M=2.90). The mode value of (4) on peer assessment variable –3- further strengthens the learners' sense of confidence to handle peer assessment.

Table 2.

Learner's feelings about the assessments before taking the assessments

| Variables | Mode | Median | Mean | Std.Dev | Skewness | Kurtosis |
|---|------|--------|------|---------|----------|----------|
| 1. I am excited about self and peer assessment | 3.00 | 3.00 | 3.30 | 0.87 | 1.231 | 1.573 |
| 2. I feel confident to self-assess my story | 2.00 | 3.00 | 2.90 | 1.021 | .218 | -.586 |
| 3. I feel confident to peer - assess a classmate's essay | 4.00 | 3.00 | 3.35 | 1.182 | -.769 | -.298 |
| 4. I think ratings resulting from peer assessment are not fair | 4.00 | 3.00 | 3.45 | .999 | -.024 | -.992 |
| 5. Having my peer assessing my story will embarrass me. | 2.00 | 3.00 | 3.20 | 1.24 | .501 | -1.401 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

As for the learners' perceptions on how unfair a peer's rating can be, the mean score of (M=3.45), and a strong SD (.999), as well as a mode score of (4) on variable 4, demonstrate that the majority thought that the peer's rating is not fair. Likewise, variable five shows that the feeling embarrassed to have your writing peer-assessed was expressed by the majority (M=3.20).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Pre-perceptions of learners' capability to self and peer assess. Table 3 indicates the descriptive statistics of the learners' pre-perceived capability to self and peer-assess. The skewness and kurtosis's statistical values show that the distribution of values is within the acceptable range +/-1; thus, the data is not heavily skewed, nor is it sharply peaked. That is, the distribution of values is within the acceptable range of normal distribution.

Table 3 indicates that more than half of learners expressed their doubts on variable six about being able to critically peer assess as the mean value was (M=3.05): *I do not feel I will not be critical in assessing a classmate's writing*, with a relatively clustered SD value of (SD=.999). However, in terms of not being able to rate accurately in peer assessment- variable 7- the mean score of (M=2.65), and low SD value of (SD=.875), and a mode of a score of (2) showed that the majority were confident in this regard.

Table 3

learners' capability to self and peer assess before taking the assessments.

| Variables | Mode | Median | Mean | Std.Dev | Skewness | Kurtosis |
|---|------|--------|------|---------|----------|----------|
| 6. I do not feel I will be critical in assessing a classmate's writing. | 3.00 | 3.00 | 3.05 | .999 | .244 | .374 |
| 7.I do not feel I will be able to rate a classmate 's story writing accurately. | 2.00 | 3.00 | 2.65 | .875 | .274 | -.781 |
| 8.I think I will not be critical in my self-assessment. | 1.00 | 3.00 | 2.75 | 1.372 | .093 | -1.261 |
| 9.I do not feel I will be able to rate my writing accurately. | 3.00 | 3.00 | 2.90 | 1.165 | -.231 | -.799 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

In terms of the perceived incapability to self-assess, variable 8: *I think I will not be critical in my writing self-assessment demonstrated a mean score of (M=2.75)*, and a mean score

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

of (M=2.90) on variable 9: *I do not feel I will be able to rate my story writing accurately* showing that prior to the self-assessment, the majority were confident that they would be critical in their self-assessment, and would rate their writing accurately.

Similar to the pre-questionnaire, this part includes two sections: feelings about self and peer assessment and the perceived incapability to self and peer assess writing as a second language.

Learners' feelings about assessment after taking the assessments. Table 4 demonstrates the descriptive statistics of learners' feelings about self and peer assessment in second language writing. The skewness and kurtosis values show that the values are close or within the satisfactory range of normal distribution.

Table 4 shows learners' levels of the following: excitement, confidence, embracement, and their perceived fairness of the assessments. The mean values table 4 show that, after taking the assessment, the majority were excited about the assessments as shown on variable 1: (M=3.30); also, learners were confident to self-assess as shown on variable 2: (M= 3.30), and also confident to peer-assess as can be seen on variable 3: (M= 3.35).

As for the learners' perceptions towards how unfair the peer rating was, variable 4, *I think ratings resulting from peer assessment were not fair*, shows that the mean score resulted was (M=2.90), which was under the median value of (3). This means the majority thought that rating was fair

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 4.

Learners' feelings about assessment after taking the assessments

| Variables | Mode | Medina | Mean | Std.Dev | Skewness | Kurtosis |
|--|------|--------|------|---------|----------|----------|
| 1.I was excited about self and Peer assessment. | 3.00 | 3.00 | 3.30 | 1.031 | .282 | -.945 |
| 2.I felt confident to self-assess my story | 4.00 | 3.00 | 3.30 | .865 | -.119 | -.726 |
| 3.I felt confident to peer -assess a classmate's essay | 4.00 | 3.00 | 3.35 | .933 | -.377 | -1.077 |
| 4.I think ratings resulting from peer assessment were not fair | 2.00 | 3.00 | 2.90 | 1.119 | .632 | -.481 |
| 5.Having had my peer assessing my story embarrassed me. | 2.00 | 3.00 | 2.05 | 1.046 | .466 | -.637 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

The lowest mean score was (M= 2.05) and was on variable 5; having *had my peer assessing my story embarrassed me*. This means that the vast minority of learners thought they did not feel embarrassed to have their writing peer-assessed. These positive attitudes were also backed up by the low mode values of 2.00 on variables 4 &5.

Post perceptions of learners' capability to self and peer assess. This part in the post questionnaire revealed the learner's perceived incapability to be critical and accurate to self and peer-assess second language writing. As mean values shown in table 5 are all under 3.00, meaning that most learners thought they were not incapable of being critical and accurate at rating during the self and peer assessment. In table 5, the descriptive statistics of the learners' post-perceived incapability to self and peer assess show all variables had skewness and kurtosis statistics within or close to the satisfactory range (+/- 1).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 5

Learners' capability to self and peer assess after taking the assessments

| Variables | Mode | Median | Mean | Std.Dev | Skewness | Kurtosis |
|---|------|--------|------|---------|----------|----------|
| 6.I think I was not critical in my writing self-assessment. | 2.00 | 3.00 | 2.70 | 1.081 | .394 | -.466 |
| 7.I do not feel I critically peer assessed my classmate's writing | 2.00 | 3.00 | 2.75 | 1.118 | .298 | -.805 |
| 8.I do not feel I rated my writing accurately. | 2.00 | 3.00 | 2.65 | 1.089 | -.021 | -1.310 |
| 9.I do not feel I rated a classmate's writing accurately | 2.00 | 3.00 | 2.35 | .813 | .541 | .190 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

The mean values on this section came out as follows: ($M=2.70$) on variable 6: *I think I was not critical in my writing self-assessment*, ($M=2.75$) on variable 7: *I do not feel I critically peer-assessed a classmate's writing*, ($M=2.65$) on the variable: *I do not feel I rated my story writing accurately*, and ($M=2.35$) on the variable *I do not feel I rated a classmate's story writing accurately*.; the mode score on all these variables came out as ($Mo= 2.00$). As all variables were to elicit perceptions about how incapable learners were to self and peer assess, the mean values below the median value of 3 on all variables show that most participants perceived themselves capable of self and peer assessing writing.

Attitude changes from pre- to post-questionnaire. Descriptive statistics of the pre-and post-mean and SD values are reported to show the attitude changes of learners towards self and peer assessment. Qualitative data from open-ended questions in the questionnaire were also utilized to elaborate on the changes. To further reveal any potential attitude changes, thematic analysis was used to analyze the interview data based on the predetermined categories of the attitudes (feelings about self and peer assessment, and perceived incapability to self and peer-assess).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Change of learners' feelings about self and peer assessment. Table 6 demonstrates the Mean and Std. Dev values of learners' feelings in the pre-and then in the post questionnaire. The pre-and post-mean values of variable 1: *Feeling excited about self and peer assessment* shows no attitude change as the mean score (M=3.30) on the pre-questionnaire and remained the same (M=3.30). Yet the post-Std. Dev value (SD=1.031) is more spread out from the mean than in the pre-questionnaire (SD=0.87), suggesting that there were probably some outliers in the post questionnaire.

Variable 2, *feeling confident to self-assess*, had a higher post mean score of (M=3.30) than its pre- mean score (M=2.90). This indicates that the learners felt more confident to self-assess after they experienced taking self-assessment. As variable 3, *Feeling confident to peer assess*, learners' attitudes did not change; this variable got a mean score of (M=3.35) on both pre- and post-questionnaire; however, the post-Std. Dev value (SD=0.933) shows that the scores were more clustered around the mean than in the pre-Std. Dev value (SD=1.182).

Table 6

Attitude change Change in the learners' feelings about self and peer assessments

| Variables | Median | Pre-Mean | Pre Std.Dev | Post Mean | Post Std. Dev |
|---|--------|----------|-------------|-----------|---------------|
| 1. Feeling excited about self and peer assessment | 3.00 | 3.30 | 0.87 | 3.30 | 1.031 |
| 2. Feeling confident to self-assess | 3.00 | 2.90 | 1.021 | 3.30 | 1.031 |
| 3. Feeling confident to peer assess | 3.00 | 3.35 | 1.182 | 3.35 | .933 |
| 4. Perceived unfairness of peer rating | 3.00 | 3.45 | .999 | 2.90 | 1.119 |
| 5. Feeling embarrassed to have your writing peer assessed | 3.00 | 3.20 | 1.24 | 2.05 | 1.046 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 6 also shows that the learners' *perceived unfairness of peer rating* on variable 4 dropped down after learners took the peer assessment with a mean score of (M=2.90) after it had been (M=3.45) in the pre- questionnaire. The same drop took place on variable 5: *Feeling embarrassed to have your writing peer assessed*, values show that the participants felt less embarrassed to have their writing peer-assessed in the post questionnaire with a mean score of (M=2.05) compared to (M=3.20) in the pre- questionnaire.

Qualitative data from open-ended questions showed a noticeable attitude change in the learners' confidence; in their responses to the question *Describe how you feel after self and peer assessing the story writing assignment*, the most recurrent answer was feeling more confident and sure about the assessments; the following are some excerpts from learners' responses:

S1: I feel more confident in assessing assignments

S2: I think I am now self-confident to review writing.

S3: Now I can do it again with less hesitation.

Interview data demonstrated that self and peer assessment involvement caused positive effects on the learners' attitudes. Learners expressed three major positive feelings about self and peer assessment: feeling more confident, feeling less nervous, and feeling more entertained. With some interview excerpts, the following three sections show the learners' attitude changes regarding their feelings and how it affected their capability to self and peer assess.

Feeling more confident. Feeling more confident was the most recurrent positive feeling that learners formed after taking self and peer assessment. Many stated that the experience cleared confusions and uncertainties that they had had before they self and peer-assessed the writing task. Others stated that they felt confident because they were able to see what is good and bad about their writing:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: I was not really confident with my writing, and then whenever the teacher said like to peer assess it, I went back and like I critique my writing, but at the same time I also found what was good about it, so I think I benefited both with like felt more confident, but I also found what was wrong with my writing.

S2: I felt confident about my paper, and then after I felt relieved because I understood my paper, my strengths and weaknesses.

Feeling more confident was expressed by others to be caused by their feeling that the experience would improve their writing's final product. They stated that self and peer assessment made them able to improve their writing and make it look much better in the revised version:

S1: I felt confident, I felt like I was going to have perfect writing.

S2: Well at the beginning, I didn't really want to look over my writing because it just thought there were too many mistakes, but then after I looked over it, I thought once I was like maybe I can improve here, but here is good, but after that I just realized writing like it constantly can be improved no-matter-what, so I think it changed my attitude of the whole writing thing, I think it improved my writing a lot.

Feeling less nervous. Feeling nervous about having their writing seen and peer-assessed by other classmates was one of the concerns learners expressed in the interviews. They said that, in the beginning, they had felt nervous about what their peers would think of their writing product; their uncertainty of how their peers would react to their writing sparked their nerves. Yet, the level of nervousness dropped down after the learner's peer-assessed their classmates' writing. This drop-in in tension occurred because, according to the participants, they were able to see and experience peer assessed writing:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: I also, before I took it, I was like nervous, because it did not know how other people would like to react to mine, but at the same time, I wanted to see how I react to other people's stories, so I think it benefited me as well, even after I found way to improve my writing based on their writing, so I think both ways I benefited with self and peer.

Another reason to feel less nervous afterward was that while they were peer assessing, they realized that their classmate's papers were also not perfect either:

S1: I was a little nervous because I did know how people would like to respond to the story, but I was happy we did because afterwards, I felt more relieved because I could see like their writings are also imperfect.

Feeling entertained. Self and peer assessments were enjoyable to some participants, they expressed that the experience was not only beneficial but also enjoyable. As one learner said:

S1: Self and peer assessment were not just useful, but I also really enjoyed the experience.

Others expressed that they enjoyed taking peer assessment more; they thought that reading others' writing was entertaining and beneficial at the same time. One learner stated:

S1: I liked reading their story because they could have been funny, and actually wanted like read more about because someone could like what they wrote, and it was also, like fun to like improve other people' writing, and tell them where to work on, and then you can also reflect back on your paper.

Changes in learners' perceived capability to self and peer assess. Table 7 shows the pre- and post-Mean, and Standard deviation values of learners' perceived incapability to self and peer assess second language writing. As for peer assessment, learners' perceived incapability to critically peer- assess on variable six went down from (M= 3.05) in the pre- questionnaire to (M=2.70) in the post questionnaire; however, the mean score went up from (M=2.65) in the pre

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

to (M=2.75) in the post questionnaire on the learners' perceived incapacity to rate a peer's writing accurately on variable 7. Therefore, learners thought they were more critical in peer assessment than before taking the assessment; conversely, their capability to rating peer's writing accurately turned out to be less in the post questionnaire. This can be explained in two ways: either their ability to rate peer's writing accurately had slightly been overestimated in the pre-questionnaire, or that accurate rating was actually a challenge.

Table 7

Change in the learner's perceived capability to self and peer assess

| Variables | Median | Pre-Mean | Pre-St. Dev | Post Mean | Post St. Dev |
|--|--------|----------|-------------|-----------|--------------|
| 6. Feeling incapable to critically peer assess | 3.00 | 3.05 | .999 | 2.70 | 1.081 |
| 7. Feeling incapable to rate a peer's writing accurately | 3.00 | 2.65 | .875 | 2.75 | 1.118 |
| 8. Feeling incapable to critically self-assess own writing | 3.00 | 2.75 | 1.372 | 2.65 | 1.089 |
| 9. Feeling incapable to rate own writing accurately | 3.00 | 2.90 | 1.165 | 2.90 | .813 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

As for self-assessment, learners' perceived incapability to *critically self-assess their writing* on variable eight dropped slightly in the post questionnaire with a mean score of (M=2.65) than in the pre- questionnaire (M=2.75). The learners' perceived incapability to *rate their writing accurately* on variable nine remained steady regarding the mean score (2.90) on both questionnaires, but with more convergent Std. Dev value (SD=.813) in the post questionnaire than in the pre one (SD=1.165). These values demonstrate that the learners' perceived capability to critically self-assess their writing slightly increased after self-assessment. Also, even though the Mean score did not change on this item: *Feeling incapable of rating own*

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

writing accurately on variable 9, the Std. Dev of (SD=813) on the post questionnaire shows more consistency of learners' scores around the Mean compared to the Std. Dev of (SD=1.165 in the pre-questionnaire.

Some interview data turned out to explain the findings shown in Table (7) that shows that the incapability to peer rate raised in the post questionnaire (M=2.75) as opposed to (M= 2.65) in the pre-questionnaire:

Difficult to peer rate. Some learners in the interviews reported that peer rating was a challenge because disputing self-given numbers was difficult; this problem occurred when learners peer-assessed the peer's writing to which self-assessment results was attached:

S1: I think the rating thing was a hard thing to do; looking at my peer's rating made it hard on me to give numbers.

S2: Well, the one with self-assessment was ok, but rating took me some time, usually number are scary to me

However, the feeling to be more capable of peer assessing writing after experiencing peer assessing classmates' writing was obviously expressed by learners during the interview. Many learners stated that they had been less certain about their ability to peer assess their peers' writing satisfactorily; yet, having had the chance to peer assess two stories boosted their belief in being able to peer-assess writing effectively. A major reason that seemed to be behind enhancing the learner's ability to self and peer-assess was the design of the assessments' forms. The forms were analytically built to be aligned with the teacher's analytical rubric--the one the teacher shared with learners for the writing task:

S1: Before self and peer assessment, I was kind of confused cuz I did not how would like assess my peer's work, if someone were to give to me, I would not know what to do, but now since I

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

like we did that chart of assessment thing, now I know what to look at if I have to assess someone else' work.

S2: I liked how the teacher gave us this chart to rate it, like it is kind of a rubric, instead of just finding mistakes, so I think I really enjoyed the process of going back to the writing, finding if it like how the climax of the story, described a lot of details about the characters, I think it was a good experience with both self and peer assessment.

Learners' perceived importance of self and peer assessment, and the washback effects on revision.

This section includes findings to address the second questions of the study: *the importance of self and peer assessment in revising second language writing:*

A: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing?

B: Are there any washback effects of self and peer assessments on the participants to revise their writing?

Descriptive statistics are reported to show an overview of the data distributions of the learners' pre- and post-perceived importance of self and peer assessment in revising second language writing. Additionally, descriptive statistics are also utilized to demonstrate changes in the learner's perceptions towards the importance of self and peer assessment in revising writing.

Qualitative data from open-ended questions in the post questionnaire and from interviews are reported to demonstrate the washback effects of self and peer assessment on learners to revise their writing.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

This part is divided into three subparts: pre-perceived importance, post-perceived importance, and the perception changes of the importance of Self and Peer assessment.

Pre-perceived importance of self and peer assessment to revise writing. Six questionnaire items were used to address learner's perceptions on how important self and peer-assessment can be in revising second language writing. The pre-questionnaire descriptive statistical values in table 8 show that the skewness and kurtosis are within the acceptable range (± 1); therefore, the distribution of values is considered within the normal distribution as no skewness values are beyond the absolute value of three, and no kurtosis is beyond the absolute value of ten (Kline, 2011).

Table 8

Pre-perceived importance of self and peer Assessments

| Variables | Mode | Median | Mean | Std. Dev | Skewness | Kurtosis |
|---|------|--------|------|----------|----------|----------|
| 10. Self and peer assessing writing are good ways to learn writing | 4.00 | 3.00 | 4.00 | .725 | .000 | .931 |
| 11. Self-assessment is necessary to revise my writing | 5.00 | 3.00 | 4.15 | .988 | -1.056 | .312 |
| 12. Peer -assessment is necessary to revise my writing. | 4.00 | 3.00 | 3.60 | .94 | -.743 | -.355 |
| 13. I think self and peer assessment will help me fix grammatical errors, and enhance my ideas | 4.00 | 3.00 | 3.45 | .999 | -.024 | -.933 |
| 14. If I had the choice not to self-assess my story, I would not self-assess it | 2.00 | 3.00 | 3.00 | 1.026 | .000 | -.671 |
| 15. If I had the choice not to peer assess my classmate's story, I would not peer assess it. | 2 | 3.00 | 2.65 | .993 | .377 | -1.077 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 8 demonstrates the Mode, Mean, and the Std. Dev values on the six variables dedicated to revealing how learners had perceived the importance of self and peer assessment before taking self and peer assessment. The Mean scores on variables 10, 11, 12, & 13 show that the learners highly valued self and peer assessment to learn and revise writing. The Mean score on variable 10: (M=4.00) indicates that learners did think that self and peer-assessment are two learning tools. The Mean scores on the following three variables 11, 12, & 13 :(Mean= 4.15, 3.60, & 3.45) indicate that learners perceived that self and peer-assessment are important in revising writing and fixing grammar and content in writing.

The last two variables--14&15--asked the learners whether they would not take self and peer assessment if they had the chance not to; for self-assessment, half of the learners (M=3.00) seemed to agree with this choice; for peer assessment shown as variable 15, more learners disputed the idea of not taking peer assessment if they had the choice not to (M=2.65).

Table 8, to conclude, demonstrated a positive pre-perception from the majority of learners towards the importance of self and peer assessment to improve second language writing and revise writing tasks.

Post perceived importance of self and peer assessment to revise writing. Like the pre-questionnaire, six variables were used to address the learner's post-perceptions on how important self and peer assessment were in revising second language writing. Statistics in table 8 show that the skewness and kurtosis are within the normally acceptable range (+/- 1) and, therefore, the distribution of values is considered in the normal range; no skewness values are beyond the absolute value of three, and no kurtosis is beyond the absolute value of ten (Kline, 2011).

Table 9 shows the descriptive statistics of learners' post perceptions on how important self and peer assessment were. The high Mode values on the first four variables, 10, 11, 12 & 13 (Mode=

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

5.00, 4.00, 5.00, &4.00), indicate that learners held favorable post perceptions towards self and peer assessment. Moreover, the Mean scores on the four variables turned to be (Mean = 4.05, 3.80, 3.80, & 3.50), all came above the Median (Median =3.00). This suggests that experience of taking self and peer assessment to revise writing was positive in the perspective of the majority of learners: the majority of learners thought that *self and peer assessing writing were good ways to learn writing, self-assessment was necessary to revise my story writing, Peer-assessment was necessary to revise my story writing, Self and peer assessments helped fix grammatical errors, and enhance my ideas.*

The last two variables asked the learners whether they would not have taken self and peer assessment if they had had the chance not to. For self-assessment shown in variable 14, less than half of the learners (M=2.45) seemed to agree with this choice, yet, for peer assessment represented on variable 15, fewer more learners (M=2.55) said they would not have taken peer assessment if they had had the chance not to.

This suggests that the majority had self and peer assessment experience that made them feel positive to have it again.

Table 9, to conclude, demonstrated a positive post-perception from the majority of learners towards the importance of self and peer assessment to revise writing tasks.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 9

Post -perceived importance of Self and Peer Assessments

| Variables | Mode | Median | Mean | Std.Dev | Skewness | Kurtosis |
|--|------|--------|------|---------|----------|----------|
| 10. Self and peer assessing writing were good ways to learn writing | 5.00 | 3.00 | 4.05 | 1.099 | -1.165 | 1.409 |
| 11. Self-assessment was necessary to revise my writing | 4.00 | 3.00 | 3.80 | 1.152 | -1.406 | 1.986 |
| 12. Peer-assessment was necessary to revise my writing. | 5.00 | 3.00 | 3.80 | 1.105 | -.343 | -1.210 |
| 13. I think self and peer assessments helped me fix my grammatical errors, and enhance my ideas | 4.00 | 3.00 | 3.50 | 1.147 | -.465 | -.399 |
| 14. If I had had the choice not to self-assess my story, I would not have self-assessed it. | 2.00 | 3.00 | 2.45 | 1.05 | .449 | -1.001 |
| 15. If I had had the choice not to peer assess my classmate's story I would not have peer assess it. | 2.00 | 3.00 | 2.55 | 1.05 | .157 | -1.144 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

Changes in the Perceived importance of Self and Peer assessment. Table 10 indicates the pre-and post- Mean and Std. Dev values of learners' perceived importance of Self and Peer assessment in revising second language writing. As can be depicted, there are positive attitude gains on variables 10, 12, & 13 variables, but a negative drop on variable 11. The mean value of Variable 10, *self and peer assessing writing, are good ways to learn writing*, went up in the post questionnaire to (4.05) as opposed to (M=4.00) in the pre-questionnaire. Variable 12: *Peer-assessment is necessary to revise my story writing* had a post Mean value of (M=3.80) in the post questionnaire versus (M=3.60) in the pre one. Variable 13, *I think self and peer assessments help me fix my grammatical errors and enhance my ideas* escalated slightly from (M=3.45) in the pre-

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

questionnaire to (M= 3.50) in the post questionnaire. The Std. Dev values on these three variables look slightly more divergent from the mean in the post questionnaire, yet the values are considered within the satisfactory range of normal distribution. Yet, the drop in the Mean value took place on variable 11: self-assessment *is necessary to revise my story writing*, from (M= 4.15) in the pre-questionnaire to (M= 3.80) in the post questionnaire.

Table 10

Change in the perceived importance of self and peer assessments

| Variables | Median | Pre-Mean | Pre-Std.Dev | Post Mean | Post Std. Dev |
|--|--------|----------|-------------|-----------|---------------|
| 10. Self and peer assessing writing are good ways to learn writing. | 3.00 | 4.00 | .725 | 4.05 | 1.099 |
| 11. Self-assessment is necessary to revise my writing | 3.00 | 4.15 | .988 | 3.80 | 1.152 |
| 12. Peer-assessment is necessary to revise my writing. | 3.00 | 3.60 | .94 | 3.80 | 1.105 |
| 13. I think self and peer assessments help me fix my grammatical errors, and enhance my ideas | 3.00 | 3.45 | .999 | 3.50 | 1.147 |
| 14. If I have the chance choice not to self-assess my story, I would not self-assess it. | 3.00 | 3.00 | 1.026 | 2.45 | 1.05 |
| 15. If I have the choice not to peer assess my classmate's story, I would like peer to assess it. | 3.00 | 2.65 | .993 | 2.55 | 1.05 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

Table 10, moreover, demonstrates learners' pre- and post-responses on whether they would not take self and peer assessment if they have the chance not to do so. As can be seen on variables 14 & 15, there was a drop in the Mean scores in both the self-assessment variable 14 (pre-Mean = 3.00 vs post Mean = 2.45) and for peer assessment variable 15 (pre-Mean Mean = 2.65 vs post Mean = 2.55). As for the pre-and post-Std. Dev, there were no many differences in

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

the values and all Std. Dev values are considered within the satisfactory range of normal distribution.

Table 10, thus, demonstrates the positive perceived importance of the usage of self and peer assessment in second language writing. Statistics show steadiness in the learner's perception towards the following: the importance of self and peer assessment to learn writing, the importance of peer assessment in revision, and the importance of self and peer assessment to fix some writing features in writing revision and willingness to take the assessments in the future.

However, the quantitative data in table 10 demonstrated a drop in the learner's perceptions towards the necessity of using self-assessment in writing revision. Qualitative data revealed a negative washback effect about self-assessment that can be linked to why the importance of self-assessment in revision was perceived less important in the post-questionnaire than it had been perceived in the pre-questionnaire. While self-assessment created positive washback, it also generated one negative washback effect on learning:

Difficult to objectively self-assess writing. Some learners found that it was a little difficult for them to be objective in their self-assessment. As explained in chapter 3, Self-assessment had been taken before peer assessment so that learners assess their own writing performance and hand the results to their peers in the peer assessment; these forced learners to be serious in their self-assessment. However, some learners reported that it was hard to self-assess and have the third-person perspective:

S1: I don't feel comfortable self-assessing myself because I could either be way too harsh on myself or extremely cocky and believe that my paper is the best thing anyone has ever written.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: when you're looking at your own work, your kind of like are you, like, you think in your mind that yours is already kind of good enough because you read it a couple times and It's not as easy to catch those mistakes

It seems that the incapability to critique one's own writing through the readers' perspective made the self-assessment a challenging task for some learners in this study:

S1: I did not know how to self-assess my story accurately.

This finding from qualitative data can be a justification for the quantitative findings regarding self-assessment. For the item on table 10, Self-assessment is necessary to revise my story writing, the Mean score in the post questionnaire dropped to 3.80 after it had been 4.15 in the pre-questionnaire.

However, self-assessment, in addition to peer assessment, created several washback effects on learning. Those washback effects are presented in the next section, and those effects are specifically linked to how much they affected revision.

The washback effects of self and peer assessments on learners to revise their writing

This section addresses part B of questions 2: *Are there any washback effects of self and peer assessment on the participants to revise their writing? If so, what are they?*

A close investigation and categorization of all qualitative data sources--10 interviews and questionnaire open-ended questions revealed several washback effects of self and peer-assessment on learners to revise their writing.

Washback effects of self and peer assessment on revision. As stated by Cheng and Curtis (2004), positive washback effects appear when participants seem to hold positive attitudes towards the assessment; consequently, participants involve themselves desirably and

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

collaboratively in the assessment. In this study, several washback effects from self and peer assessment appeared to the surface. Some washback effects were found to be caused by either self or peer assessment; other washback effects seemed to be caused by both assessments.

The washback effects concluded from the qualitative data are these: helping to meet the writing genre requirements, bridging writing gaps through reading, editing through self-assessing, and expanding lexicon reservoir through peer assessment, becoming more creative through peer assessment, and building metacognitive awareness.

Helping to meet the writing genre requirements. Since self and peer assessment forms addressed the components of the writing genre learners have to meet in the writing task (story or narrative), learners expressed that the forms were effective to address the components of the story writing genre. This section in the self-assessment form was detailed and specific at what story elements to consider and what details to add to each element:

S1: I liked how the teacher gave us the form, it is kind of a rubric, instead of just finding mistakes, so I think I really enjoyed the process of going back to the writing, finding if, it, like, how the climax of the story, described a lot of details about the characters, I think it was a good experience with both self and peer assessment.

S2: The forms, you can make your climax like more suspenseful maybe instead of just putting it straight into the story,

S3: I think they were helpful because they went over the different areas I need to correct.

Self and peer assessment forms scaffolded learners to address the story writing genre requirement; this was appealing to learners. It helped them move smoothly from one point to another and, therefore, process the story and its requirements more effectively:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: I liked the form when it went through each point, for example, the climax, and so I will go specifically to the climax and read that, it helped me go specifically in pinpoints, or like go to each section, and make sure each section is good, so instead of reading the whole thing and try to figure it out.

S2: I think I liked going specifically on the components of the story, it made my thoughts organized, like pointing out specifically the climax or resolution or rising actions, so I think pointing out specifically.

Self and peer assessments' forms informed the learner's first draft writing performance, and the rated scale with its descriptors informed the learners about how satisfactory their performance in each component was. These analytically organized genre components kept learners aware of their writing performance:

S1: I saw that there should be improvement, like, more emotions with the characters, or like the ending, making the ending to be more suspenseful.

Therefore, it was clear that self and peer assessment effectively set up learners to address the writing task genre requirements through the good design of assessment forms; learners' ideas were kept organized to meet the genre requirements.

Bridging writing gaps through reading. In addition to fulfilling the genre requirement through the forms, those missed requirements, or to be called here writing gaps, are bridged through the action of reading. Reading over one's own writing and reading peers' writing who wrote within the same writing genre frame can reveal writing gaps; this was clearly reported by a number of learners. During self-assessment, learners read their first draft, and they were able to catch the writing gaps they had not noticed in the first draft. As for peer assessment, having had read a peer's writing enabled the learners to reflect on their writing gaps:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: Um, self-assessment, I think I like going over my work to see what I needed to fix and what I didn't catch the first time when I read it, and then peer assessment, I like to read in other people's stories and like seeing what I could fix on my own and also what the other person could fix.

S2: So, for self-assessment, I liked how like, basically after I wrote everything in my story, I can like read it back and see what I would have done wrong, and I should have fixed, so it was like me seeing as a third person just like someone else sees my story.

The reflection of one's own writing through reading peers' writing was more likely because learners had to meet a number of story writing requirements such as description, logical sequences of actions, etc. And thus, the chance to see how peers did with those requirements was an eye-opening learning experience:

S1: I know one of the stories I peer assessed was really descriptive, and the other one was like the plot was good like the sequence was really good, so it helped me realize like oh your story is not in any order and was not really descriptive, and so I could understand what was going on at that time so yea, I think I really enjoyed it, and also going back, and just reading other people's stories help me get mine better, and so I liked that too

S2: I really liked to do like whatever I was reading other's stories, it really helped me build my plot, like it helped me realize like oh my ending is kind shaky, it needs a little bit more help, or like the beginning like oh I was not as descriptive as other people were, and I need to add more description, so I really enjoyed it.

S3: I think some of the things I saw other people doing, I realized I probably should have done that too.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

So, both self and peer assessment provided chances for active reading where learners can detect their own writing gaps autonomously and then enhance and better and stitch those gaps later on in their revised writing version.

Editing through self-assessing and expanding lexicon reservoir through peer assessment. Self-assessment was found to heavily prompt learners to catch their grammatical errors. Learners reported that when they self-assessed their own writing, they were able to spot grammatical errors, such as spelling or wrongly tensed sentences:

S1: I think for self-assessment, it kind of helped catch structural mistakes, spelling and grammar mistakes, I would change present to past and so things like that.

S2: so, for self-assessment, it helped with grammatical mistakes because when I read it over, I could see any issues with grammar.

S3: When I self-assessed, I found that they helped a lot, because they made my paper better like I fixed the grammar.

Although the correct tenses to address dialogues and narrations in the story—the writing task--were promoted in the forms of self and peer assessment, the majority of learners addressed the tense correctness in their self-assessment:

S1: For self-assessment, I liked the form when it went through each point, for example, grammar, so I would read through it and look specifically for grammar mistakes.

S2: I could see that I had many problems with present and past verbs, and self-assessment kind of helped fix that.

Peer assessment, on the other hand, appeared to help learners learn new vocabularies and phrases:

S1: I think peer helped build my vocabulary a little bit.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: I can see words I never thought of, and used them, my peer's writhing showed me many words I really needed.

Moreover, seeing many other descriptions through peer assessing scaffolded the learners to be more descriptive in their revised writing version; they would seemingly copy and/or assimilate their peers' descriptions to fix their own:

S 1: Peer assessment helped me fix my content and add descriptions to my story.

S 2: I did my best to say things about the characters and things in my story, and peer assessment helped me improve those descriptions and add many more.

A peer's feedback, likewise, on the content made problems about descriptions appear to the surface. The first writing version is usually filled with vague phrases and descriptions, and one may not be aware of those vague descriptions even if it is self-reviewed, but when those weakly expressed phrases are viewed through another reader's perspective, they are easy to spot:

S1: I think may be going back on my own story, and looking at my sentences and also like, because like other people did not really know what I was trying to say with my story sometimes, it made me have to come back and realize like oh it is really clear there, oh like I need to go back there and like add more descriptions like it very foggy what is going on in here, so, it helped me like see my weaknesses, so yea.

S2: I, like, we have different grammatical knowledge at different levels so like some of my peer may know specific like some concepts a lot better than I do so that was like helpful.

In short, it seems that self and peer assessment contributed to different writing areas in the learners' revised writing versions. Self-assessment helped them realize and improve grammatical gaps, while peer assessment helped enhance content through more lexical variety

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

and descriptions. It can be argued here that the implementation of both self and peer assessment worked in a complementary way.

Becoming more creative through peer assessment. In addition to expanding lexicon reservoir, peer assessment provided learners with new ideas to integrate in their own writings. A number of participants reported that during peer assessing their peers' writing, they would learn some new thoughts and ideas to use it in their story writing later on:

S1: I also like the peer one because it gave me ideas like what I have not thought about my own.

S2: I feel that my writing improved, and I got new ideas from peers and fixed old mistakes.

Thus, this experience of peer assessing classmates' stories generated new thoughts that led them to be more creative with their own story writing:

S3: But with peer assessment it kind of helped my writing with being more creative, being open to more ideas.

Likewise, the inspiration for creativity took place when participants received feedback from peers. Some learners reported that they received inspiring comments and feedback from their peers that prompted them to enhance the level of creativity of their original thought:

S1: so, I understood my story, but when someone reads it, like sometimes, like when an outsider, they do have to see my exact idea that I had, so it is not like as clear to them, so like, to me, when I was reading it was so clear, beautiful, someone, like, when my peer was reading it, they would sometimes say hey, this can be better, so that was really helpful because I would have caught that with their comments, and make ideas much better.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: I think yes cuz someone might point out something I did not notice before, or they can give me more ideas to improve my story that I might not have thought of.

To conclude, peer assessment had positive washback back effects by inspiring learners to be more creative in their writing and prompting learners to amend some ideas for betterment.

Building metacognitive awareness. Having the chance to self-review their own writing as well as to peer review other's writings made learners aware of their own writing properties, their strengths and weaknesses. The two assessment experiences created a metacognitive awareness in a way learner became familiar with what they are able and not able to do with writing:

S1: I think the whole thing helped in learning my style of writing, components

S2: Then, like, the forms showed me what I was better at and like my weaknesses, so I think it is actually it was a good way to improve your story before submitting to the teacher, it like really helps.

S3: I understood my paper, my strengths and my weaknesses.

Being aware of their writing strengths and weakness made learners understand their own writing features, and this let them to feel both comfortable and motivated:

S4: Self and Peer assessment told you what your strengths were and so that motivated me to write and improve, I like that.

S5: I felt relieved because I understood my paper, my strengths and my weaknesses.

S6: I feel like self-assessing showed what I can do and cannot do with my writing, I feel like this kind of made me feel ok, and not to stress out.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

In short, the experiences of self-assessing, peer assessing and being peer assessed helped learners to learn about their writing properties, weaknesses and strengths which led them to improve their writings and feel motivated to write and learn more about writing.

Learners' perceived effectiveness of using self-assessment with peer assessment and its washback effects on learning

This section reports findings to address the third question of this study:

The effectiveness of considering and not considering the peer's self-assessment results in peer assessment:

A: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

B: Are there any washback effects of considering and not considering peer's self-assessment results in peer assessment? If so, what are they?

Descriptive statistics demonstrate an overview distribution of the learners' pre- and post-perceptions about the effectiveness of considering and not considering peer self-assessment results in peer assessment. Moreover, descriptive statistics are also presented to show the changes in the learners' perceptions (from pre to post).

Qualitative data from open-ended questionnaire questions and from interviews was synthesized to demonstrate the washback effects on learning by considering and not considering peer self-assessment results in peer assessment.

This part is divided into three subparts: pre-perceptions, post-perceptions, and changes in the perceptions about the effectiveness of considering and not considering the results of peer' self-assessment results in peer assessment.

Pre-perceived effectiveness of considering and not considering the results of a peer writer's self-assessment in peer assessment. Table 11 demonstrates the descriptive statistics of the learners' perceived effectiveness of considering the results of the peer self-assessment during peer assessment. The distribution of the skewness and kurtosis for the five variables in the table are considered to be within the satisfactory range of normal distribution (+/-1). The skewness and kurtosis values were both considered within the guideline frame by Kline (2011), i.e., no skewness values are beyond the absolute value of three, and no kurtosis is beyond the absolute value of ten.

Table 11 includes the five variables that aimed to capture the learners' perceptions about the contribution of peer self-assessment results in peer assessment. The questions were posed to the learners from two perspectives: as a peer assessor and a peer-assessed as all learners did peer assessment and were peer-assessed. Thus, it is inevitable to examine the perceptions through those two scopes.

The first two variables—16 & 17--in table 11 indicate how helpful the consideration of self-assessment in peer assessment from the peer assessor and peer-assessed standpoints. The mean values on both variables--16 &17--show that most learners thought peer self-assessment results could be helpful both ways: as a peer assessor and as a peer-assessed. However, the mean value is higher from the peer assessor's standpoint, variable 15 (M= 3.45), than from the peer assessment standpoint, variable 16- (M=3.15).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 11

Pre-perceived effectiveness of considering, and not considering the results of a peer writer's self-assessment in peer assessment

| Variables | Mode | Median | Mean | Std.Dev | Skewness | Kurtosis |
|--|------|--------|------|---------|----------|----------|
| 16. As a peer assessor, having the results of a classmate's self-writing assessment is helpful to peer assess a classmate's writing | 4.00 | 3.00 | 3.45 | .887 | -1.592 | 1.854 |
| 17. Having my peer consider my self-assessment results may lead to better peer assessment results in my writing | 3.00 | 3.00 | 3.15 | .875 | .732 | .402 |
| 18. As a peer assessor, having the results of a classmate's self-writing assessment may restrict my peer assessment. | 3.00 | 3.00 | 3.00 | .973 | .000 | -.159 |
| 19. As a peer assessor, I do not feel comfortable disputing a classmate's self-assessment results. | 4.00 | 3.00 | 3.10 | 1.021 | -.218 | -.586 |
| 20. If I had the choice not to take the self and peer assessments together, I would only take one of them to assess my story. | 3.00 | 3.00 | 2.90 | .968 | .217 | -.060 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

As for standard deviation values for the two variables, variable 15 had std. Dev (SD=.877) and variable 16 had Std. Dev (SD=.875); those two values show an acceptable normal distribution of values around the means.

The third variable in table 11 --variable 18--refers to how restrictive a peer's self-assessment results can be during peer assessment; the mean value shows that half of the learners

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

thought it is restrictive ($M=3.00$). Variable 19 states that most learners felt that they would feel uncomfortable disputing their peer assessment results during peer assessment ($M=3.10$).

Finally, variable 20 shows that less than half of the learners demonstrated their unwillingness to take both self-and peer- assessments to assess writing ($M=2.90$). Examining the standard deviation for the last three variables – 17, 18, &19, they had values of $Sd=.873$, $Sd=1.021$, and $Sd=.968$, which indicate an acceptable normal distribution of values around the means.

Thus, the pre-questionnaire in table 11 shows that learners, in general, demonstrated an acceptance to use peer's self-assessment results in peer assessment, but they are a little uncomfortable with disputing their peer's self-assessment results. Also, half of them expressed their expectations that peer' self-assessment results could restrict their peer assessment.

Post-perceived effectiveness of considering or not considering the results of peer self-assessment in peer assessment. Table 11 shows the descriptive statistics of the learners' post perceived effectiveness of considering the peer's self-assessment results in peer assessment. Skewness and kurtosis values are considered within the satisfactory range of normal distribution (± 1). Even though some items were beyond ± 1 , the skewness and kurtosis values were both considered within the guideline frame by Kline (2011), i.e., no skewness values are beyond the absolute value of three, and no kurtosis is beyond the absolute value of ten. That is, the observed variables were seen as normally distributed.

Table 12 includes the five variables that aimed to capture the learners' perceptions about the consideration of peer's self-assessment results in peer assessment. The questions were posed to the learners from two perspectives: as a peer assessor and a peer-assessed as all learners did peer assessment and were peer-assessed.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The first variable--16-- asked learners how helpful the peer's self-assessment results from the perspective of a peer assessor: *As a peer assessor, having the results of a classmate's self-story assessment helped me peer assess the classmate's story*. The mean value came out to show that more than half of learners thought it did help: $M=3.10$. As for the perspective of being peer assessed--variable 17-- *Having had my peer consider my self-assessment results helped my peer assess my writing*, the mean score came out as $M=3.50$ to indicate that the vast majority of learners thought results of their own self-assessment helped their peers assess their writing. It can be that participants were satisfied with the peer's quality they received--the peer's assessments to which they attached their self-assessment results.

Variable 18, *As a peer assessor, having had the results of a classmate's self-story assessment restricted my peer assessment*, shows that more than half of the participants thought that the peer's self-assessment results were restrictive in peer assessment ($M=3.15$). This makes sense since relying on other sources can restrict one's sources.

Variable 19, *As a peer assessor, I did not feel comfortable disputing a classmate's self-assessment results*, demonstrated that less than half of learners felt uncomfortable in disputing their peer's self-assessment results ($M=2.45$). Hence, based on the finding of this variable19, the element of interpersonal relational restriction is excluded from being the reason why some learners felt restricted during peer assessment considering their peer's self-assessment results.

The last variable--20--in table 12, *If I have the choice not to take the self and peer assessments together, I will only take one of them to assess my story*, shows that the experience of considering a peer's self-assessment results was that more participants found it helpful and few found it not helpful. ($M=2.70$). Thus, this indicates that the overall experience of considering the peers' writer's self-assessment results had a lasting positive effect on learners.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Table 12.

Pos-t perceived effectiveness of considering, and not considering the results of a peer writer's self-assessment in peer assessment

| Variables | Mode | Median | Mean | Std. Dev | Skewness | Kurtosis |
|--|------|--------|------|----------|----------|----------|
| 16.As a peer assessor, having the results of a classmate's self-writing assessment helped me peer assess the classmate's writing | 4.00 | 3.00 | 3.10 | .968 | -.604 | -.850 |
| 17.Having had my peer consider my self-assessment results helped my peer assess my writing. | 4.00 | 3.00 | 3.50 | .889 | -1.251 | 2.285 |
| 18.As a peer assessor, having had the results of a classmate's self-writing assessment restricted my peer assessment | 3.00 | 3.00 | 3.15 | .875 | .732 | .402 |
| 19.As a peer assessor, I did not feel comfortable disputing a classmate's self-assessment results. | 2.00 | 3.00 | 2.45 | .945 | 1.409 | 1.821 |
| 20.If I have the choice not to take the self and peer assessments together, I will only take one of them to assess my story. | 2.00 | 3.00 | 2.70 | 1.261 | .109 | -1.252 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

Change in the perceived effectiveness of the results of peer's self-assessment in peer assessment. Table 13 demonstrates the pre-and post-Mean and Standard deviation values of learners' perceived effectiveness of considering and not considering the results of a peer's self-assessment results in peer assessment. As mentioned previously in table 10 and table 11, the questions in this part of the questionnaire were posed to elicit answers from two perspectives: peer assessor and peer assessed.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The first variable--16--: *as a peer assessor, having the results of a classmate's self-story assessment is helpful to peer- assess a classmate's story*, showed a drop in the mean value in the post questionnaire (M= 3.10) as opposed to mean value in the pre- questionnaire (M=3.45). Now it is clear that fewer learners in the post questionnaire believed in the effectiveness considering the peer's self-assessment results in peer assessment. However, even though the post mean value (M=3.10) came out smaller than the pre- mean value (M=3.45), the post value still represents the majority. The standard deviation values on the pre-and post- mean clearly show a normal distribution of values around the means.

Conversely, variable 17: *Having my peer consider my self-assessment results may help my peer assess my writing demonstrated* a rise in the mean value in the post questionnaire (M=3.50) compared to the mean value (M=3.15) in the pre- questionnaire. This rise in the mean value in the post questions suggests that more learners were more satisfied with the results of peer assessment they received in their writing considering their self-assessment results. That is, the vast majority of learners were happy with the peer' feedback they got with the consideration of their self-assessment results. Thus, the findings on the first and second variables--16 & 17-- propose two perspectives: learners found it more helpful to give self-assessment results to a peer in peer assessment than peer assessing with the consideration of a peer' self-assessment results. Learners seemed to like what their peers pointed out and commented on their first draft by considering their self-assessment.

Variable 18: *As a peer assessor, having the results of a classmate's self-story assessment can restrict my peer assessment* showed a slight increase of 0.15 in mean value (M=3.15) in the post questionnaire. This finding can probably explain why the mean value dropped on the first- variable 16 in table 13 as fewer learners favored using a peer's self-assessment results to peer-

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

assess a peer's writing. The restrictiveness could be generated by the idea of considering the comments pointed out in the self-assessment form.

Variable 19: *As a peer assessor, I do not feel comfortable disputing a classmate's self-assessment results* demonstrated a drop in the mean value in the post questionnaire (M =2.45) as opposed in the pre-questionnaire (M=3.00). Hence, considering the findings on variable 18 in which learners showed that peer's self-assessment results were restrictive, and findings on variable 19 excludes a reason why (feeling uncomfortable to dispute a peer' self-assessment results). As expressed in the above paragraph, it is probably that the content of the peer's self-assessment results was restrictive.

Table 13

Changes of the perceived effectiveness considering of a peer writer's self assessment results

| Variables | Median | Pre-Mean | Pre-Std. Dev | Post Mean | Post Std. Dev |
|---|--------|----------|--------------|-----------|---------------|
| 16. As a peer assessor, having the results of a classmate's self-writing assessment is helpful to peer assess a classmate's writing. | 3.00 | 3.45 | .887 | 3.10 | .968 |
| 17. Having my peer consider my self-assessment results may help my peer assess my writing. | 3.00 | 3.15 | .875 | 3.50 | .889 |
| 18. As a peer assessor, having the results of a classmate's self-writing assessment can restrict my peer assessment. | 3.00 | 3.00 | .973 | 3.15 | .875 |
| 19. As a peer assessor, I do not feel comfortable disputing a classmate's self-assessment results. | 3.00 | 3.10 | 1.021 | 2.45 | .945 |
| 20. If I have the choice not to take the self and peer assessments together, I will only take one of them to assess my story | 3.00 | 2.90 | .968 | 2.70 | 1.261 |

(N=20). Based on a 5-point Likert scale from 1 (Strongly disagree) to 5 (Strongly agree)

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Variable 20 in table 13: *If I have the choice not to take the self and peer assessments together, I will only take one of them to assess my story* showed a drop in the mean value in the post questionnaire (M=2.70) than in the pre- questionnaire (M=2.90) suggesting that the experience of considering a peer's self-assessment was generally appealing to learners. Qualitative data synthesized from open-ended questions and interviews are presented in the next section to show the washback effects of considering and not considering the results of self-assessment in peer assessment; these washback effects are also to insightfully explain the findings on table 13.

Washback effects of considering the results self-assessment in peer assessment

The consideration of the peer's self-assessment results in peer assessment created three major washback effects on learning: reinforcing peer's self-assessment critique, expanding on a peer's self-critique, and restricting a peer's decision-making.

Reinforcing peer's self-critique. Several participants expressed the idea that a consideration of a classmate's, or peer's, self-assessment resulted in the peer's self-assessment subconsciously directing their attention to address only the gaps that the peer pointed out in the self-assessment form. They stated that those self-assessment comments made them focus on repeating and reflecting on those comments:

S1: I was looking at what they wrote for themselves and try to help them with that.

S2: For peer assessment with self-assessment, it really helped, like, me go deeper in my peer's self-feedback, I felt like I helped my peer with the problems they wrote about in the self-assessment, so yea.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S3 -Well, the one with the self-assessment, I can see what they thought about their paper, and then I can give my feedback on it.

Also, a peer's self-critique reinforcement was noticed by learners after their peers assessed the first draft; those learners stated that the peer assessment results they received on the two papers-- one to which their self-assessment was attached to, and the one that did not have self-assessment with--were different.

S1 The comments, the one with the self-assessment, they were very similar to the comments I gave myself, and they added relevant comments.

S2 I feel like the one with self-assessment stressed myself comments and also opened my eyes to more ways to fix my mistakes, and the one that was with no self-assessment was sort of different, it was, like, new different ideas.

Expanding on peer's self-critique. In addition to reinforcing the self-critiqued writing areas when they considered the results of their peer' self-assessment results, those peers offered solutions and suggestions on those self-critiqued writing areas. This was noticed by a number of learners in their peer- assessed writing:

S1: yea I think one comment that I put in myself assessment was that may be my ending was not good, and maybe seemed to be a little bit rushed, and so the person that peer-assessed my story with the self-assessment, she considered that, and so, she said oh add this, add that, that will help, that will make it better, so yea, I think self-assessment is really important because like if I have pointed that out, someone else may not figure that out, so it helped out realize and be on the same page.

S2: The one with the self-assessment, they were able to see the things I struggled with and gave me good ideas.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Similarly, some peer assessors noticed that the existence of the peer' self-assessment made them expand on the peer's self-assessment results by giving suggestions and ideas:

S1: But then when I had it with the self-assessment form, I can like guide the story maybe better because I can see like areas, they needed help on, they really, like, for example, they put a lot of time like on the setting, or describing the characters, but they were not able to make the ending very good, then I can give suggestions about those things.

S2: self-assessment form was um, like, I looked at their comments and ratings, and tried to, like, come up with ways to help them fix things.

Restricting peer's decision-making. While a peer's self-assessment results inspired peer assessors to produce new ideas and comments to address writing issues highlighted in the self-assessment, a peer's self-assessment also restricted peer assessors from making autonomous decisions. Many learners reported that as they were peer assessing their classmate's draft, the peer' self-assessment results influenced their assessment decisions and became heavily based on their peer's self-assessment results:

S1: I feel like the first one I feel like looking at what they thought it was wrong with their own writing and kind of use that to influence my decision making,

S2: I feel like my answers were more biased because I was looking at what they wrote for themselves.

The word "*feel restricted*" was one of the most common phrases expressed by peer assessors when considering the peer's self-assessment results. Yet, that feeling of being restricted made peer assessment to more dedicated to providing solutions for self-critiqued writing areas expressed in the peer' self-assessment results:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: so, when I got with the self-assessment, it was like, I felt very restricted because I could not like, I had to take what other people were saying and give feedback on their issues extensively.

S2: I felt restricted with self-assessment, I felt I had to help my peer with they thought they have problems with their story.

Hence, it can be concluded that the existence of peer' self-results in peer assessment restricted learners to only address their peer's self-criticized writing areas. This finding is aligned with the quantitative finding presented in table 12 –variable 18--in which learners showed that self-assessment restriction. Yet, this created a positive washback effect in that it made peer assessment more personalized for the purpose of addressing self-viewed writing issues; however, it also created a negative washback effect in that peer assessors' decisions were restricted.

Washback effects of not considering the results self-assessment in peer assessment

The nonexistence of peer' self-assessment results in peer assessment caused several washback effects: unbiased peer assessment and independent peer's decision making.

Independent peer's assessment. Peer assessing writing without the consideration of a peer's self-assessment results made peer assessment more independent. That is, peer assessors had no pressure to make biased decisions that could be caused by a peer's self-assessment results. Learners expressed that their peer assessment decisions were not based on any predetermined self-assessment results and that it made their peer assessment decisions entirely based on their own opinions:

S1: So, when I did on my own, with no self-assessment, it allowed me to actually do it without a biased opinion or anything.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: the one I could not see what they did was based purely of my own opinion, and so I feel if I did not see it, it would not have influenced my opinion.

Thus, the absence of peer' self-assessment results excluded the chances of having biased opinions in peer assessment decisions; learners had no peer's self-viewed writing "problems that may affect the final peer's decisions.

Fresh peer's perspective. While the absence of peer self-assessment results resulted in independent peer decisions, as shown in the previous point, it also led to making peer assessors have *a fresh start and a fresh perspective* about their peer's writing:

S1: I feel like without the self-assessment, it was like a fresh start, and just do comments without someone else's comments.

S2: When I had it without the self-assessment, I had to do everything on my own, I was reading the story with any help, I did not have any of the author's input or anything, so it was like a fresh set of eyes or something

This fresh peer perspective allowed the peer assessor to explore a peer's writing freely; consequently, peer assessors clearly understood peer writing.

S1: Without my peer's self-assessment, I noticed things they have never said, and so it was clearer to me what is wrong.

The fresh perspective made a peer make more confident peer decisions:

S1: Then, like without the self-assessment, it allowed me to be like, oh this is wrong, maybe they might have thought it was perfect, but now they may go back and think that is wrong, you are now confident about your decisions.

The draft that was given to the peer with no self-assessment resulted in new peer assessment results that they missed addressing in their self-assessment:

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S1: ".....but the one without my self form, it was something completely different, something that I had not thought of".

S2: um. I found new feedback than in the peer form, but the peer form with my self-feedback gave me kind of similar stuff".

In short, the washback effect of not considering a peer's self-assessment results turned out to allow peer assessors to make independent peer assessment decisions and also having a new start perspective to navigate a peer's writing. These two findings can be combined as "independent fresh peer assessment."

The washback effects of self and peer assessment process

One of the big themes that emerged and that is worth mentioning from the interviews is the washback effects of the process of self and peer assessment. As explained in chapter 3, the process was to self-assess their writing and peer-assess two classmates' writing (with the self-assessment and another without self-assessment). This process resulted in three major washback effects.

Getting multiple feedback perspectives. Having three perspectives about their own writing was appealing to some learners. The process allowed learners to receive three sets of feedback from self-assessment and from two peers; this process increased the chances of having various perspectives to address the writing gaps and increase the writing quality:

S1: Self-assessing my paper was really helpful to finalize everything, and then with peer assessment, I had another view on myself writing evaluation, because when someone peer assesses your paper, they look at it in a different way than you, so like my story is perfect, nothing wrong with it, but someone comes and read it and say hey this part is good, but it can be better, and then the other peer have me another new view on my writing so yea

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: I mean I like having both, just that like I can compare my own review with my peers' comments, I think they self and peer assessment both showed me what I needed to do to make my paper look better

Thus, the process resulted in having three perspectives about one's writing (self, peer with self, and peer), and this provided more feedback and more emerging ideas about how to make writing better.

Getting more familiarized with self-writing gaps. The process of first self-assessing one's own writing and then having it peer-assessed was a good order in the learners' standpoint because it gave learners the chance to be familiar with their own writing performance gaps before someone else stressed those writing gaps. This seemed to have a positive washback effect on learning as learners can become more familiar with their own writing errors; subsequently, they become more autonomous to tackle those errors:

S1: I think self-assessment is beneficial to catch your own writing before someone else helps you because if you realize your wrong like your own mistakes, and then people, like I mean my peer confirm my mistakes, it is I think it will help you in future writing than if someone were just to tell you about your mistakes, I think it is beneficial to first look over your writing to see what needs to be changed, or something you u did wrong before someone look at it, so I liked that way that we did first and then peer.

S2: I thought both are really important, self-assessment really made my mind see the problems in my story, and then peer review even brought them up again.....oh yea they are both really beneficial to improve your writing, so... I actually did like it.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Hence, the self and peer assessment process scaffolded learners to be more aware and familiar with their writing issues because those writing issues will probably be addressed again in one of the peer assessment results that learners received.

Making peer assessing easier. It was concluded that self-assessment made peer assessment an easier task. That is, taking self-assessment first and then handing the results to the peer for peer assessment made peer assessment smoother. It seems that self-assessment results facilitated pointing out peer comments and feedback so that peers do not have to do things all over again. Some learners reported that while peer-reviewing a peer's paper with the peer's self-assessment results, they only looked at what the peer thought of her/his paper and tried to help tackle those issues. They claimed that the self-assessment was done properly, and all the writing issues the peers pointed out in their self-assessment were actually there.

S1: I saw that the self assessment form had all what needed to be said, so, um, I looked for those things and just offered some ideas, and yea, so

S2: The peer assessment with self-assessment was easier than the other peer assessment, I think when someone self-assess his own writing, they will make sure to do it properly, and for me, I trusted my peer's self-comments, so mine was like, you know, I figured out some comments to help my peer um, fix his story issues.

When asking learners, *Why do you think self-assessment can make peer assessment better?* Their answers were very logical:

S1: cuz when you do it for yourself, then you are already making it better for the person, and then when it is already better for the person to peer assess it, it will be more improved

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

S2: I like the order of doing the self to peer because you were able to fix as much as you can, and then you give it to your peer, so instead of giving it like to your peer with mistakes that you could have fixed by yourself.

In short, the process facilitated one of the peer assessments' tasks as peers depended on the peer's self-assessment results, and that not only saved time and effort but also expanded on the self-assessment results since peers seemed to focus on those self-assessment results.

Summary

In short, the findings demonstrated that learners' feelings towards self and peer assessment came out to be more positive in the pre-questionnaire. Similarly, the learners' perceived capability to self and peer assesses appeared to be enhanced after taking the assessments. In terms of the perceived importance of considering self- with peer- assessment, quantitative data showed a drop in the necessity of using self-assessment. When looking for an explanation in the qualitative data, it appeared that self-assessment made peers reflect extensively on self-assessment.

Qualitative data revealed that self and peer assessment created positive and negative washback effects on learning, yet more positive washback effects resulted. Self and peer assessment helped meet the writing genre requirements, aided in bridging writing gaps through reading, supported editing through self-assessing, and benefited from expanding the lexicon reservoir through peer assessment, thereby helping the student become more creative through peer assessment. It also helped in building metacognitive awareness, yet learners faced difficulty to objectively self-assess writing. As for the washback effects of considering or not considering self-assessment results in peer assessment, data showed that considering self-assessment results

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

made peers reinforce a peer's self-critique and expand on that critique, but it restricted a peer's decisions. On the other hand, not considering self-assessment results led peers to make independent decisions and have a fresh perspective.

Washback of the assessment process was an emerging theme from the interview data; the process implemented in the study resulted in three positive washback effects: getting multiple feedback perspectives, getting familiarized with own writing gaps, making peer assessment easier.

Chapter 5: Discussion

Overview

This chapter aims to interpret and discuss the findings of the three research questions. Those questions sought to look into the washback effects of self and peer assessment in Arabic as a second language writing/composition. As demonstrated in chapter 4, the study revealed the learners' attitudes and perceptions--their feelings, their perceived capability to self and peer assesses, their perceived importance of self and peer assessments, and their perceived effectiveness of considering and not considering the peer's self-assessment results in peer assessment. Additionally, washback effects on revision and the washback effects of considering and not considering the peer's self-assessment results were investigated. The following sections of this chapter seek to integrate pieces of quantitative data with their corresponding pieces of qualitative data to explain and interpret finding more comprehensively. The chapter starts with addressing the pre-and post-attitudes of learners; then, it discusses the learners' perceived importance of the assessments and the washback effects of the assessments on revising writing. Finally, the chapter addresses the learner's perceptions about the idea of considering a peer's self-assessment in peer assessment and the resulting washback effects of this idea on learning.

RQ1: What are the participants' pre- and post-attitudes towards self and peer assessment?

Beliefs and perceptions are related to attitudes (Nkosana (2009), and, therefore, the study considered revealing the learners' attitudes. Those attitudes about self and peer assessment were mostly positive; quantitative data showed that learners' feelings about self and peer assessments and their perceived capability to self and peer assess were either natural or more positive after

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

taking the assessments. As for the qualitative data, it came out to strengthen the quantitative finding: building more confidence and feeling less nervous. Qualitative data even showed an emerging washback effect: feeling entertained to self and peer assess. It is worth mentioning that learners' feeling of entertainment was probably a result of them comparing their experience taking self and peer assessment in writing versus their previous experiences in writing with no self and peer assessment.

Learners' attitudes towards the assessments are vital in exploring how learners are doing with the assessment, and having positive attitudes is a good sign towards learning. Dong (2020) investigated the link between the participants' perceptions and learning, and the conclusion was that positive perceptions and attitudes were signs of positive effects on learning. Therefore, the attitude gains achieved in the post questionnaires and revealed in the qualitative data may indicate that self and peer assessments were working positively towards learning.

When they were asked about how excited they were taking peer assessment, their attitudes did not change; the pre-and post-mean score remained the same ($M=3.30$ on a scale of 1 to 5). Similarly, no attitude change took place when learners were asked how confident they felt before and after peer assessment ($M=3.35$). This may be because learners had previous experiences with peer assessment, as indicated in the demographic information in chapter 3. On the other hand, the attitudes changed on self-assessment variables; for example, on the variables *feeling confident to self-assess*, learners' attitudes went from a mean score of (2.90) before the self-assessment to a mean score (3.30) after the self-assessment. Thus, the introduction of self-assessment provided a confidence-raising experience for learners to self-assess their writing. Moreover, it is worth mentioning that assessing writing critically was a feeling that was developed through the experience; the feeling of being capable of critically self and peer assess

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

writing grew up from pre-assessment experience to the post-assessment experience. This finding resonates with those reported by Hung (2012), where Hung found that introducing the alternative assessment-digital writing profile project- resulted in a positive washback effect that learners enhanced their critical learning skills.

As reported in the interviews, another contributing factor to positive attitudes was the good design of the assessments. Learners in the interviews reported that the assessment forms walked them through what they needed to address during self and peer assessments. The assessment forms were analytically designed to address the different components of the writing task learners needed to consider in their writing. The fundamental component that made the assessment forms effective was that the forms mirrored the teacher's rubric. So, as the teacher had shared the rubric with learners, the assessment forms worked accordingly. This finding was in line with what Lee (2017) recommended: peer feedback can be best implemented when it is guided by a clear peer and assessment form in which students are kept to focus on the requirements of the assessment.

Feeling capable of rating a peer's writing accurately was one thing that quantitative data showed to have increased after taking peer assessment; however, some learners expressed that it was hard for them to rate the peer's writing to which the peer's self-assessment results were attached. It seems that disputing self-assessment results was a challenge. However, learners primarily developed positive attitudes towards self and peer assessment. Qualitative data also confirmed those positive attitudes: building more confidence and feeling less nervous. Qualitative data even showed an emerging washback effect: feeling entertained taking self and peer assessment.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The investigation of feelings and perceptions was used to reveal the attitudes; revealing the attitudes is necessary to reveal the washback back since learners' beliefs are essential in the validity of the washback effect (Green 2007;2013). After revealing learners' attitudes towards the assessments, the next step was investigating the washback effects of the assessment on learning. This part of the study is addressed in question 2:

RQ 2: The importance of self and peer assessment in revising second language writing:

a: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing?

b: Are there any washback back effects of self and peer assessment on the participants to revise their writing?

For the perceived importance of the assessments, learners perceived self and peer assessments as important. The quantitative data showed that learners valued using self and peer assessment in their writing. Unsurprisingly, all Means values showed a positive gain in the post questionnaire; yet one mean score on the variable *Self-assessment is necessary to revise my story writing* dropped from 4.15 to 3.80 (on a scale of 1 to 5). This finding contradicts another finding in the same questionnaire where learners were asked this: *If I have the choice not to self-assess my story, I would like to self-assess it*; the mean score on this variable dropped from 3.00 to 2.45, which means that learners viewed self-assessment as an important assessment tool with peer assessment. Even though the drop from 4.15 to 3.80 on the variable *Self-assessment is necessary to revise my story writing* is a negative sign, the post mean score (3.80) still means that the majority viewed self-assessment as important. The drop of 0.35 could mean something wrong was going on with the self-assessment. To reveal the possible reasons behind this drop, it is

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

inevitable to uncover the washback effects of the assessment on learners. Unsurprisingly, qualitative data demonstrated several washback effects (to be discussed in the following paragraphs); yet, one negative washback effect turned out to be caused by self-assessment. That negative washback effect was the difficulty in being objective in self-assessing one's writing; this finding can logically be associated with the statistical drop of 0.35 in the post Mean score of the variable *self-assessment is necessary to revise my story writing*. Qualitative data showed that it was difficult for some participants to be objective in self-assessing their own writing; consequently, this finding can be related to the drop in the Mean score on the variable *self-assessment is necessary to revise my story writing*.

The washback of self and peer assessments in this study was evident in the qualitative data.

Aldersons and Wall (1993) suggested in their Washback 15 hypothesis the following:

Tests that have important consequences will have washback effects,

Tests that do not have important consequences will have no washback effect.

These two hypotheses were used to show the relationship between high stakes tests and their influences on teaching, learning, and stakeholders; yet, in this study, the two hypotheses are adapted to show the relationship of the assessments (self and peer) and their washback effects on learners and learning. So, in terms of the assessment's influence on learners, self and peer assessments did affect the learners' perceptions; the post questionnaire, discussed in the above paragraph, showed that the perceptions were changed positively. Additionally, washback effects existed as well.

The investigation of washback effects of the assessments on revision showed that self and peer assessments had some common learning washback effects and some separate washback effects. That is, self and peer assessments were found to cause similar effects on learning and

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

found to have different effects as well. Hence, this finding seems to partially resonate with what Black et al. (2004) stated that self and peer-assessment are complementary assessment tools as they positively affect different areas of learning. This may be due to the assessment design as well. Both assessments were designed to address the exact writing task requirements, and the building components of the two assessments were somehow similar. An example of the assessments' common learning areas was *helping to meet the writing genre requirements*; this washback effect resulted from the two assessments because the two assessments prompted learners to target the story writing genre requirements. Another common washback effect was *learners' building of met-cognitive awareness* of writing- strengths, and weaknesses--which seemingly was caused by the assessment forms' informative, analytical design. This finding supports what Bardine & Fulton (2008) stated: self and peer assessment lead to building metacognitive awareness about learning. Another washback effect commonly caused by self and peer assessment was *bridging writing gaps* resulting from the extensive reading learners had to do to self and peer-assess writing. However, self-assessment was found to help with editing grammatical errors; the logical explanation for this is that catching one's own grammatical gaps seemed to be easier than objectively criticizing one's own ideas. This can be argued to be true if the fact that learners found it difficult to objectively self-assess writing. (See the discussion of the first question). Peer assessment was found to help with *expanding a vocabulary mental dictionary and becoming more creative*; these two washback effects were more likely created through the exposure to read other's ideas, and this finding can definitely be related to the collaborating writing theory as Hu & Lam (2010) stated that peer assessment provides collaboration opportunities in which learners support each other.

R3: The effectiveness of considering and not considering the peer's self-assessment results in peer assessment:

a: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

b: Are there any washback effects of considering and not considering peer's self-assessment results in peer assessment? If so, what are they?

The third question of this study was to look into the learners' perceptions about considering peer's self-assessment results in peer assessment and the potential washback effects that can be generated when considering and not considering the peer's self-assessment results. As elaborated in chapter 3, learners had to peer-assess two classmates' papers, one with the consideration of the peer's writer's self-assessment results and another classmate's paper that had no self-assessment results.

The examination of the qualitative data showed contradictory findings. The consideration of peer's self-assessment results was not favored from the perspective of peer assessors, whereas it was favored from the standpoint of the peer writers. The qualitative data revealed that the washback effects of considering self-assessment results were experienced differently through two perspectives. When learners had to peer-assess while considering the peer's self-assessment results, they had to comply with the results of the self-assessment. Subsequently, most learners had to reinforce and expand on the peer's self-critiqued writing areas while also providing suggestions and solutions. The peer assessors felt restricted; therefore, they did not favor the idea.

On the other hand, when learners received back their papers that had been peer-assessed with the consideration of their self-assessment results, learners found it helpful that their peers

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

considered their comments and their "writing gaps"; they may have felt that peer assessment had been personal, and hence, this created a positive attitude towards the idea of considering self-assessment results. As for the washback effects of not considering self-assessment results in peer assessment, qualitative data showed that peer assessors experienced no restrictions in expressing their own viewpoints during peer assessment. They reported that their decisions were independent, and the experience felt like a fresh start.

So, the two experiences demonstrated that every learner has peer-assessed two classmates' papers, one with and another without self-assessment results. This means that every learner had her/his paper peer-assessed twice, one with, and another without self-assessment results

As detailed in chapter three, the whole process of self and peer assessment was that every learner would initially self-assess their first writing draft using the self-assessment form. Then, they have that first draft peer-assessed twice by two different peers—one with the consideration of self-assessment results and another without self-assessment. One of the washback effects of this process was that learners received three feedback perspectives on the same paper. Those three perspectives can be viewed as self-feedback, peer feedback inspired by self-feedback, and independent peer feedback. It is clear what self-feedback is, but why are there two different peers' feedback perspectives? Why are they different? As explained in the findings of question two of this study that self-assessment was found to affect peer's decisions; the same thing is to be concluded here: self-assessment forced peers, in one way or another, to produce feedback that mirrors self-feedback. Then, the independent peer feedback perspective was a result of one peer assessing another peer's writing without any pressure from the first peer's self-assessment results.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

The question is whether this washback effect on learning was a positive one or a negative one. According to learners' reports in the interview, this process positively affected their writing. Taking self-assessment first was found to facilitate the peer assessment -the one that would consider the results of self-assessment results. This particular finding seems to be aligned with what Lee (2017) suggested. Lee stated that when self-assessment is preceded, learners have the chance to reflect on their writing, diagnose the gaps and strengths in their writings, then take those gaps to peer assessment to focus on, and so the assessment is now more purposeful and personal (p.97). In addition to this effect, this study revealed--due to its design--that the lack of consideration of peer' self-assessment, in one step of this process, created a fresh peer perspective too.

To conclude, the study adapted 4 of the 15-hypotheses suggested by Alderson and Wall (1993). The study attempted to show that alternative assessments, namely, self and peer assessments, can have washback effects. Based on the findings, it is a valid move adopting 4 of the 15 hypotheses of Wall and Alderson (1993) Alderson & Wall (1993):

A test will influence learning.

A test will influence how learners learn.

A test will influence the degree and depth of learning.

Tests that have important consequences will have washback.

(Alderson and Wall 1993, pp.120-121)

These four hypotheses were proven in this study: Self and peer assessments influenced learning; learners expressed their opinions about how useful and un-useful the assessments were. Self and peer assessments influenced how learners learn; they made learners approach writing a certain way throughout the assessment process. Self and peer assessments influenced the degree

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

and depth of learning; they made learners go deep into the assessment forms' analytical design.

Finally, self and peer assessment had important consequences, so they had a washback.

Chapter 6: Conclusion

Overview

Second language assessment and testing have shown that washback research has grown over several decades in its body of literature. Washback research has mainly focused on the effects on large-scale standardized tests. However, the body of literature has not shown enough washback studies dealing with classroom-based alternative assessments; additionally, the literature has not contained, so far, a single washback study on self and peer assessment in second language writing. Thus, this study bridges the gap by investigating whether washback effects can exist in alternative assessments, namely, self and peer assessment in second language writing. Having that said, the study, specifically, examined washback effects on Arabic language learners writing in Arabic.

After re-introducing the purpose and the methodology of the study, this chapter highlights the findings' conclusions and the significance of those findings in the field of teaching second language writing. Additionally, the chapter presents the limitations that challenged the study and some recommendations for similar future studies.

Purpose

Revealing what learners think of my assessments is one of my reliable means to validate my assessment tools. As a second language teacher who likes teaching writing, I find that self and peer assessments are inevitable teaching means. I pay much attention to designing, adopting, and adapting self and assessment materials; yet, I have always questioned what my students think about them. After some research investigating the washback effect and how washback can

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

reveal how well assessment works if learners are given the freedom to express their opinions, I had the idea of conducting a washback study on self and peer assessments in second language writing. The central questions of my study are these:

1. What are the participants' pre-and post-attitudes towards self and peer assessment?

2. The importance of self and peer assessment in revising second language writing:

A: What are the participants' pre- and post-perceived importance of self and peer assessment to revise their writing:

B: Are there any washback effects of self and peer assessment on the participants to revise their writing? If so, what are they?

3. The effectiveness of considering and not considering the peer's self-assessment results in peer assessment:

A: What are the participants' pre- and post-perceptions about considering peer's self-assessment results in peer assessment?

B: Are there any washback effects of considering and not considering peer's self-assessment results in peer assessment? If so, what are they?

Methodology

The nature of the study entailed using a mixed method. Revealing the attitudes and the perceptions about the assessments were statistically demonstrated, and through the quantitative method, I was able to justify some of those statistics and be able to reveal the washback effects of the assessments. Additionally, the qualitative method created an emergence of other themes. Interviews made me dig inside the learner's mind to elicit more in-depth information about their opinions and perceptions, and experiences taking self and peer assessment.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Quantitatively, I administered a 19-item questionnaire (pre and post) that targeted different areas of the learner's attitudes and perceptions about the assessments (feelings, perceived ability to take the assessments, perceived importance of the assessments, and perceived effectiveness to consider peer's self-assessment in peer assessment). Then, qualitatively, I utilized open-ended questions in the post questionnaire as well as interviewing ten learners. The interview participants were selected based on their contribution to the assessment: I chose the best ten contributors among learners to get richer data from the interviews to reveal the washback effects of the assessments on various learning areas. Each interview session lasted around 10-13 minutes, and I reached out to the participants via zoom application.

The participants were high school students who study Arabic as a second language. Their Arabic level was advanced, and so the writing task which I integrated for the study was at the advanced level. The writing task was an assigned writing task in their writing course (writing story). After they wrote the first draft of the story, I administered the pre-questionnaire, and then I administered the assessments. They were then allowed to integrate the assessment results into their second draft, and finally, I interviewed the participants. Afterward, interview data were transcribed and analyzed through the Vivo application.

Findings

The findings of this study can be shown into two parts: learners' attitudes and perceptions and washback on different areas of learning.

The pre-and post-questionnaires revealed that the attitudes and perceptions were primarily positive. The learners' feelings became more positive about the assessments after taking the assessments, and learners felt more excited, more confident, and less embarrassed, as shown in

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

the post-questionnaire. Qualitative data confirmed those feelings and also revealed that learners even felt entertained. As for the learner's perceived capability to self and peer assess, data shows, in the post questionnaire, that they were more critical in self and peer assessment than they had thought, yet they were less accurate in peer rating and neutral in self-rating than they had had thought before taking the assessments. Also, qualitative data showed contradictory findings of the ability to self-assess writing. Some expressed their incapability to be objective in their self-critique, while others said otherwise.

Post questionnaire variables on the perceived importance of self and peer assessment showed that most learners viewed that self and peer assessment were important, yet the data showed a contradictory finding for the variable: Self-assessment is necessary to revise my story writing. The mean score dropped from 4.15 down to 3.80 on a scale of 1-5 to tell that learners did not value self-assessment. On the other hand, on another related variable, If I have the chance choice not to self-assess my story, I would not self-assess it; the mean score dropped from 3:00 down to 2:45 to show that learners valued self-assessment.

Qualitative data examination revealed that self and peer assessments generated similar washback effects. The similar effects were these: helping to meet the writing genre requirements, bridging writing gaps through reading, building met-cognitive awareness. The qualitative data also showed that self and peer assessment had different washback effects: the washback effects that differed were these: editing through self-assessment, expanding lexicon reservoir through peer assessment, and becoming more creative through peer assessment. These washback effects were all positive, and the only negative washback effect that emerged was the difficulty to objectively self-assess writing.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Finally, the perceived effectiveness of considering a peer's self-assessment results in peer assessment was found in the post questionnaire to be more helpful from the peer-assessed perspective but not helpful in the peer assessor's perspective. The qualitative data from interviews explained how that happened. The consideration of the peer's self-assessment results in peer assessment reinforced and expanded on the peer's self-critique, but it restricted the peer evaluator's decisions; thus, the peer evaluator viewed peer' self-assessment results as "restricting." However, the peer being assessed found the assessment to be helpful. On the other hand, the absence of the peer's self-assessment results made the peer's decision more independent and fresher; thus, both the peer assessor and the peer-assessed found it helpful. This whole process had the following effects on learning: getting multiple feedback perspectives, getting more familiarized with own writing gaps, making peer assessment easier.

Significance

The study offers another proof, like the one presented in Hung (2012), that washback effects can exist in alternative assessments just like they exist in large scale-standardized tests. Hence, the study highlights the importance of investigating the washback effects of alternative assessments. Learners' opinions about the assessments they have to take should never be neglected. Therefore, digging into the washback effects of the assessments will give a chance for learners to express their opinions and reveal their experiences with the assessments. This will grant the teacher a comprehensive overview of the mechanism of the assessment, and subsequently, the teacher can determine the strong and the weak areas in the assessment. The process of investigating the washback effects of the assessments will make the learners feel heard and considered, which boosts their motivation for learning.

Limitations

Despite the attempt to triangulate the data, revealing the washback effects of using both self and peer assessment was limited by my incapability to have an actual look into the learners' first and final drafts to see how the assessments affected the learners' writing performance from the first draft to the final draft. This would have allowed me to compare drafts. Digging into learners' papers would have also given me a magnifying glass to detect how learners peer-assessed with the consideration of the writer's self-assessment results versus without the writer's self-assessment results. Looking at the student's papers would have most likely confirmed or disputed some of the findings that learners claimed.

Another limitation of this study was that I did not consider the teacher's perspective on how the assessment affected learners; the teacher's input would have allowed me to reveal other hidden effects. Since the teacher is usually familiar with the learner's progress, it would have been a good idea to talk to the teacher and ask questions about the learners' progress, their feedback about the assessments, the process, and the design, etc. Additionally, considering the teacher's perspective can be informative about the long-term washback effects on learning, I could have interviewed the teacher a few weeks later to ask about whether learners were still motivated to integrate self and peer assessments in any writing activities.

Recommendations for Future Research

Given the limitations expressed in the above section, I recommend collecting the learner's first and final draft; this will unquestionably strengthen the findings. Having a deep look at the learners' performance would clarify and justify some of the quantitative findings. Sometimes, some statistical findings cannot be easily verified through interviews; this happened to me when

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

a student said she needs to check the paper to remember what type of feedback she received from her peer. The other recommendation is to interview the teacher as this would allow for richer data about the learner's feedback about the assessments. Another recommendation is to give enough time for learners to familiarize themselves with assessment forms. I asked the teacher to let the learners have a look at the design of the format and give them more time to ask questions about the content. A final recommendation is making sure that the assessment forms are aligned with the teacher's rubric and that learners are given a chance to pilot it.

Ending

Tending to prefer teaching second language writing, I conducted this study with enthusiasm for the field with hopes of reaching the end with a conclusion that I can make my writing students more involved and heard. As the study was all about what the learners thought about what they experienced in taking self and peer assessment, I validated the effectiveness of the assessment materials by considering what the learners reported and expressed. I have formed a conviction that being communicative in teaching writing is not merely following the writing process parameters but also making the process valid and reliable. This is what I tried to achieve during my experience of teaching in general and writing in particular.

Having had the belief that washback effects do not only exist in large-scale standardized tests but also in classroom alternative assessments, I began the study confidently by building on the famous similar study of Hung (2012) in which he concluded that the washback effect strongly existed in E-portfolio writing projects. I followed the same methodology protocols known to be used with washback studies that dealt with large-scale standardized tests, and my conclusion was

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

that washback effects do exist in alternative assessments. This study made me now even more considerate of my learners, who will be more actively involved in all the learning decisions.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

References

- Abbasian Boroojeni, F., Tavakoli, M., & Vahid Dastjerdi, H. (2017). Washback effect of general English test of Ph. D. entrance exam on science and humanities students: Perceptions and practices. *Research in Applied Linguistics*, 8(1), 109-132.
- Alderson, J. C. 2004. 'Foreword' in L. Cheng, Y. Watanabe, and A. Curtis (eds.). *Washback in language testing: Research contexts and methods*. London: Lawrence Erlbaum.
- Alderson, J. C, & Banerjee, J. (2001). Language testing and assessment (Part1). *Language teaching*, 34, 213-236
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115-129.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A case study. *Language Testing*, 13, 280-297.
- Andrade, H. G., and B. A. Boulay. 2003. "Role of Rubric-referenced Self-Assessment in Learning to Write." *The Journal of Educational Research* 97 (1): 21–30. doi:10.1080/00220670309596625.
- Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & evaluation in higher education*, 32(2), 159-181
- Allen, D. (2016 a). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language testing in Asia*, 6(1), 7.
- Allen, D. (2016 b). Japanese cram schools and entrance exam washback. *The Asian Journal of Applied Linguistics*, 3(1), 54-67.
- Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. *International Journal of Qualitative Methods*, 18, 1609406919874596.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Ashraf, H., & Mahdinezhad, M. (2015). The Role of Peer-assessment versus Self-assessment in Promoting Autonomy in Language Use: A Case of EFL Learners. *Iranian Journal of Language Testing*, 5(2), 110-120.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- BAKER, E. L. (1991). Alternative assessment and national policy. Paper presented at the National Research Symposium on Limited English Proficient Students' Issues: Focus on Evaluation and Measurement, Washington, D.
- Baksh, A., Mohd Sallehuddin, A. A., Tayeb, Y. A., & Norhaslinda, H. (2016). Washback Effect of School-based English Language Assessment: A Case-Study on Students' Perceptions. *Pertanika Journal of Social Sciences & Humanities*, 24(3).
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13. 257-279.
- Bardine, F. & Fulton, P. (2008). Analyzing the benefits of revision memos during the writing and revision process. *Clearing House*, 81(4), 149–154. doi: 10.3200/TCHS.81.4.149–154
- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, 86(1), 8-21.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL quarterly*, 32(4), 653-675
- Brown, H. D. (2001). *Teaching by Principles: An interactive approach to language pedagogy* (2nded.). San Francisco State University: Longman.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. Psychology Press.
- Brown, S., Rust, C., & Gibbs, G. (1994). *Involving students in the assessment process. Strategies for diversifying assessments in Higher Education*.
- Bing, X. U. (2016). A Study of the Effects of Student Self-assessment on the EFL Writing of Chinese College Students. *Journalism and Mass Communication*, 6(2), 91-107.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, 86(1), 8-21.
- Burkšaitienė, N., & Teresevičienė, M. (2008). Integrating alternative learning and assessment in a course of English for law students. *Assessment & Evaluation in Higher Education*, 33(2), 155-166.
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In *Washback in language testing* (pp. 25-40). Routledge.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-121. *Studies in language testing 21*. Cambridge: Cambridge University Press.
- Caulk, N. (1994). Comparing teacher and student responses to written work. *Tesol Quarterly*, 28(1), 181-188.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education*. Education.
https://doi.org/10.1111/j.1467-8527.2007.00388_4.x
- Collimore, L. M., Paré, D. E., & Joordens, S. (2015). SWDYT: So What Do You Think? Canadian students' attitudes about peerScholar, an online peer-assessment tool. *Learning Environments Research*, 18(1), 33-45.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Damankesh, M., & Babaii, E. (2015). The washback effect of Iranian high school final examinations on students' test-taking and test-preparation strategies. *Studies in Educational Evaluation*, 45, 62–69. <https://doi.org/10.1016/j.stueduc.2015.03.009>
- Dyer, B. (1996). L1 and L2 composition theories: Hillocks' environmental mode and task-based language teaching. *ELT Journal*, 50(4), 312-317.
- Diltz, J. (2006). Words to voice: Three approaches for student self-evaluation. *Teaching English in the Two-Year College*, 34(1), 41–45.
- Dong, M. (2020). Structural relationship between learners' perceptions of a test, learning practices, and learning outcomes: A study on the washback mechanism of a high-stakes test. *Studies in Educational Evaluation*, 64, 100824.
- Eckstein, M. A., & Noah, H. J. (1993). A comparative study of secondary school examinations. University of London, International Centre for Research on Assessment.
- Elbow, P. (1994). About voice and writing. In P. Elbow (Ed.), *Landmark essays on voice and writing*. (pp. xi– xlvii). Mahwah, NJ: Hermagoras Press
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta- analysis. *Review of Educational Research*, 59(4), 395-430.
- Fitzgerald, J., & Markham, L. R. (1987). Teaching children about revision in writing. *Cognition and instruction*, 4(1), 3-24.
- Ferris, D. R., & Hedgcock, J. (2013). *Teaching L2 composition: Purpose, process, and practice*. Routledge.
- FREDERIKSEN, J.R. and COLLINS, A. (1989). A system approach to educational testing. *Educational Researcher*, 18, 9, 27-32.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Freeman, M. (1995). Peer assessment by groups of group work. *Assessment & Evaluation in Higher Education*, 20(3), 289-300.
- Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment & Evaluation in Higher Education*, 24(4), 365-377.
- Goodrich, H. (1996). Student self-assessment: At the intersection of metacognition and authentic assessment (Doctoral dissertation, Harvard Graduate School of Education).
- Gregory, K., Cameron, C., & Davies, A. (2000). *Knowing what counts: Self-assessment and goal setting*. Merville, Canada: Connections Publishing.
- Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39-51.
- Green, A. (2007 a). *IELTS washback in context: Preparation for academic writing in higher education* (Vol. 25). Cambridge University Press
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Hamp-Lyons, L. 1997. 'Washback, impact and validity: ethical concerns. *Language Testing* 14/3: 295–303
- Hanrahan, S.J. & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research and Development*, 20(1), 53–73. doi: 10.1080/07924360120043658
- Hoa, D. T. P. (2020). Washback of English Proficiency Test in Classroom Activities at National University of Arts Education. *VNU Journal of Science: Education Research*, 36(1).

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Honsa, S. (2013). Self-assessment in EFL writing: A study of intermediate EFL students at a Thai university. *Voices in Asia Journal*, 1(1), 34-57.
- Hu, G., & Lam, S. T. E. (2010). Issues of cultural appropriateness and pedagogical efficacy: Exploring peer review in a second language writing class. *Instructional science*, 38(4), 371-394.
- Hughes, A. (1988). Introducing a needsbased test of English language proficiency into an English medium university in Turkey in A.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press
- Hughes, I.E., & Large, B.J. (1993). Staff- and peer-group assessment of oral communication
- Hughes, A. (1993). Washback and TOEFL 2000. Unpublished manuscript, University of Reading Educational Testing Service (ETS). University of Reading.
- Hughes, I. E., & Large, B. J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, 18(3), 379-385.
- Hung, S. T. A. (2012). A washback study on e- portfolio assessment in an English as a Foreign Language teacher preparation program. *Computer Assisted Language Learning*, 25(1), 21-36.
- Iraji, H. R., Enayat, M. J., & Momeni, M. (2016). The effects of self-and peer-assessment on Iranian EFL learners' argumentative writing performance. *Theory and Practice in Language Studies*, 6(4), 716-722.
- James, R., McInnis, C., & Devlin, M. (2002). *Assessing learning in Australian universities*. Melbourne: The University of Melbourne Centre for the Study of Higher Education.
- Jhangiani, R. S. (2016). The impact of participating in a peer assessment activity on subsequent academic performance. *Teaching of Psychology*, 43(3), 180-186.
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, 39(3), 387-406.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Keeves, J., & Sowden, S. (1992). Analyzing qualitative data. *Methodology and Measurement in International Educational Surveys*
- Kennedy, S., & Lui, R. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concordia Working Papers in Applied Linguistics*, 4(10), 23-29.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.
- Lam, H.P. (1994). Methodology washback and insider's view. In D. Nunan, R. Berry, & V. Berry (Eds.), *Bringing about change in education* (pp. 83-99). Hong Kong: University of Hong Kong.
- Larson, R.L. (2000). Revision as self-assessment. In J.B. Smith & K.B. Yancey (Eds.), *Student self-assessment and development in writing: A collaborative inquiry*. (pp. 97-104). Cresskill, NJ: Hampton Press
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer Singapore.
- Lewthwaite, M. (2007). Teacher and student attitudes to IELTS writing tasks: positive or negative washback. *UGRU Journal*, 5, 1-16.
- Lin, S. S., Liu, E. Z., & Yuan, S. M. (2002). Student attitudes toward networked peer assessment: Case studies of undergraduate students and senior high school students. *International Journal of Instructional Media*, 29(2), 241-255.
- Liu, J., & Hansen, J. (2002). *Peer response in second language writing classrooms*. Ann Arbor: University of Michigan Press.
- Liu, E. Z. F., & Yuan, S. M. (2003). A study of students' attitudes toward and desired system requirements of networked peer assessment system. *International Journal of Instructional Media*, 30(4), 349-355.
- Liu, N. F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher education*, 11(3), 279-290.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Lu, R., & Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning*, 6(2).
- Luxia, Q. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language testing*, 22(2), 142-173.
- Luoma, S., & Tarnanen, M. (2003). Creating a self-rating instrument for second language writing: From idea to implementation. *Language Testing*, 20, 440- 465.
- MADAUS, G. F. (1988). The influence of testing on the curriculum. In: Tanner, L. N. (Ed.). *Critical issues in curriculum: Eighty-seventh yearbook of the National Society for the Study of Education*. Chicago: University of Chicago Press, pp. 83-121.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 075-100
- McGarr, O., & Clifford, A. M. (2013). 'Just enough to make you take it seriously': exploring students' attitudes towards peer assessment. *Higher education*, 65(6), 677-693.
- McCarthy, P., Meier, S., & Rinderer, R. (1985). Self-efficacy and writing. *College Composition and Communication*, 36, 46
- McGarr, O., & Clifford, A. M. (2013). 'Just enough to make you take it seriously': exploring students' attitudes towards peer assessment. *Higher education*, 65(6), 677-693.
- 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.
- Nimehchisalem, V., Chye, D. Y. S., & Jaswant Singh, S. K. A. (2014). A Self-Assessment Checklist for Undergraduate Students' Argumentative Writing. *Advances in Language and Literary Studies*, 5(1), 65-80.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- O'Malley, J. M., and Valdez Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. New York: Addison Wesley.
- Oscarson, A. D. (2009). *Self-Assessment of Writing in Learning English as a Foreign Language: A Study at the Upper Secondary School Level*. Goteborg Studies in Educational Sciences 277.
- Messick, S. (1996). Validity and washback in language testing. *Language testing*, 13(3), 241-256.
- Nelson, G. L., & Carson, J. G. (1998). ESL students' perceptions of effectiveness in peer response groups.
- MESSICK, S. (1989). Validity. In: LINN R. (Ed.) (3rd Ed.) *Educational Measurement* New York, NY: Macmillan, pp. 13-103. *Journal of second language writing*, 7(2), 113-131.
- Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of second language writing*, 15(2), 118-141.
- Montrul, S. (2011). Introduction: The linguistic competence of heritage speakers. *Studies in Second Language Acquisition*, 33(2), 155-161
- Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). Conducting semi-structured interviews. *Handbook of practical program evaluation*, 492.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2), 199-218.
- Nkosana, L. (2009). The washback effects of curriculum and assessment reform in Botswana. Marang: *Journal of Language and Literature*, 19(1), 59-79.
- Pan, Y. C., & Newfields, T. (2012). Tertiary EFL Proficiency Graduation Requirements in Taiwan: A Study of Washback on Learning. *Electronic Journal of Foreign Language Teaching*, 9(1).
- Pan, Y. C. (2008). A critical review of five language washback studies from 1995-2007: Methodological considerations. *JALT Testing & Evaluation SIG Newsletter*, 12(2), 2-16.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Pathak, A., & Intratat, C. (2016). Use of semi-structured interviews to investigate teacher perceptions of student collaboration. *Malaysian Journal of ELT Research*, 8(1), 10.
- Patchan, M. M., Schunn, C. D., & Clark, R. J. (2011). Writing in natural sciences: Understanding the effects of different types of reviewers on the writing process. *Journal of Writing Research*, 2(3), 365-393. doi:10.17239/jowr-2011.02.03.4
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational psychologist*, 36(2), 89-101.
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of second language writing*, 8(3), 265-289.
- Planas Lladó, A., Soley, L. F., Fraguell Sansbelló, R. M., Pujolras, G. A., Planella, J. P., Roura-Pascual, N., ... & Moreno, L. M. (2014). Student perceptions of peer assessment: an interdisciplinary study. *Assessment & Evaluation in Higher Education*, 39(5), 592-610.
- Pavlenko, A. (2008). 18 Narrative Analysis. *The Blackwell guide to research methods in bilingualism and multilingualism*, 311.
- POPHAM, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Praver, M., Rouault, G., & Eidswick, J. (2011). Attitudes and affect toward peer evaluation in EFL reading circles. *Reading*, 11(2), 89-101.
- Raes, A., Vanderhoven, E., & Schellens, T. (2015). Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Studies in Higher Education*, 40(1), 178-193.
- Raworth, K., Sweetman, C., Narayan, S., Rowlands, J., & Hopkins, A. (2012). *Conducting semi-structured Interviews*. Oxfam

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Roskams, T. (1999). Chinese EFL students' attitudes to peer feedback and peer assessment in an extended pairwork setting. *RELC Journal*, 30(1), 79-123.
- Romeo, L. (2008). Informal writing assessment linked to instruction: A continuous process for teachers, students, and parents. *Reading & Writing Quarterly*, 24(1), 25-51.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment.
- Safa, M. A. (2014). The Washback Effects of Task-based Assessment on the Iranian EFL Learners' Grammar Development. *Procedia-Social and Behavioral Sciences*, 98, 90-99.
- Saglam, A. L. G. (2018). Can exams change how and what teachers teach? investigating the washback effect of a university english language proficiency test in the turkish context. *Eurasian Journal of Applied Linguistics*, 4(2), 155-176.
- Schunn, C., Godley, A., & DeMartino, S. (2016). The reliability and validity of peer review of writing in high school AP English classes. *Journal of Adolescent & Adult Literacy*, 60(1), 13-23.
- Saville, N., & Hawkey, R. (2004). The IELTS impact study: Investigating washback on teaching materials. *Washback in language testing: Research contexts and methods*, 73-96.
- Shih, C. M. (2007). A new washback model of students' learning. *Canadian Modern Language Review*, 64, 135-162.
- Schendel, E., & O'Neill, P. (1999). Exploring the theories and consequences of self-assessment through ethical inquiry. *Assessing Writing*, 6(2), 199-227
- Shirzadi, D., & Amerian, M. (2020). Washback effects of multiple-choice, cloze and metalinguistic tests on EFL students writing. *Indonesian Journal of Applied Linguistics*, 9(3), 536-544.
- Singh, K., & Terry, J. (2008). Fostering students' self-assessment skills for sustainable learning.
- SMITH, M. L (1991). Meanings of test preparation. *American Educational research Journal*, 28, 521-42.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language testing*, 13(3), 298-317.
- Sommers, N. & Saltz, L. (2004). The novice as expert: Writing the freshman year. *College Composition and Communication*, 5(6), 124–149.
- Stewart, D. C. (1972). *The Authentic Voice: Pre-writing Approach to Student Writing*. C. Brown.
- Stiggins, R. J. (1987). Design and development of performance assessments. *Educational measurement: Issues and practice*, 6(3), 33-42.
- Sullivan, G. M., & Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*. <https://doi.org/10.4300/JGME-5-4-18>
- Tannenbaum, J. 1996. . ED395500, Washington, DC: ERIC Clearinghouse on Languages and Linguistics.
- Tsagari, D. (2009). *The complexity of test washback: An empirical study* (Vol. 15). Peter Lang.
- Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276.
- Tsagari, K. (2006). *Investigating the washback effect of a high-stakes EFL exam in the Greek context: Participants' perceptions, material design and classroom applications*. Lancaster University.
- Van Gennip, N. A., Segers, M. S., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280-290.
- Vasu, K., Nimehchisalem, V., Fung, Y. M., & Sabariah, M. R. (2018). The usefulness and effectiveness of argumentative writing self-assessment checklist in undergraduate writing classrooms. *International Journal of Academic Research in Business and Social Sciences*, 8(4), 202-219.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

- Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning?. *Assessment & Evaluation in Higher Education*, 34(2), 221-230.
- Villamil, O. S., & de Guerrero, M. C. M. (2006). Socio-cultural theory: A framework for understanding the socio-cognitive dimensions of peer feedback. In K. Hyland & F. Hy land(Eds.i Feedback in second language writing: Contexts and issues (pp. 23-42). New York: Cambrids University Press
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological proem*. Cambridge, MA: Harvard University Press.
- Wall, D. (1997). Impact and washback in language testing. *Encyclopedia of language and education*, 7, 291-302.
- Wall, D., & Alderson, J. C. (1993). Examining washback: the Sri Lankan impact study. *Language testing*, 10(1), 41-69.
- Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, 19, 80-96.
- Warne, B.M. (2008). Writing steps: A recursive and individual experience. *English Journal*, 97(5), 23–27.
- Wen, M. L., & Tsai, C. C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education*, 51(1), 27-44.
- Yancey, K. B. 1998. *Reflection in the writing classroom*. Logan, Utah: Utah State University Pr
- Yancey, K.B. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication*, 50(3), 483–501
- Zou, Y., Schunn, C. D., Wang, Y., & Zhang, F. (2018). Student attitudes that predict participation in peer assessment. *Assessment & Evaluation in Higher Education*, 43(5), 800-811.

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Appendices

Appendix A: Peer assessment form

| | 4: Excellent | 3: Good | 2: Average | 1: Needs improvement | Score |
|---|--|--|---|--|-------|
| Exposition (Setting) | | | | | |
| Your peer was able to | write many vivid descriptive words to tell the audience when and where the story took place | writes some vivid descriptive words to tell the audience when and where the story took place. | write only little descriptions about when and where the story took place. | Mention, where and when the stories took place, but no further description is included. | |
| Exposition (Characters) | | | | | |
| Your peer was able to | describe the characters and clearly describe them. Readers could describe the characters accurately. | describe most of the characters. Readers would have some idea of what the characters looked like. | briefly describe characters so that the reader knows a little about the characters. | Describe only some characters Other about the characters are left out. Also, it is hard to tell who the main characters are. | |
| Conflict (Problem) | | | | | |
| Your peer was able to | to make it very easy for the reader to understand the problem the main characters face, and why it is a problem. | make it fairly easy for the reader to understand the problem the main characters face and why it is a problem. | make it easy for the reader to understand the problem the main characters face, but it is not clear why it is a problem, | somehow show that the story has a problem, but the problem is hard to and confusing to follow through the story. | |
| Conflict (Rising action and Climax) | | | | | |
| Your peer was able to | make the story very interesting that shows a developed rising suspense and climax of the conflict. | make the story somehow interesting that shows a developed rising suspense and climax. | make only some parts in the story interesting, but the rising suspense and climax need to be developed in some scenes. | only show few interesting scenes, but not able to show good rising suspense and climax | |
| Resolution (Falling action) | | | | | |
| Your peer was able to | make events leading to the end of the story suspensefully sequenced, and well developed. | make events leading to the end of the story well developed. | make events leading to the end of the story developed in some parts, not in other parts | make events leading to the end somehow developed, but they are not well sequenced in many places | |
| Resolution (End of the story) | | | | | |
| Your peer was able to | make the ending very surprising and suspenseful | make the ending surprising | make the ending good, but typical. | make an ending, but it is not quite good. | |
| Writing Mechanics (Content organization) | | | | | |
| Your peer was able to | make the story well organized; one idea or scene follows another in a logical sequence with clear transitions. Also, the use of pictures is organized and purposeful to the context. | make the story pretty organized with good transitions. But some ideas and scenes seem out of place. Also, the use of pictures is somehow organized but your peer needs to develop this a little bit. | make the story organized in some parts, but hard to follow in other parts. The transitions are sometimes not clear. The use of pictures is ok, but it is not purposeful to the context in many cases. | make only few organized ideas, and the story line seems to be random. Your peer also missed to insert pictures. | |
| Writing Mechanics (content Clarity) | | | | | |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | | | | | |
|---|--|--|--|--|-------------------------|
| Your peer was able to | write clearly meaningful sentences and phrases. The story has no unclearly expressed sentences except for few un-disturbing cases | Managed to write clearly meaningful sentences and phrases, but there are few vague unclear sentences | write meaningful sentences and phrases, but there are many unclear sentences and phrases. | write few meaningful sentences and phrases, but there are many confusing unclear sentences and phrases that make some parts of the story difficult to understand. | |
| Mechanics (Vocabulary and structures) | | | | | |
| Your peer was able to | use all vocabularies suitably to describe the story plot and characters with very few spelling errors. She/he used some complex structures in my descriptions. | use most of vocabularies suitably to describe the story plot and characters with very few spelling errors. Your peer used some compound structures in my descriptions. | use vocabularies suitably in some cases to describe the story plot and characters, but there are cases where she/ he needs to develop them. The story contains some disturbing spelling errors. Also, most of the story descriptions are simple. | use vocabularies suitably in very few instances. Your descriptions of the story plot and the characters are not quite understandable in many spots in the story. The story contains so many simple grammatical structures. | |
| Writing Mechanics (Tenses) | | | | | |
| Your peer was able to | accurately use the past tense to narrate, and use the tenses correctly in dialogues between characters. | mostly, use the past tense correctly to narrate, and use the other tenses correctly in dialogue between characters. | in some instance, use of past tense correctly to narrate, and also correctly use the different tenses in dialogues, but there many instances where I failed to do so | correctly use past tense to narrative in a few cases. Also, the different tenses in dialogues between characters are used correctly in a few places in the story. | |
| | | | | | Total score/40 |
| Summative assessment | | | | | |
| In your viewpoint, what are the strengths of this story? | | | | | |
| In your viewpoint, what are the weaknesses of this story? | | | | | |
| Any other comments that your peer should consider: | | | | | |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Appendix B: Self assessment form

| Exposition (Setting) | Yes:4 | Yes:3 | Yes: 2 | Not sure: 1 | Score |
|--|---------------|---------------|---------------|--------------------|-------------------------|
| I wrote many vivid descriptive words to tell the audience when and where the story took place. | | | | | |
| Exposition (characters) | Yes: 4 | Yes: 3 | Yes: 2 | Not sure: 1 | |
| I elaborated on the characters, and clearly described them, so readers would have a solid idea of what they look like and their roles in the story | | | | | |
| Conflict Problem | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I made it very easy for the reader to understand the problem the main characters faced and why it is a problem. | | | | | |
| Conflict Rising action and Climax | Ye: 4 | Yes: 3 | Yes: 2 | Not sure: 1 | |
| I made the story very interesting at showing developed suspenseful rising actions to the climax. Also, I made the climax very engaging. | | | | | |
| Resolution Falling action | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I made the events leading to the end of the story suspensefully sequenced and developed. | | | | | |
| Resolution (End of the story) | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I made the ending surprising and suspenseful. | | | | | |
| Writing Mechanics (content organization) | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I made the story well organized; one idea or scene follows another in a logical sequence with clear transitions. Also, the use of pictures is organized and purposeful to the context. | | | | | |
| Writing Mechanics (content Clarity) | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I wrote clearly meaningful sentences and phrases. The story has no unclearly expressed sentences except for few un-disturbing cases | | | | | |
| Writing Mechanics (Vocabulary and structures) | Yes: 4 | Yes:3 | Yes: 2 | Not sure: 1 | |
| I suitably used accurate and varied vocabulary to describe the story plot and characters with only very few spelling errors. I used complex structures in my descriptions. | | | | | |
| Writing Mechanics (Tenses) | Yes: 4 | Yes:3 | Yes: 2 | Not sure:1 | |
| I accurately used the past tense to narrate the story and used the tenses correctly in dialogues between characters. | | | | | |
| | | | | | Total score/40 |
| General comments | | | | | |
| 1. What are your strengths in this assignment? Please critically analyze your strengths by presenting specific examples: | | | | | |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

2. What are your weaknesses in this assignment? Please critically analyze your weaknesses by presenting specific example:

Appendix C: Pre-Questionnaire

I would like to sincerely thank you in advance for completing this questionnaire. Please do not write your name anywhere on this questionnaire.

Part1: Demographic information

What is your gender?

Female/ Male

Do you speak Arabic in the household?

Yes/No

Do you write in Arabic outside the class (e.g., you text in Arabic)

Yes/No

Have you ever self-assessed your own writing?

Yes/No

Have you ever peer assessed your own writing?

Yes/No

Part2: Feelings about self and peer assessment

1. I am excited about self and peer assessment.

Strongly Disagree Disagree Neutral Agree Strongly Agree

2. I feel confident to self-assess my story

Strongly Disagree Disagree Neutral Agree Strongly Agree

3. I feel confident to peer -assess a classmate's story

Strongly Disagree Disagree Neutral Agree Strongly Agree

4. I think ratings resulting from peer assessment are not fair

Strongly Disagree Disagree Neutral Agree Strongly Agree

5. Having my peer assessing my story will embarrass me

Strongly Disagree Disagree Neutral Agree Strongly Agree

Part 3: Pre-perceptions of learners' incapability to self and peer assess

6. I do not feel I will be critical in assessing a classmate's writing.

Strongly Disagree Disagree Neutral Agree Strongly Agree

7. I do not feel I will be able to rate a classmate 's story writing accurately.

Strongly Disagree Disagree Neutral Agree Strongly Agree

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| |
|---|
| 8. I think I will not be critical in my self-assessment. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 9. I do not feel I will be able to rate my writing accurately. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| Part 4: Pre-perceived importance of Self and Peer Assessments |
| 10. Self and peer assessing writing are good ways to learn writing Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 11. Self-assessment is necessary to revise my writing Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 12. Peer -assessment is necessary to revise my writing. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 13. I think self and peer assessment will help me fix grammatical errors, and enhance my ideas Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 14. If I had the choice not to self-assess my story, I would not self-assess it. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 15. If I had the choice not to peer assess my classmate's story I would not peer assess it Strongly Disagree Disagree Neutral Agree Strongly Agree |
| Part 5: Pre- perceptions about the effectiveness of considering and not considering self-assessment in peer assessment |
| 16. As a peer assessor, having the results of a classmate's self-writing assessment is helpful to peer assess a classmate's writing. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 17. Having my peer consider my self-assessment results may lead to better peer assessment results in my writing Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 18. As a peer assessor, having the results of a classmate's self-writing assessment may restrict my peer assessment. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 19. As a peer assessor, I do not feel comfortable disputing a classmate's self-assessment results. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 20. If I had the choice not to take the self and peer assessments together, I would only take one of them to assess my story. Strongly Disagree Disagree Neutral Agree Strongly Agree |

Appendix C: Post-Questionnaire

| | | | | | | |
|---|--|-------------------|----------|---------|-------|----------------|
| <p>I would like to sincerely thank you in advance for completing this survey. As teachers learn more about how students learn, we can learn how to teach better. Please do not write your name anywhere on this questionnaire</p> | | | | | | |
| Part1: Post feelings about the self and peer assessment | | | | | | |
| 1. | I was excited about self and Peer assessment. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 2. | I felt confident to self-assess my story. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 3. | I felt confident to peer -assess a classmate’s essay. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 4. | I think ratings resulting from peer assessment were not fair. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 5. | Having had my peer assessing my story embarrassed me. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| Part 2: Post perceptions of learners’ incapability to self and peer assess | | | | | | |
| 6. | I think I was not critical in my writing self-assessment. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 7. | I do not feel I critically peer assessed my classmate's writing. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 8. | I do not feel I rated my writing accurately. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 9. | I do not feel I rated a classmate's writing accurately. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| Part3: Post -perceived importance of Self and Peer Assessments | | | | | | |
| 10. | Self and peer assessing writing were good ways to learn writing. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 11. | Self-assessment was necessary to revise my writing. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 12. | Peer-assessment was necessary to revise my writing. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 13. | I think self and peer assessments helped me fix my grammatical errors, and enhance my ideas. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| 14. | If I had had the choice not to self-assess my story, I would not have self-assessed it. | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| |
|---|
| 15. If I had had the choice not to peer assess my classmate's story, I would not have peer assessed it. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| Part4: Post-perceived effectiveness of considering not considering the results of peer self-assessment in peer assessment. |
| 16. As a peer assessor, having the results of a classmate's self-writing assessment helped me peer assess the classmate's writing. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 17. Having had my peer consider my self-assessment results helped my peer assess my writing. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 18. As a peer assessor, having had the results of a classmate's self-writing assessment restricted my peer assessment. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 19. As a peer assessor, I did not feel comfortable disputing a classmate's self-assessment results. Strongly Disagree Disagree Neutral Agree Strongly Agree |
| 20. If I have the choice not to take the self and peer assessments together, I will only take one of them to assess my story. Strongly Disagree Disagree Neutral Agree Strongly Agree |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

Appendix C: Story Grading Rubric

| | 4 | 3 | 2 | 1 |
|--------------------------|---|---|---|--|
| Exposition | | | | |
| Setting | Many vivid descriptive words are used to tell when and where the story took place. | Some vivid descriptive words are used to tell the audience when and where the story took place. | The reader can figure out when and where the story took place, but the author did not supply details. | The reader has trouble figuring out when and where the story took place. |
| Characters | The main characters are named and clearly described in the text as well as pictures. Most readers could describe the characters accurately. | The main characters are named and described. Most readers would have some idea of what the characters looked like | The main characters are named. The reader knows very little about the characters. | It is hard to tell who the main characters are. |
| Conflict | | | | |
| Problem | It is very easy for the reader to understand the problem the main characters face, and why it is a problem. | It is fairly easy for the reader to understand the problem the main characters face and why it is a problem. | It is fairly easy for the reader to understand the problem the main characters face, but it is not clear why it is a problem, | It is not fairly clear what the problem the main characters face. |
| Rising action and Climax | The story clearly and interestingly shows developed rising suspense and climax of the conflict | The story somehow shows developed rising suspense and climax | The story's rising suspense and climax are developed in some scenes, but not in other scenes. | The story does not show developed rising suspense and climax |
| Resolution | | | | |
| Falling action | Events leading to the end of the story are well developed. | Events leading to the end of the story are somehow | Events leading to the end of the story are developed in some parts, not in other parts | Events leading to the end of the story are not developed and confusing. |
| End of the story | The ending is surprising and suspenseful | The ending is somewhat surprising | The ending is unoriginal | The ending is missing |
| Writing mechanics | | | | |
| Organization | The story is well organized. One idea or scene follows another in a logical sequence with clear transitions. | The story is pretty organized. One idea or scene may seem out of place. Clear transitions are used. | The story is a little hard to follow. The transitions are sometimes not clear. | The Ideas and scenes seem to be randomly arranged. |
| (Tenses) | The author used the past tense to narrate accurately and used the tenses correctly in dialogues between characters. | The author mostly used the past tense correctly and mostly used the tenses correctly in dialogues between characters. | The use of past tense to narrate and the different tenses in dialogues are not correct in many instance | The use of past tense to narrative and different tenses in dialogues is mostly incorrect |
| Vocabulary and Spelling | All vocabulary used are suitable to | Most of vocabularies used are suitable to | Some of the vocabulary used are | A lot of vocabulary used to describe the |

THE WASHBACK EFFECT OF SELF AND PEER ASSESSMENT

| | | | | |
|--|---|--|--|---|
| | describe setting and Characters. No or very few spelling errors in the story. | describe setting and Characters. spelling errors in the story. | suitable to describe the characters and seating, but some are out of place. Many spellings errors. | setting and characters are not suitable, a lot of disturbing spelling errors. |
| | | | | Score = / 90 |

