The Practice of Cross-Grading in Assessing Writing: The Case of EFL Teachers and Students in a Saudi Arabian Context

Abdullah Alshakhi[1]

[1] English Language Institute, King Abdulaziz University, Jeddah, Saudi Arabia

Correspondence: Abdullah Alshakhi, English Language Institute, King Abdulaziz University, Jeddah, Saudi Arabia. E-mail: amalshaikhi1@kau.edu.sa

Abstract

This qualitatively based research study utilized a combination of multiple methods, which aimed at investigating the efficacy and reliability of employing cross-grading when assessing English as a Foreign Language (EFL) tertiary level learners' writing. It further explored the perceptions of the EFL teachers and learners regarding the cross-grading practices to provide a clearer understanding of this relatively unexplored line of research enquiry. It was set to answer the following research question: In what ways does cross-grading practice contribute to assessing EFL writing? The participants of this study were conveniently selected where the sample included four language instructors from different ethnic and cultural backgrounds, as well as four Saudi EFL learners. Semi-structured interviews were individually conducted with all eight participants. In addition, four one-on-one feedback sessions between language instructors and learners were observed to assess feedback effectiveness after the cross-grading sessions. The data analysis revealed that instructors had difficulty explaining the feedback on their learners' papers since they did not grade their students' papers themselves. Furthermore, students felt they did not benefit from the feedback sessions because they could not fully understand the external grader's markings and, thus inhibiting the learner's ability to improve and develop their writing. The study concluded with some pedagogical implications for the EFL writing assessment context.

Keywords: writing assessment, cross-grading, English as a Foreign Language (EFL), qualitative research, case study, blind marking, cross-assessing

1. Introduction

The Kingdom of Saudi Arabia (KSA) has witnessed rapid changes in all fields, especially in higher education, which is undoubtedly amongst the areas KSA is advancing to fulfill its 2030 Vision (Nurunnabi, 2017). To achieve the ambitious goals laid out in the 2030 Vision, KSA needs to establish its global position and meet its international needs in a world where English serves as a twenty-first-century *lingua franca* (Seidlhofer, 2013). In the past five decades, English as a Foreign Language (EFL) in KSA has witnessed numerous transformations as tertiary institutions strive towards improving the effectiveness of teaching and learning in the EFL classroom (Elyas & Picard, 2018). For the past two decades, the concept and practices of co-teaching has been largely promoted as a common model for collaboration at many educational institutions around the world (Bauwens & Hourcade, 1991; Cook & Friend, 1991; Sapon-Shevin & Schniedewind, 1992). However, many researchers have called for an in-depth critical analysis of the discourse and context of co-teaching (Brownell, Adams, Sindelar, Waldron, & Vanhover, 2006; Dieker & Murawski, 2003; Trent, 1998), the practice continues to be associated with various challenges and some researchers have argued that this practice has largely failed to reform the traditional education models (Scruggs et al., 2007).

While researchers have focused their attention on several areas of EFL teaching and learning in the Saudi context over the past two decades, one controversial area has been EFL writing assessment (Obeid, 2017). Teachers, as well as learners, have expressed concerns with regards to the writing assessment in KSA. On the one hand, teachers perceive a possibility for grade inflation due to the institution occasionally demanding higher grades than might be warranted (Al-Barrak & Al-Razgan, 2016). Furthermore, teachers find teaching writing in a second language to be a burden and an unpleasant process. Weigle (2007) echoed these sentiments, "Teachers often feel that assessment is a necessary evil rather than a central aspect of teaching that has the potential to be beneficial to both teacher and students" (p. 194). On the other hand, since EFL learners usually feel graded unfairly (Obeid,

2017), they question the true outcome of a writing assessment (Oraif, 2016) and the efficacy of the written corrective feedback provided by the teachers (Rajab et al., 2016). Notwithstanding the aforementioned research, writing and classroom-based assessment have been an area of research largely underrepresented in EFL contexts. Furthermore, external grading (cross-grading) remains unchartered in writing assessment research in the Saudi context.

Cross-grading was devised to reduce grade variation and grade inflation and provide an overall fair grade for the learners (Hunter & Docherty, 2011). However, a shortage of corroborating empirical studies exists both globally and in the Saudi context. Therefore, as its theoretical framework, this study utilized Foucauldian discourse analysis (FDA). According to Kendall and Wickham (1999), Foucauldian discourse analysis (FDA) can be described in the following five steps:

> The first step is a simple recognition that discourse is a body of statements that are organized in a regular and systematic way. The subsequent four steps are based on the identification of rules on: how those statements are created, what can be said (written) and what cannot, how spaces in which new statements can be made are created and making practices material and discursive at the same time (p. 42).

Currently in KSA, writing is primarily taught with a prescribed curriculum and textbook. It is mainly a top-down approach with a focus on the process of teaching writing in an EFL setting. Assessment is mainly exam driven, with a focus on mechanics, grammar, and word choice. Due to the low proficiency levels of students, cohesion and organization are not focused on until more advanced levels of language study. Benahnia (2016) has reported the difficulties mother tongue speakers of Arabic encounter when beginning to learn how to write in a foreign language, particularly English. Students who have not been exposed to learning languages with a different writing system, such as English, encounter various difficulties while some may continue to struggle throughout the process.

In the context of this study, the teachers are native English speakers from various English-speaking countries. During cross-grading sessions, teachers are required to use analytical rubrics which are then shared with students on the day of the feedback sessions. The academic management department is there to supervise the paper exchange among teachers, as well as to handle appeals from students' complaints. After careful consideration of the aforementioned research and current situation in KSA, this research focuses on the power relationships embedded in the interviews to explore the efficacy of the cross-grading practice in assessing tertiary EFL writing in the Saudi context.

1.1 Literature Review

As a key component in education, assessment comes in many forms. Generally and for this study, the definition of assessment: is "The process of gathering, interpreting, recording, and using information about pupils' responses to an educational task" (Harlen et al., 1992, p. 217). To take it a step further, while the concept of writing assessment remains a wide and varied one, it can also include a variety of classroom activities which range from basic testing and grading to more elaborate feedback. Specifically, written corrective feedback is communicated along with test results, and then both are used in making the assessment decisions. However, a pedagogical dilemma exists where teachers consciously realize the difficulty of being both, a teacher and an evaluator of their students, more like a coach and a judge at the same time (Connolly, 1998). This dilemma revolves around how teachers justify the grade given to their students and, at the same time, perform their other professional duty of encouraging their students to improve. Thus, since pedagogical strategies are constantly evolving to meet the demands of globalization in KSA's evolving tertiary education, new strands of discourse have woven their way throughout academic discussions. As such, one notable change in the practice of assessment has been the shift from the old practice of assessing what the learners can do (write) to the wider practice of assessing how to continue advancing the writing of the learners. In examining these changes in writing assessment, Michel Foucault's CDA serves to address complexities while still parsing the multiple perspectives embedded in power dynamics involved in the EFL classroom.

As one of the most influential social scientists today, Michel Foucault's theories regarding affective history allows a complex mapping of interpersonal relationships as well as a branching of multiple discourses (Schirato et al., 2010). Quite simply, it is used to understand how people interact and co-construct knowledge in terms of making sense of what is really going on. Foucault's CDA does not focus on absolutes, but instead on a truth that emerges from the discourses involved within a specific context as in the EFL classroom in KSA tertiary institutions (Foucault, 2013). One well-established perspective is, undeniably, how the practice of assessment is typically one of the teachers' least favorite topics and tasks when it comes to their professions. Conversely, students also do not typically fully appreciate assessment and often feel uneasy about elements of the assessment process (Segers & Tillema, 2011; Tavakol & Doody, 2015). While some teachers embrace a more traditional form of classroom

writing assessment (e.g., the task of essay writing), Kubiszyn and Borich (2013) proposed a more contemporary approach to classroom writing assessment practice (e.g., group writing assessment activities, research activities, and portfolio-based exercises). Relatedly, Brown et al. (2004) introduced another new strand into the discourse by presenting the efficacy of cross-grading when they showed its potential as a form of writing assessment that yields the end goal of creating reliable and valid results across several classes of the same course. As they proposed, cross-grading allows one instructor to assess the work of another instructor's learners with the end goal of creating reliable and valid results across several classes of the same course. By showing this potential in assessing EFL learners' writing in the case of their study, cross-grading moved from discussion to practice as Saudi students were enrolled in a language institute for one year to study English before entering the university.

Nevertheless, to ensure a successful foundation for cross-grading, all assessments should be contextually designed and grounded in ethics (Moore et al., 2009). Without either one or both, an assessment fundamentally lacks in both purpose and integrity and is, therefore, rendered useless (Crusan, 2010). With this being said, studies have shown that teachers' assessment practices can have a certain negative effect on the overall students' achievement and outcomes in the classroom (Alkharusi et al., 2014). Measuring the various qualities, effectiveness, and significance of classroom assessment are necessary to evaluate the outcomes of the teachers' assessment practices on students' achievement in the classroom (Allen et al., 2013).

Presently, very little research has been conducted on cross-grading and its effects on teachers and students in KSA at the tertiary level. This is the one of the reasons why this research is important to shed light on the current situation. Recently, Alshakhi (2018) examined how an analytic rubric would be more beneficial for students than a holistic one regarding cross-grading procedures. Literature analysis and first-hand research revealed that an analytic rubric would allow for increased contextual-based learning, increased flexibility, and empower teachers to be more involved with the assessment process of their own students. The shift from using analytic rubrics, as opposed to holistic ones, would benefit not only the students, but the teachers and coordinators in the department as well. Peaci (2020) is also in agreement with these findings and suggests that testing instruments should not only be used for the passing and failing of students as this tends to demotivate them, but rather the evaluation and testing of students should be constructive and positive. This supports that writing examinations need to be less focused on assessment through holistic purposes and therefore should be evaluated more analytically. A more analytical assessment of students' writing can result in increased perception of fairness and allows teachers to provide proper and constructive feedback to their students. Alghammas and Alhuwaydi (2020) emphasized the importance of building good relationships and rapport among teachers and students, and that teachers should try to know their students more personally. Teachers should be able to identify the strengths and weaknesses of their students and be able to address their needs during and after the lesson. A survey revealed that over 75% of participants confirmed the importance of positive teacher-student relationships. Finally, Alshakhi and Phan (2020) discussed many of the challenges and frustrations teachers have encountered with regards to cultural and institutional differences in Saudi Arabia. They reported on the difficulty expatriate teachers have experienced when trying to interact with students, resentments of privilege and deficiency, responses to challenges from a variety of cultural and social differences, as well as a feeling of helplessness to work with a prescribed curriculum with a top-down approach in regards to policy and practice.

In following the previously mentioned Foucauldian principles and as the discourse surrounding cross-grading progresses, exploring the perceptions of EFL teachers' classroom assessment practices is important as a means of developing learners' acquisition of the English language as well as allowing the teachers an opportunity to voice their concerns in that regard. Additionally, it is equally important to explore the learners' perception of the positive and negative attributes of the cross-grading process and how it impacts their overall EFL learning. As a result, this study aims to examine the following research questions:

1) How do Saudi tertiary EFL students perceive native English-speaking teachers' cross-grading practices in assessing their English academic writing?

2) How do native English-speaking teachers perceive cross-grading practices in assessing their Saudi tertiary EFL students' English academic writing?

2. Method

2.1 Research Design

This research is a qualitative case study on the current situation of EFL teachers and their students in regard to writing assessment through the use of cross-grading. Richards (2009) has revealed that qualitative studies have been conducted in numerous ways in language teaching and learning, including: teaching approaches, identity and socialization, narrative/lives, teacher beliefs and reflections, learner strategies, and differences of agendas between

teachers and learners to name a few. Duff (2008) has explained the importance and relevance of using case studies in educational and linguistic settings with the hopes of shaping educational policies and practices. Case studies can reveal situational complexities and themes in the realm of applied linguistics. The findings are not meant to be generalized, but can be used to closely examine the current situation in its specific context. This methodology looks less at the cause-and-affect and more on the evolving power relationships. Thus, Foucault's unique contribution to the social sciences provides invaluable tools for the study of modern classroom dynamics in the Saudi context. This research was chosen in order to shed light on the current situation of cross-grading in the foundation levels of an English language program at the tertiary level in KSA. It is an investigation as to whether cross-grading is appropriate and beneficial in the current academic setting as very little regarding this subject and context has previously been published.

2.2 Participants

The teachers and tertiary students who participated in this study were conveniently selected, which means the sampling is considered "a non-probability sampling technique in which subjects are selected because of their convenient accessibility and proximity to the researcher" (Quinlan, 2011, p. 221). These participants were chosen because they were the easiest to recruit for the study, and the researcher did not intend to select participants that are representative of the entire population. Due to the cultural restrictions in KSA where male and female teachers and students are not permitted to mix, only male participants were chosen for this study. The identities of the participants were protected and all the extracts in this study are presented using designated pseudonyms.

All four language instructors from the same institution were from Western, English-speaking countries, but they still had different ethnic and cultural backgrounds. The teachers were selected because the department employs native English speakers from various countries, but more importantly to investigate their point of view in regard to using cross-grading. The teachers selected come from more traditionally liberal grading systems and different contexts than what is currently being conducted in KSA. Due to their diverse teaching backgrounds and education, it was deemed an important aspect to look more closely at in this case study. All of the teachers had between five-ten years of teaching experience.

Jacob was from Australia and had a Masters in TESL along with experience in teaching in both his home country and KSA. Paul was from England and was CELTA certified. He also had the most EFL experience since he had taught in Asia for many years. Patrick was from Ireland and was the only one who had prior experience teaching in the Middle East. Lastly, Roger was from the U.S.A. and had the least experience teaching EFL. However, he was also well-trained to write items for exams. All four of the teachers taught the highest levels of EFL (3 and 4) in the university's academic program.

The four Saudi EFL learners studying in KSA already had at least one to two years of tertiary instruction in English for general purposes and were each in a section taught by the participating teachers. The students were between 18−19 years old and had approximately a low A2−low B1 proficiency level in English. Ahmad was in Jacob's Level 4 class; Fahad was in Paul's Level 3 class; Anas was in Patrick's Level 3 class; and Mohammad was in Roger's Level 4 class. Students from each of the teachers' classes were chosen as participants so that the discourse from both sides of the classroom power structures could be compared and analyzed. Furthermore, only sections of the two highest levels of writing classes were chosen because the feedback sessions were more dynamic with more fluent interaction between the teachers and students.

2.3 Data Collection

Considered the most common type of qualitative data collection and a favorite amongst researchers, a semi-structured interview style was selected for gathering data (Patten & Newhart, 2017). The semi-structured interviews were individually conducted with the participants who agreed to take part in this research study. Additionally, after teachers participated in cross-grading exchanges with colleagues, four one-on-one feedback sessions between the language instructors and the learners were observed to evaluate the effectiveness of the feedback sessions. One teacher and one learner were selected from each class, so the observations made pertained to two participants in each session. The advantages of using observations as a qualitative data collection method were elaborated by Creswell (2012), who stated that observations are commonly used by researchers to collect data. Through the use of observations, researchers are able to vary their roles throughout the data collection process. The duration of the interviews ranged from 35 to 50 minutes, and the feedback sessions observed were 50 minutes each. All interviews and observation sessions were audio-recorded for data analysis purposes.

2.4 Data Analysis

This research study was conducted for the purpose of exploring the discursive elements of the co-teaching practice

which is currently practiced in a tertiary English language program in KSA. The social theorist Michel Foucault's critical theories on discourse, power, dominance, and construction of the subject (McHoul et al., 2015), has informed the theoretical foundation and framework for this study. His CDA theory provided the main methodological structure for the data analysis in this study. Additionally, thematic coding was used in order to analyze the data. Gibbs (2007) explains that thematic coding is a type of qualitative analysis. It involves identifying common themes and ideas which can then be organized and categorized to create a framework for research.

2.5 Ethical Considerations

Lincoln and Guba (1985) identified four key areas of trustworthiness to consider while conducting research: credibility, transferability, dependability, and confirmability. Schwandt (2007) explains the importance of credibility by maintaining the input from participants and interpreting it in a way that maintains its original meaning without altering it from the researcher's perspective. In terms of transferability, Creswell (2007) highlights the importance of using thick and rich descriptions where possible to allow readers to decide if the information and conclusions can be applied to other contexts. Dependability revolves around the responsibility of the researcher and good documentation by the researcher which follows a logical process (Schwandt, 2007). Finally, interpretations must be based on the facts of the data and not simply the imagination of the researcher (Schwandt, 2007). All of these key areas of trustworthiness were used and carefully examined throughout the process of this research.

3. Results and Discussion

This section presents the results of the two research questions of this study:

1) How do Saudi tertiary EFL students perceive native English-speaking teachers' cross-grading practices in assessing their English academic writing?

2) How do native English-speaking teachers perceive cross-grading practices in assessing their Saudi tertiary EFL students' English academic writing?

The students expressed their unhappiness with the cross-grading practice and how it affects them. They were upset that the teachers who they regularly have classes with were not the ones responsible for assessing their work. They also expressed they were unhappy because if they had follow-up questions about their grade, they were unable to get any explanations as the teacher who assessed them was not their regular classroom teacher. Students also shared their dissatisfaction with the system and found that it does not support students in their learning.

The teachers were equally unhappy with the cross-grading practices and revealed that they often feel helpless when it comes to trying to improve the system due to the lack of support. They felt that top management doesn't listen to their concerns, and it appears they lack confidence in the teachers. Student appeals to the cross-grading assessments are always considered and the grades are endlessly rechecked which is ultimately a waste of time and resources.

The importance of fairness in writing assessment is clearly stated by Silva and Matsuda, "The language tester has no more inherent right to decide what is fair for other people than anyone else does. Nevertheless, the language tester does have the responsibility to use all means to make any language test she or he is involved in as 'fair' as possible" (2012, p. 124). As such, ethical codes related to assessment in general and to specifically writing assessment are vital and well-established in the policies and practices of institutions. These ethical codes ultimately result in evaluations that are fair, accurate, beneficial, and feasible at the same time (Green et al., 2007). This study reveals a consistency in the views among the students and among the teachers that cross-grading is not beneficial for teachers or students. Cross-grading may serve a purpose in different contexts and settings, but for the students and teachers involved in the scope of this research, it is deemed unfavorable and unfair. Even more than that, similar points are made across the teacher-student power dynamic.

3.1 Cross-Grading as a Learning Obstacle for Students

The EFL students felt that being taught by one teacher and graded by another was more than unfair. They viewed the cross-grading system as one that obstructed their learning and negatively affected their learning outcomes. Indeed, the students had strong views on what the purpose of grading should be. For example, Ahmad asserted, "Grading and feedback should mean to help us learn our mistakes and improve on them. It should be a helping tool, but it's never the case." With cross-grading not only making the feedback more difficult to understand, but also to benefit from, Mohammad lamented, "Instead of making things easy for us, cross-grading system makes it more difficult to comprehend the written feedback of the teachers." Overall, the students seemed to experience a disconnect between the classroom instruction and the feedback they received.

Since the teachers were having difficulty conveying and explaining the feedback comments to the students, the students received little useful information from the comments on their papers or from feedback sessions. Fahad conveyed his perspective:

> It is not acceptable to be graded by teachers who do not teach you. I'm so upset about the grades when I cannot discuss it with my classroom teacher and when I'm not allowed to go and see the grader. I prefer to ask my teacher about the feedback, and I am always comfortable with my classroom teacher.

Anas also expressed his concern with being unable to discern the feedback he received after the teachers had cross-graded the papers, "It is hard to understand the comments on my writing. I asked my teachers several times, but he doesn't have a convincing explanation because he was not the one who graded my paper."

One interesting aspect of this comment was how Fahad referred to his relationship with his teacher. Having teachers provide accurate, easy-to-understand feedback was important, but the rapport that the students had with their classroom teachers also played an important role in how students reacted to grades being assigned by an unknown teacher. Indeed, Mohammad explained the impact of this apparent grading anonymity on how students perceived the educational institution:

> Students' learning should be important for the teachers and management at the institute, but here the system doesn't support students and students' learning at all. There is a serious lack of facilitation and cross-grading is a great example of that.

These students' views were in line with the teachers' experiences of the grading system in this EFL context.

The cross-grading scheme appears to be the outcome of the top-down management system, which does not give space to the teachers' and students' input. Moreover, the EFL learners do not find it easy to understand the written feedback of the grader as they are more accustomed to their classroom teacher's teaching style and way of giving feedback. These factors contributed to the challenge of the students who expected the management to facilitate their learning and remove any hurdles in the process of getting feedback and grades on their writing quizzes. These findings suggest that professional and qualified EFL teachers need to establish themselves and engage in ethical practices, without which our profession might be discredited.

3.2 Inconsistency in Cross-Grading

In a language education setting, teachers perform multiple roles in the writing assessment process (Chung & Nam, 2018; Vogt & Tsagari, 2014). In the past decade, educational institutions worldwide have been moving away from the so-called objective testing as the major form of assessment and adopting more complex tasks that are more context-specific and more target-oriented (Philpott, 2016). Thus, institutions must consider how fairness and equity may be conceptualized in different forms of assessment (Morgan & Watson, 2002). Researchers have reported potential inequity sources that resulted from certain discrepancies with teachers being involved in assessment. The sources of inequity included: inconsistent application of assessment standards, systematic bias and poorly designed tasks (Danielson, 2017). Moreover, the findings of this study support the teachers' responses suggesting that the feedback given to the learners lacks consistency. Within this study's context, teachers interpreted the writing rubrics very differently which reflected a serious deficit in their understanding of the rubrics. As a result, the cross-grading practice did not benefit the EFL learners, which was a fact well-known to teachers. However, they also did not feel they had the support of the administration to address their concerns and correct these deficiencies.

3.3 Lack of Support and Motivation in Cross-Grading

Writing assessment procedures can undoubtedly be complex. In order to ensure credibility in the grades awarded to learners, these procedures must also be regulated. The teachers should be informed of the assessment policies at the beginning of the academic year/semester, and all assessment procedures must conform to established practices. Additionally, these policies should be designed solely to regulate the writing assessment process and not disempower the teachers by a top-down control and accountability, which may further exaggerate and complicate the issue of assessment (Ingersoll, 2009). Berryhill et al. (2009) echoed this sentiment by cautioning that any institution with a strong drive towards a set of academic goals has the potential to see a greater number of burned-out teachers. One participating teacher from the U.S., Roger, expressed his frustration with the cross-grading approach taken in this study, the lack of support from the administration, and the apparent futility in trying to change the system. He pointed out:

> I think the teachers are helpless to improve the system. It will be a great idea to scrap the existing system and allow the classroom teachers to grade the writing scripts of their students. It will help both the teachers and the students. Unfortunately, the top management does not listen to the teachers. They should listen to the

teachers' concerns. We often raise our concerns, but it doesn't help.

Clearly, Roger's frustration comes through, yet it is interesting to note how he wishes that teachers could grade their own students, which is also in line with the students' sentiments. Indeed, Mohammad from Roger's class was the one student who explicitly brought up the lack of support for students and students' learning.

The other teacher who specifically brought up issues with not feeling supported was Jacob. Like Roger, Jacob taught a higher-level class out of the two levels included in this study. He also resented that while holding teachers accountable, the administration never really went beyond a superficial evaluation of the cross-grading system. He revealed:

> Teachers are always criticized for the failure of students. EFL students and top leadership in our context never look into the root cause of the problems. I think teachers try their level best to be fair-seeming and grade papers on merit; however, the students are never pleased and neither is the top management. Sadly, there's no support from the top management and the students' appeals are always considered for re-checking of the grades. It appears as leadership lacks confidence in teachers.

Interestingly, Jacob tied the students' dissatisfaction with their grades to the administration's lack of support of the teachers. His observations seem to take a broader overall view with evaluative statements regarding the entire system of assessment he found himself working within. Paul shared his concerns that were similar to Jacob's in terms of being able to communicate with the administration, "There has to be communication channels established with the administration for the sake of improving the grading process."

Furthermore, the EFL teachers in this study seemed to find the cross-grading process as one that sapped their motivation and failed to hold their interest. Patrick, a teacher from Ireland, observed:

> I'm never motivated to grade the scripts of students whom I have not taught because the students always complain about the grades and the feedback provided to them. As a result, a lot of time is wasted on dealing with students' appeals. You will have to justify the awarded grades.

Like the other teachers, Patrick believed that the administration or top leadership was not cognizant of the grading complexities and was not ready to improve the current system. They considered it part of the top-down system that did not include the teachers' say.

Overall, the above excerpts illustrate the lack of administrative support for the EFL teachers who tried to be as fair as possible in their approach while working within a problematic system. As Webb (2006) asserted, the institutional improvements could be achieved through the transparent accountability procedures, in which "Teachers will work more effectively when data of student performances are made public and accountability decisions (i.e., incentives and sanctions) are levied in relation to the quality of those performances" (p. 201). However, the lack of teachers' voices in the existing grading system made it difficult for teachers to do their job professionally.

3.4 Offering a Solution and Strengthening Cross-Grading

The interviewed teachers did not simply complain about cross-grading and criticize the system they were working in. Instead, most offered solutions based on their experiences in prior positions. For example, Jacob observed how the lack of training had a negative impact on the entire cross-grading endeavor:

> As we have never been trained on the rubrics and their application, we all see the same rubric differently. You know, teachers have their own justifications for the comments they make on the learners' writing scripts. It spoils the whole effort. (Jacob)

By clearly defining the problem, Jacob provides a base from which to explore the options for addressing the matter. However, Roger's take on the challenges acknowledges how the tools being used with cross-grading may not need to be replaced:

> Teachers do not grade the script in a standardized manner. The existing rubrics may be appropriate for grading the scripts; however, the EFL teachers need to be properly trained to grade papers in a standardized fashion. (Roger)

Now, the strand of discourse can be traced, and its evolution tracked. Beginning with focusing on the institution's lack of support, the conversation evaluated the rubrics provided with statements being made concerning their quality. Then, the conversation shone a light back on the teachers coming full circle to where it began.

The next logical step in this new discourse thread would be to begin exploring exactly what changes would need to be made, so teachers could be properly trained, which would then eliminate the inconsistency and ambiguity in the

cross-grading writing assessment. Paul turned towards a well-established standardized test:

> IELTS writing and its grading is a great example of a standardized grading. The graders are often trained in a rigorous manner that enables them to grade the IELTS writing script accurately. Interestingly, there's no or less discrepancy in the grading of IELT examiners.

Through this comment, Paul is moving the discussion towards seeking an actual solution. Once again, these quotes suggest that the teachers take action in the form of required training that would greatly improve the quality of feedback given to EFL learners. In other words, no longer is it solely the lack of support from ambivalent administrations; rather, for educational institutions to eliminate sources of inequities, engaging teachers to become actively involved in writing assessment must be a top priority.

4. Conclusion

This study aimed to explore various areas of the cross-grading assessment process as perceived by EFL teachers as well as students. However, these issues need to be addressed holistically as the cross-grading process should not merely take the teachers' and students' beliefs and perceptions into consideration, but also the institution and its assessment policies. The main purpose for institutions to implement cross-assessing/co-assessing practices is to eliminate inconsistencies, biases, and grade inflation or deflation (Brownlie & Schnellert, 2009; Hunter & Docherty, 2011; Lynch, 1996). However, the findings of this study, substantiating previous relevant research, highlight several issues that negatively affect the assessment process when the institution implements the cross-grading procedure. Clearly, from the accounts given by the teachers as well as the students who participated in this study, issues were voiced about concerns relating to the cross-grading practices. Teachers were concerned about the logistical issue of time management when carrying out the cross-grading process, which might consume a large amount of time. The issue could be further amplified by the cases where some students may object to the grade or the feedback provided. Thus, the process may exhaust even more time from the teachers.

Furthermore, the teachers expressed their reservations about cross-grading where they believed that it created a gap between themselves and their students since they were not the ones providing the feedback. Additionally, the students conveyed their concerns regarding cross-grading. They felt there was disconnectedness with the cross-grader and a feeling of powerlessness regarding their grades, if they felt they were unfairly graded. They expressed dissatisfaction with their grades when cross-graded since they believed they were not objectively or fairly assessed. Therefore, students expressed their desire to receive much-needed constructive feedback which they felt was lacking when their writing was cross-graded. Likewise, Pitt and Winstone (2018) found that students preferred more personal and individualized feedback who stated, "anonymous marking might undermine the learning potential of feedback, and minimize the strength of the relationship between lecturers and students, which may minimize the role of dialogue in the feedback process" (p. 1183). Indeed, in several accounts in this study, students expressed their frustration and dissatisfaction with the cross-grader's feedback. As Alghammas and Alhuwaydi (2020) pointed out the important nature of positive student-teacher relationships, it would be necessary to align this with the assessment process as well. If students' own teachers are not allowed to assess and provide feedback to their own students during the cross-grading procedure, it certainly is not beneficial in having a positive effect on their relationship. At best, we have discovered students' frustrations with not being able to receive helpful feedback after cross-grading procedures due to the fact that an unknown teacher has assessed their work.

As such, we need to reiterate that assessment is generally accepted in the educational settings as a required measure of learning as well as a vital procedural requirement for obtaining credentials (Amin, 2011). Additionally, experienced and effective teachers should not perceive the assessment process as merely an add-on to their teaching tasks, but rather as an integral component of their teaching. Furthermore, writing assessment teachers need to be aware of any issues that might mar the assessment process. Thus, they need to eliminate any bias in their grading practices and provide meaningful and constructive feedback to the students they are cross-grading; thereby, they are helping those students benefit from the assessment process. Above all, tertiary level institutions adopting cross-grading should carry out thorough evaluations of the process and careful consideration of both teachers' and students' perceptions and beliefs to identify and establish best practices regarding writing assessment.

The current study has provided some evidence into the process of cross-grading where teachers as well as students expressed their concerns regarding certain issues related to this process. Due to the qualitative and small-scale nature of this study, the number of participants were few. This makes it difficult to generalize the results to a larger population. The lack of previous research in this particular context of tertiary students involved with cross-grading practices in KSA also made it difficult to compare previous cross-grading practices to this context. It is worth noting, this study might not have covered all the elements and factors which are related to cross-grading. Hence, the study recommends a deeper exploration of the area of cross-grading in education in order to determine more

conclusive results. It would be beneficial to conduct a study to explore the effects of cross-grading with the provision of feedback regarding the perceptions of teachers and students. It could be beneficial in the future for learning and building relationships with each other. The study also recommends conducting more longitudinal research and a wider consideration of cross-grading at other tertiary-level institutions worldwide with a broader base of teacher and student participants. Additionally, the study recommends involving management and course administrators who can provide significant value added to the study of the cross-grading process. Finally, the above recommendations can best be achieved via conducting a mixed-methods approach with randomly selected participants where data collection tools can include surveys, focus groups, and observations.

References

Al-Barrak, M. A., & Al-Razgan, M. (2016). Predicting students final GPA using decision trees: A case study. *International Journal of Information and Education Technology*, *6*(7), 528−533. https://doi.org/10.7763/IJIET.2016.V6.745

Alghammas, A., & Alhuwaydi, A. (2020). The weaknesses of English writing skills among undergraduate Saudi students majoring in English at Qassim University: A perspective of English faculty. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3621267

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014). Classroom assessment: Teacher practices, student perceptions, and academic self-efficacy beliefs. *Social Behavior and Personality: An international Journal*, *42*(5), 835−855. https://doi.org/10.2224/sbp.2014.42.5.835

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. *School Psychology Review*, *42*(1), 76−98. https://doi.org/10.1080/02796015.2013.12087492

Alshakhi, A. (2018). Revisiting the writing assessment process at a Saudi English language institute: Problems and solutions. *English Language Teaching*, *12*(1), 176. https://doi.org/10.5539/elt.v12n1p176

Alshakhi, A., & Phan, L. H. (2020). Emotion labor and affect in transnational encounters: Insights from Western-trained TESOL Professionals in Saudi Arabia. *Research in Comparative & International Education*, *15*(3), 305−326. https://doi.org/10.1177/1745499920946203

Amin, N. (2011). Critique and care in higher education assessment: From binary opposition to Möbius congruity. *Alternation*, *18*(2), 268−288. http://alternation.ukzn.ac.za/Files/docs/18.2/14%20Ami.pdf

Bauwens, J., & Hourcade, J. J. (1991). Making co-teaching a mainstreaming strategy. *Preventing School Failure: Alternative Education for Children and Youth*, *35*(4), 19−24. https://doi.org/10.1080/1045988x.1991.9944254

Benahnia, D. A. (2016). Teaching Writing to EFL/ESP Arab Learners: A Socioculturally-Based Approach. *Sino-US English Teaching*, *13*(4). https://doi.org/10.17265/1539-8072/2016.04.004

Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The effects of education accountability on teachers: Are policies too-stress provoking for their own good? *International Journal of Education Policy and Leadership*, *4*(5), 1−14. https://eric.ed.gov/?id=EJ898886

Brown, G. T. L., Glasswell, K., & Harland, D. (2004). Accuracy in the scoring of writing: Studies of reliability and validity using a New Zealand writing assessment system. *Assessing Writing*, *9*(2), 105−121. https://doi.org/10.1016/j.asw.2004.07.001

Brownell, M. T., Adams, A., Sindelar, P., Waldron, N., & Vanhover, S. (2006). Learning from collaboration: The role of teacher qualities. *Exceptional Children*, *72*(2), 169−185. https://doi.org/10.1177/001440290607200203

Brownlie, F., & Schnellert, L. (2009). *Collaborating to support all learners in English, social studies, and humanities*. Portage & Main Press.

Chung, S. J., & Nam, Y. (2018). Language assessment literacy of Korean EFL teachers: An investigation of their training experiences and needs. *Modern English Education*, *19*(1), 38−48. https://doi.org/10.18095/meeso.2018.19.1.04

Connolly, J. (1998). Cross-teaching/cross-grading: An experiment in teaching writing. *College Teaching*, *46*(1), 20. https://doi.org/10.1080/87567559809596227

Cook, L., & Friend, M. (1991). Principles for the practice of collaboration in schools. *Preventing School Failure:*

Alternative Education for Children and Youth, *35*(4), 6−9. https://doi.org/10.1080/1045988X.1991.9944251

Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.

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

Crusan, D. (2010). *Assessment in the second language writing classroom*. University of Michigan Press. https://doi.org/10.3998/mpub.770334

Danielson, C. (2017). Assessment for learning—For teachers as well as students. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 191−213). Routledge. https://doi.org/10.4324/9781315086545-7

Dieker, L. A., & Murawski, W. W. (2003). Co-teaching at the secondary level: Unique issues, current trends, and suggestions for success. *The High School Journal*, *86*(4), 1−13. https://doi.org/10.1353/hsj.2003.0007

Duff, P. (2008). *Case study research in applied linguistics*. Lawrence Erlbaum Associates.

Elyas, T., & Picard, M. (2018). A brief history of English and English teaching in Saudi Arabia. In C. Moskovsky & M. Picard (Eds.), *English as a foreign language in Saudi Arabia: New insights into teaching and learning English* (pp. 78−92). Routledge. https://doi.org/10.4324/9781315688466-3

Foucault, M. (2013). *Archaeology of knowledge*. Routledge. https://doi.org/10.4324/9780203604168

Gibbs, G. R. (2007). *Thematic coding and categorizing*. In *Analyzing qualitative data* (pp. 38−55). SAGE Publications, Ltd. https://doi.org/10.4135/9781849208574.n4

Green, S. K., Johnson, R. L., Kim, D.-H., & Pope, N. S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, *23*(7), 999−1011. https://doi.org/10.1016/j.tate.2006.04.042

Harlen, W., Gipps, C., Broadfoot, P., & Nuttall, D. (1992). Assessment and the improvement of education. *The Curriculum Journal*, *3*(3), 215−230. https://doi.org/10.1080/0958517920030302

Hunter, K., & Docherty, P. (2011). Reducing variation in the assessment of student writing. *Assessment & Evaluation in Higher Education*, *36*(1), 109−124. https://doi.org/10.1080/02602930903215842

Ingersoll, R. M. (2009). *Who Controls Teachers' Work?: Power and Accountability in America's Schools*. Harvard University Press.

Kendall, G., & Wickham, G. (1999). *Using Foucault's methods*. London: Sage Publications. https://doi.org/10.4135/9780857020239

Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement: Classroom applications and practice* (10th ed.). Wiley Global Education.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage. https://doi.org/10.1016/0147-1767(85)90062-8

Lynch, T. (1996). *Communication in the language classroom: A book about classroom interaction and how to teach communication skills to language students*. Oxford University Press Oxford.

McCalman, I., & Pickering, P. A. (2010). From realism to the affective turn: An agenda. In I. McCalman & P. A. Pickering (Eds.), *Historical reenactment: From realism to the affective turn* (pp. 1−17). Palgrave Macmillan. https://doi.org/10.1057/9780230277090_1

McGowen, R. (2002). Power and humanity, or Foucault among the historians. In C. Jones & R. Porter (Eds.), *Reassessing Foucault: Power, medicine and the body* (pp. 101−122). Routledge. https://doi.org/10.4324/9780203019481-9

McHoul, A., McHoul, A., & Grace, W. (2015). *A Foucault primer: Discourse, power and the subject*. Routledge.

Moore, C., O'Neill, P., & Huot, B. (2009). Creating a culture of assessment in writing programs and beyond. *College Composition and Communication*, *61*(1), W107−W132. https://library.ncte.org/journals/CCC/issues/v61-1/8315

Morgan, C., & Watson, A. (2002). The interpretative nature of teachers' assessment of students' mathematics: Issues for equity. *Journal for Research in Mathematics Education*, *33*(2), 78−110. https://doi.org/10.2307/749645

Nurunnabi, M. (2017). Transformation from an oil-based economy to a knowledge-based economy in Saudi

Arabia: The direction of Saudi vision 2030. *Journal of the Knowledge Economy*, *8*(2), 536−564. https://doi.org/10.1007/s13132-017-0479-8

Obeid, R. (2017). Second language writing and assessment: Voices from within the Saudi EFL context. *English Language Teaching*, *10*(6), 174−181. https://doi.org/10.5539/elt.v10n6p174

Oraif, I. M. (2016). The right approach in practice: A discussion of the applicability of EFL writing practices in a Saudi context. *English Language Teaching*, *9*(7), 97−102. https://doi.org/10.5539/elt.v9n7p97

Patten, M. L., & Newhart, M. (2017). *Understanding research methods: An overview of the essentials* (9th ed.). Taylor & Francis. https://doi.org/10.4324/9781315213033

Peaci, D. (2020). Writing evaluation in university English preparatory programs: Two universities of Turkey and Saudi Arabia. *Dil ve Dilbilimi Çalışmaları Dergisi*, *16*(1), 253−264. https://doi.org/10.17263/jlls.712798

Philpott, D. F. (2016). Redefining assessment in contemporary classrooms: Shifting practices and policies. In S. Scott, D. E. Scott & C. F. Webber (Eds.), *Assessment in education: Implications for leadership* (pp. 113−133). Springer. https://doi.org/10.1007/978-3-319-23398-7_6

Pitt, E., & Winstone, N. (2018). The impact of anonymous marking on students' perceptions of fairness, feedback and relationships with lecturers. *Assessment and Evaluation in Higher Education*, *43*(7), 1183−1193. https://doi.org/10.1080/02602938.2018.1437594

Quinlan, C. (2011). *Business research methods*. Cengage Learning.

Rajab, H., Khan, K., & Elyas, T. (2016). A case study of EFL teachers' perceptions and practices in written corrective feedback. *International Journal of Applied Linguistics & English Literature*, *5*(1), 119−131. https://doi.org/10.7575/aiac.ijalel.v.5n.1p.119

Richards, K. (2009). Trends in qualitative research in language teaching since 2000. *Language Teaching*, *42*(2), 147−180. https://doi.org/10.1017/S0261444808005612

Sapon-Shevin, M., & Schniedewind, N. (1992). If cooperative learning's the answer, what are the questions? *Journal of Education*, *174*(2), 11−37. https://doi.org/10.1177/002205749217400203

Schirato, T., Danaher, G., & Webb, J. (2012). *Understanding Foucault: A critical introduction* (2nd ed.). Sage Publishing.

Schwandt, T. A. (2007). *The SAGE dictionary of qualitative inquiry*. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412986281

Scruggs, T. E., Mastropieri, M. A., & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional Children*, *73*(4), 392−416. https://doi.org/10.1177/001440290707300401

Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, *37*(1), 49−54. https://doi.org/10.1016/j.stueduc.2011.03.008

Seidlhofer, B. (2013). *Understanding English as a Lingua Franca - Oxford applied linguistics*. Oxford University Press. https://doi.org/10.1002/9781405198431.wbeal0243

Silva, T., & Matsuda, P. K. (2012). *On second language writing*. Routledge. https://doi.org/10.4324/9781410600899

Spradley, J. P. (2016). *Participant observation*. Waveland Press.

Tavakol, M., & Doody, G. A. (2015). Making students' marks fair: Standard setting, assessment items and post hoc item analysis. *International Journal of Medical Education*, *6*, 38−39. https://doi.org/10.5116/ijme.54e8.86df

Thanh, N. C., & Thanh, T. T. L. (2015). The interconnection between interpretivist paradigm and qualitative methods in education. *American Journal of Educational Science*, *1*(2), 24−27. http://files.aiscience.org/journal/article/pdf/70380008.pdf

Trent, S. C. (1998). False starts and other dilemmas of a secondary general education collaborative teacher: A case study. *Journal of Learning Disabilities*, *31*(5), 503−513. https://doi.org/10.1177/002221949803100508

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, *11*(4), 374−402. https://doi.org/10.1080/15434303.2014.960046

Webb, P. T. (2006). The choreography of accountability. *Journal of Education Policy*, *21*(2), 201−214. https://doi.org/10.1080/02680930500500450

Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing, 16*(3), 194–209. https://doi.org/10.1016/j.jslw.2007.07.004

Copyrights