



IMAM KHOMEINI
INTERNATIONAL UNIVERSITY



Print ISSN: 2676-5357
Online ISSN: 2676-5965

On the Effect of Diagnostic Self-, and Peer-assessment on Reading Comprehension: Examining EFL Learners' Diagnostic Rating Accuracy across Various Genres

Masoome Azmoode Sis Abad¹, Dr. Gholam-Reza Kiany², Dr. Gholam-Reza Abassian³

¹PhD candidate of TEFL, Science and Research Branch, Islamic Azad University, Tehran, Iran

Email: Azmoode@srbiau.ac.ir

²ELT Professor at TarbiatModares University, Tehran, Iran. (Corresponding author)

Email: kiany_gh@modares.ac.ir

³TEFL Associate Professor at Imam Ali University, Tehran, Iran

Email: gabbasian@gmail.com

Article info Abstract

Article type: **Abstract**

Research
article

Received:
29/04/2023

Accepted:
01/11/ 023

The present research responds to the call for implementing diagnostic assessment through learner involvement. This study followed two aims: first, examining the distinctive effects of diagnostic self- and peer-assessments on EFL learners' reading comprehension; second, exploring students' diagnostic rating accuracy in various reading genres. To achieve this, a number of diagnostic reading comprehension tests and a checklist were developed, the construction of which was based on the learners' challenging reading subskills. The participants include 60 English translation university students. During 12 weeks of instruction, the rating accuracy of the subjects' diagnostic self- and peer-assessments was investigated while the instructor ratings were taken as the yardstick. Data analysis, using T-test and MANOVA confirmed that the two groups had improvement in reading comprehension but there was no significant difference between the two groups' gain. In addition, no statistically significant difference was detected among the accuracy of diagnostic self-, peer- and instructor ratings except for the assessment of the main idea and supporting details subskill in the descriptive genre and cause and effect subskill in the narrative genre. This study can offer evidence for conducting diagnostic assessments through learner engagement to provide students with appropriate feedback and remedial instruction.

Keywords: *Diagnostic assessment, Reading Comprehension, Reading Genres, Reading Subskills, Rating accuracy*

Cite this article: Azmoode Sis Abad, M., Kiany, Gh. R., & Abassian, Gh. R. (2024). On the effect of diagnostic self-, and peer-assessment on reading comprehension: Examining EFL learners' diagnostic rating accuracy across various genres. *Journal of Modern Research in English Language Studies*, 11(2), 177-202.

DOI: 10.30479/jmrels.2023.18703.2204

©2024 by the authors. Published by Imam Khomeini International University. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0>



1. Introduction

For assessing reading skills, a great number of assessment designs often report a single score for the overall performance of a particular learner; consequently, they fail to provide fine-grained information on the problematic areas of reading that need attention on the part of teachers and learners (Lee & Sawaki, 2009). As the complexity inherent in reading construct makes it a challenging skill for L2 learners (Zhou & Siriyothin, 2011), the availability of such detailed information on learners' strengths and weaknesses can help instructors diagnose the students' specific reading deficiencies and do instructional planning that suits their learning needs.

Since the need for such detailed diagnostic information is widely called for, diagnostic assessment has become an active research area during the last few years. Through diagnostic assessment, learners' strengths and weaknesses in a specific linguistic and communicative competence domain are identified in order to offer related diagnostic feedback and remedial instruction (Lee & Sawaki, 2009). Despite the significance, diagnostic assessment has been almost poorly theorized (Alderson et al., 2015). Furthermore, the majority of the conducted research in the realm of diagnostic assessment are based either on the retrofitting approach (eg., Chen & Chen, 2016; Javidanmehr & Anani Sarab, 2019), or on computer-based diagnostic tests (Alderson 2005; Urmston et al., 2013) which despite their convenience, these approaches mostly result in less accurate diagnostic inferences about the learners' performance.

The scarcity of well-designed diagnostic research in SFL assessment is partly due to the fact that the role of diagnostic tests has not been clearly justified in educational assessment (Liu, 2014); consequently, there exist a few real diagnostic tests in language learning contexts (Alderson, 2010). In recent years some diagnostic tests of language proficiency have been developed [eg., DIALANG (Alderson & Huhta, 2005)]; however, the majority of such tests are run a priori, thus they rarely provide an ongoing account of the learners' abilities in instructional settings. Therefore, the demand for developing purpose-built diagnostic tests has been substantially noticed in educational assessment (Jang 2005, 2009; Sawaki et al., 2009). More importantly, for the application of the existing diagnostic tests, a prominent principle of diagnostic assessment, which is benefitting from diverse stakeholders' views (eg., self, peer, teacher ...) in diagnosis, has been almost overlooked (Alderson et al., 2015).

Since little is known about how to do diagnosis in high-level reading skills and resources (Harding et al. 2015), a range of different reading skills and subskills have been the focus of diagnostic tests in recent years. For distinguishing

salient features of reading subskills, the topic and the genres of the texts need to be considered (Harding et al. 2015); thus, for developing recent diagnostic tests like DELTA (Urmston et al., 2013) particular attention has been given to a number of reading subskills across different genres. Nevertheless, text types and characteristics of reading genres were barely noticed in EFL reading instructions. Therefore, in the current study a number of diagnostic reading tests, tapping the learners' weaknesses in particular subskills, were developed based on texts in four different reading genres. Due to the importance of stakeholders' involvement in the diagnostic assessment process (Alderson et al., 2015), this study chose to focus on self, peer, and instructor involvement. The present research aimed at examining the distinctive effects of learners' diagnostic self-, and peer-assessments on promoting reading comprehension while probing their diagnostic rating accuracy in various reading genres. It is noteworthy that employing such diagnostic approaches would be highly significant in teaching and learning contexts since learners can gain more awareness concerning their own strengths and weaknesses in a skill.

This study tried to address the following research questions:

1. Is there any statistically significant difference between the effect of diagnostic self- and peer assessment on promoting reading comprehension of Iranian EFL learners?

2. Is there any statistically significant difference between the accuracy of self-, peer-, and instructor ratings in diagnostic assessment of Iranian EFL learners' reading comprehension in four reading genres?

- 2-1. Is there any statistically significant difference between the accuracy of self-, and instructor-ratings in diagnostic assessment of Iranian EFL learners' reading comprehension in four reading genres?

- 2-2. Is there any statistically significant difference between the accuracy of peer-, and instructor-ratings in diagnostic assessment of Iranian EFL learners' reading comprehension in four reading genres?

- 2-3. Is there any statistically significant difference between the accuracy of self-, and peer-ratings in diagnostic assessment of Iranian EFL learners' reading comprehension in four reading genres?

2. Literature Review

2.1. L2 Reading Comprehension

Reviewing the literature reveals that so far numerous researchers have addressed the challenging nature of reading comprehension (eg., Goodman, 1967; Kintsch, 1988; Rost, 1993) and they put forward different conceptualizations concerning the construct, components and processes of this skill. For instance, while some discriminate between the reading process and product (Gray, 1987; Myers, 1991), others refer to a dichotomy between reading approaches such as top-down and bottom-up (Alderson, 2000, Grabe, 2009) or the interplay between the two (Kintsch, 2005). In fact, L2 reading comprehension is a multifaceted construct encompassing different cognitive, linguistic, and non-linguistic skills covering low-level skills to higher-level (Harding et al., 2015). However, in spite of the universal cognitive mechanisms engaged in the process of reading comprehension, offering a comprehensive definition that suits various contexts might not be feasible (Toprak & Cakir, 2020).

Is reading considered as a multi-divisible skill or as a single global construct? This is highly a controversial issue and there is currently no universally accepted premise in this regard. As Rost (1993) noted, contradictory positions about the divisible nature of reading skills had already been formulated: while some believe in general-factor theories (Goodman, 1976; Vacca, 1980) and consider reading as one single, global, integrated skill, others in divisible view of reading believe in multiple-factor theories (Davis, 1944; Spearritt, 1972). Although the literature predominantly suggests the latter view (eg., Chen et al., 2023; Elahi, 2016; Farhady & Daftarifard, 2006; Tengberg, 2018), there isn't yet any consensus among the advocates on the number, type and scope of reading skills (Karakoc, 2019; Song, 2008). To do research on comprehension subskills, different statistical approaches such as factor analyses or intercorrelations have been utilized; nevertheless, still there is little agreement on how the processes of reading comprehension can be categorized validly (Tengberg, 2018).

Notably, in many of the current assessment practices, the divisible view of reading skill is advocated; in diagnostic assessment research, for example, the underlying reading subskills are differentiated for various learning objectives like offering diagnostic feedback (Aryadoust, 2017). In spite of a growing research interest in multi divisible view of reading, not only there isn't any agreement among expert judges on the kind of subskills tested by different test items, but also there isn't a specific hierarchy of difficulty among different reading subskills; therefore, reaching to such understanding is highly dependent on the text genre, text topic, readers' knowledge and purposes of reading (Harding et al., 2015).

However, so far little has been done to investigate the aforementioned areas in reading subskills research. Reading genre, for instance, has attracted particular research attention in recent years (Esfandiari & Jafari, 2021; Toledo, 2005; Yin, 2018; Zhou & Siriyothin, 2011), but the majority of the studies are either on genre analysis or on genre-based instruction and exploring reading genres in relation to more specific notions such as reading subskills is a less tapped research area. Considering that texts with diverse genres invoke different processes, investigating such an interactive relationship between reading subskills and text genres is of paramount importance (Jang, 2009).

Owing to the fact that in SFL reading high-level skills and resources are involved, diagnostic inquiries in such areas would certainly be welcome (Harding et al., 2015). Therefore, in the current research a number of high-level reading subskills (distinguishing between fact/opinion, cause/effect, and main ideas/supporting details), which were diagnosed to be challenging for the participants, were the focus of the research. In addition, four widely-used reading genres in Iranian EFL tertiary education (argumentative, expository, descriptive, and narrative) were concerned in this study.

2.2. Application of Diagnostic Assessment to Reading Assessment

Understanding L2 learners' problems in reading comprehension requires precise identification and diagnosis of their weaknesses in this domain. Yet, it is one of those pedagogical insights that is easier said than acted upon; therefore, the challenges of conducting diagnostic reading assessments have repeatedly been addressed previously (e.g., Alderson, 2005, 2010; Harding et al., 2015). Despite the significance, research in the realm of diagnostic assessment is at its early stage and until recently the area of second and foreign language (SFL) assessment lacked a sound theoretical basis for diagnostic assessment practices (Alderson et al., 2015). However, in recent years a theory of diagnosis encompassing a set of principles has been proposed by Alderson et al. (2015).

Reviewing the related literature reveals that few diagnostic reading studies (eg., Nikmard & Tavassoli, 2020; Poulaki et al., 2020) have been done in the EFL contexts. In fact, the majority of the diagnostic reading research (eg., Javidanmehr & Anani Sarab, 2019; Ranjbaran & Alavi, 2017; Ravand, 2015) had been conducted in the area of cognitive diagnostic assessment, for performing which researchers had utilized ex post facto designs and applied cognitive diagnostic models to non-diagnostic tests. Though, estimating cognitive strengths and weaknesses of learners in a skill based on such designs suffer from a basic problem, because for using models of cognitive diagnostic assessment one needs to employ new test development procedures which are in accordance with the key features of these

models (Jang, 2009). Furthermore, a few diagnostic tests have been designed so far, most of which are criticized for providing pre-programmed and general feedback to learners and other stakeholders (Harding et al., 2015).

In the design of many diagnostic assessment studies, learner involvement has been hardly employed and only in a few diagnostic inquiries, (eg., Alderson, 2005; Harding et al., 2015, Jang, 2005) self-assessment was incorporated. Yet, involving learners in assessment procedures has been largely valued in educational assessment, asserting that students' participation enhances the assessment practices and reinforces autonomy and motivation in learning (Oscarson, 1989). With a close look at the body of research, we find advocates discussing the reliability and validity of learners' assessment in general (eg., Han, 2018; Ma & Winke, 2019), and the accuracy of their assessment in particular (e.g. Birjandi & Siyyari, 2010; Han & Riazi, 2017; Lu, 2018); nevertheless, one can find fewer studies on self- and peer- assessment accuracy in receptive skills, like reading (e.g. Ashton, 2014; Paleczek et al., 2015). With respect to such gaps, research on diagnostic self- and peer-assessment of reading comprehension with an eye on rating accuracy is a noticeable lack in diagnostic assessment studies.

3. Method

3.1. Participants

Sixty Iranian EFL learners (18 to 23 years old) took part in the study; they were English translation students at Islamic Azad University, Shahr Qods Branch. The participants were taking a four-credit course, Reading Comprehension (I). The subjects were selected based on their performance on the Oxford Placement Test (OPT) ($M = 33.95$, $SD = 5.271$) and a researcher-made reading comprehension pretest ($M = 24.12$, $SD = 4.11$) in two intact classes.

In the pilot study, 40 students participated. They majored in English translation at the same university. Moreover, two external raters who were ELT instructors were invited to this study. In addition to the experts, the instructor/ researcher who was a PhD candidate in TEFL also rated the students' samples.

3.2. Instruments

3.2.1. *English Language Proficiency Test*

The OPT, version 1.1 UCLES (2001) was administered to check the homogeneity of the participants.

3.2.2 Reading Comprehension Pretest and Posttest

The pretests and posttests had three passages with 39 items which were constructed based on the learners' challenging reading subskills.

3.2.3. Reading Comprehension Passages

The main reading materials including 18 passages, six of which were used for pretest and posttest, were taken from well-known books including Read This 2 (2010) and For and Against (1968), the simplified version. The texts were selected in four different reading genres and had approximately the same level of difficulty, computed through Fry's Readability Index (11 on average).

3.2.4. Diagnostic Reading Comprehension Tests

In this research, 12 diagnostic reading comprehension tests were developed by the researchers. For each test, 10 items were constructed, five of which geared at testing one particular pair of subskills.

3.2.5. Diagnostic Reading Comprehension Checklist

To assess the participants' strengths and weaknesses in the intended reading subskills diagnostically, a checklist was designed by the researchers. It had five levels of abilities (excellent, good, fair, poor, and non-reader) for rating the learners' performance (see Appendix A).

3.3. Procedure

The present research was a pre-experimental one with a two-group pretest-posttest design. It had two distinct phases; a pilot study and a main study.

3.3.1. Pilot Study

A number of diagnostic reading comprehension tests and a checklist were developed during the pilot study.

3.3.1.1. Developing Diagnostic Reading Comprehension Tests, the procedures proposed by Alderson et al. (1995) were used for developing diagnostic reading comprehension tests. In fact, the steps which were applicable for developing low-stake, classroom-based tests were taken:

- test specifications
- item writing and revising
- piloting and analysis
- training the raters
- monitoring raters reliability
- validation

3.3.1.1.1. Test Specifications, the steps that were taken to specify the characteristics of the tests are listed below:

3.3.1.1.1.1. Developing the L2 Reading Subskill List, first a reading comprehension subskills list was created according to the previous literature (eg., Alderson and Huhta 2005; Fletcher, 2006; Jang, 2009; Urmston et al., 2013). Then, the reading comprehension courses offered in the study context were explored, so that the course content and their requirements were examined thoroughly. These efforts revealed that the focus of the instruction in Reading Comprehension (I) was mainly on more basic reading subskills so that a number of related subskills were identified and some reading activities were designed accordingly. As the next step, the students read some short texts (focusing on specific subskills), answered the comprehension questions, and identified the ones that tested a particular subskill. The analysis revealed that “using knowledge of syntax” and “understanding explicitly stated information” were the easiest reading subskills for the participants whereas distinguishing between “fact/opinion”, “cause/effect”, and “locating main ideas/supporting details” were difficult for them. Therefore, this study ultimately focused on the challenging reading comprehension subskills.

3.3.1.1.1.2. Preparing the Reading Passages, for selecting reading passages, Nuttall’s (1996) criteria were applied; it includes suitability of content, exploitability, and readability.

3.3.1.1.1.3. Test Methods, for specifying the test methods, the characteristics of “truly” diagnostic tests outlined by Alderson and Huhta (2011), were used. Thus, “multiple-choice question” and “short answer format” were considered.

3.3.1.1.1.4. Marking Criteria, for developing the diagnostic reading comprehension checklist, the researchers benefited from Bachman’s (1990) guidelines for designing rating scales, so the following procedures were used:

- targeting the intended reading subskills
- defining the reading subskills operationally
- categorizing the abilities in five levels of performance
- describing the features of each performance level
- setting the cut-off score to each performance level

3.3.1.1.2. Item Writing and Revising, based on the test specifications (see Appendix B), the test items were constructed. Then, the items and the whole tests were evaluated with respect to the initial test specifications.

3.3.1.1.3. Piloting and Analysis, the constructed tests were piloted (based on the instruction similar to the main study), and basic statistical techniques for

analyzing tests and item performance were applied.

3.3.1.1.3.1. Reliability of the Diagnostic Reading Tests and Checklist, after administrating the tests, the reliability of the reading pretests ($r = .887$, $r = .903$, & $r = .912$) and posttests ($r = .914$, $r = .899$, & $r = .849$) and 12 midterm tests were estimated (see Table 1). Further, the item facility (IF) and item discrimination (ID) of all the tests were investigated. The checklist reliability was also estimated (.804), using Cronbach's alpha.

Table 1

Reading Tests Reliability Statistics

Midterm Tests	Cronbach's Alpha	Item N0
Test one	.850	10
Test two	.802	10
Test tree	.963	10
Test four	.852	10
Test five	.944	10
Test six	.934	10
Test seven	.909	10
Test eight	.884	10
Test nine	.829	10
Test ten	.922	10
Test eleven	.871	10
Test twelve	.858	10

3.3.1.1.4. Training the Raters, to brief the raters on the purpose of the research, a two-hour session was held for them.

3.3.1.1.5. Monitoring Raters Reliability, this step is related to ways to ensure inter-rater and intra-rater consistency (Alderson et al., 1995). For the three raters who assessed the performance of the learners, the inter-rater reliability was estimated. There were significant agreements between the first and second rater ($r(38) = .883$), first and third rater ($r(38) = .757$); and second and third rater ($r(38) = .779$), which all represent a large effect size, $p = .000$). To estimate the intra-rater reliability, 20 samples were randomly selected and the instructor rated them once again. A significant agreement between the first and second ratings ($r(18) = .882$) was shown.

3.3.1.1.6. Validation, the content validity of diagnostic reading tests and the

checklist were confirmed by the experts after some revisions were made. Further, to explore the underlying constructs of the tests, an exploratory factor analysis was run on the first pretest. The test had 15 items and measured three constructs of cause/ effect, main idea/ supporting details, and fact/opinion. First, the sampling adequacy and sphericity assumptions were examined. The KMO index of .783 which was higher than the minimum acceptable criterion of .60 (Field, 2018) was adequate sample size. The significant results of Bartlett's test ($\chi^2 (105) = 249.71, p < .05$) showed that the data and its correlation matrix were factorable. The variance of the three factors, which were extracted from the test, was 59.62.

The factor loadings of the 15 items are shown in Table 2. However, item 8 which was designed to load under the second factor, had its loading under the first factor. Basically, it results from the inadequate sample size of the pilot study, because factor analysis which is essentially a large sample technique requires sufficient sample size to guarantee the reliability of the factor solution (Denis, 2021).

Table 2

Matrix of Rotated Component for First Reading Pretest

	Component		
	1	2	3
First factor: Fact and Opinion (Cronbach's alpha = .837)			
q13	.796		
q15	.764		
q14	.758		
q12	.692		
q8	.619	.555	
q11	.611		
Second factor: Main Idea and supporting details (Cronbach's alpha = .838)			
q6		.823	
q10		.823	
q9		.671	
q7		.633	
Third factor: Cause and Effect (Cronbach's alpha = .725)			
q2			.767
q4			.716
q3			.655
q1			.623
q5			.574

After proving the construct validity of the first pretest, its correlations with other tests were also computed. All Pearson correlations were significant (i.e. $p < .05$), enjoying large effect sizes. Thus, the tests which were employed in this study had criterion-referenced validity. In the main study, also some tests were randomly selected and their underlying construct validity was computed again, using “principal axis factoring”. All the selected tests enjoyed construct validity.

3.3.2. Main Study

The OPT and reading comprehension pretest were administered at the outset of the research. The participants received related training and brief instructions to use the diagnostic reading checklist. The instructional treatment took 12 weeks. In fact, each pair of reading subskills was taught and tested in four different reading genres during four subsequent sessions. Every session the learners read an unseen text silently, answered the comprehension questions and filled out the diagnostic checklist. To do so, in the diagnostic self-assessment group the students identified the questions that tested a particular subskill and rated both their own level of diagnosis and reading comprehension. In the other group, the learners identified the items that tested the intended subskills and exchanged their answer keys in pairs to rate their peer’ diagnostic level and reading comprehension. The instructor assessed all the students’ samples and provided them with feedback in a weekly basis throughout the course; so that the learners could ponder upon any possible assessment evaluative mismatches. Finally, the reading comprehension posttest was administered.

3.3.4 Data Analysis

For data analysis, the obtained data were analyzed through SPSS statistical software and appropriate parametric statistics (i.e., T-test and MANOVA) were applied.

4. Results and Discussion

4.1. Results

Before testing the null hypotheses, the normality of the OPT, the pretest, and the posttest of reading comprehension were explored through skewness and kurtosis indices. The normality of the data was assured (Table 3) since the indices were lower than ± 2 (George & Mallery 2020).

Table 3
Normality Indices of OPT and Reading Pre-, Post Tests

Group		N	Skewness		Kurtosis	
		Statistic	Statistic	Std. Error	Statistic	Std. Error
Diagnostic Self-assessment	OPT	30	.408	.427	.691	.833
	Pretest	30	-.406	.427	-.147	.833
	Posttest	30	-.296	.427	-.710	.833
Diagnostic Peer-assessment	OPT	30	-.496	.427	.429	.833
	Pretest	30	.260	.427	-.499	.833
	Posttest	30	-.451	.427	.841	.833

4.1.1. Research Question One

An independent-samples T-test was run to answer the first question. Before that, the equality of variance (Levene's test) was confirmed, it was 0.33. Based on the results, there was no significant difference between the diagnostic self-assessment group ($M= 14.5$, $SD= 1.60$) and diagnostic peer-assessment group ($M=14.50$, $SD= 1.46$; $t(29) = 0.379$, $p=.7$, two-tailed). The effect size was also very small (eta squared = .002).

4.1.2. Research Question Two

To answer the second research question, three sub-questions were formulated. The assumptions of equality of variance (Levene's test) and homogeneity of variance-covariance matrices (Box's test) were checked and none of them were violated.

4.1.2.1. Investigation of the First Sub- question, the results of MANOVA for investigating the rating accuracy of diagnostic self-assessment in an argumentative genre show no statistically significant difference between the two groups on the combined dependent variables, $F(3, 60) = 2.47$; Wilks' Lambda = .88, $p = .071$; partial eta squared = .11. MANOVA in descriptive genre also shows no statistically significant difference between the groups, $F(3, 60) = .55$; Wilks' Lambda = .97, $p = .647$; partial eta squared = .02. For investigating how accurate the diagnostic self-assessment group was in expository genre, the results of MANOVA show a statistically significant difference between the groups, $F(3, 60) = 4.78$; Wilks' Lambda = .79, $p = .005$; partial eta squared = .20. Considering the results for the dependent variables separately, Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of cause and effect subskill in expository genre (see Table 4). Investigating the rating accuracy of the diagnostic self-assessment

group in the narrative genre through MANOVA shows a statistically significant difference between the groups, $F(3, 60) = 15.18$; Wilks' Lambda = .55, $p = .000$; partial eta squared = .44. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of cause and effect subskill in narrative genre.

Table 4
Between-Subjects Effects Tests

Source	Dependent Variable	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Group	diagnosis3.Cause. Expos	7.350	1	7.350	10.239	.002	.150
	diagnosis7.Fact. op.Expo	1.204	1	1.204	1.781	.187	.030
	diagnosis11.Main. sup.Exp	.600	1	.600	.572	.452	.010

4.1.2.2. Investigation of the Second Sub-question, the results of MANOVA for probing rating accuracy of diagnostic peer-assessment in argumentative genre show a statistically significant difference between the groups, $F(3, 60) = 5.39$; Wilks' Lambda = .77, $p = .002$; partial eta squared = .22. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of main idea and supporting details in argumentative genre. Investigating rating accuracy of diagnostic peer assessment in the descriptive genre through MANOVA shows a statistically significant difference between the groups, $F(3, 60) = 3.61$; Wilks' Lambda = .83, $p = .019$; partial eta squared = .16. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of cause and effect subskill in the descriptive genre. The results of MANOVA in the expository genre show a statistically significant difference between the groups, $F(3, 60) = 11.67$; Wilks' Lambda = .61, $p = .000$; partial eta squared = .38. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of cause and effect subskill in the expository genre. For investigating the rating accuracy of this group in the narrative genre, the results of MANOVA show a statistically significant difference between the groups, $F(3, 60) = 16.43$; Wilks' Lambda = .53, $p = .000$; partial eta squared = .46 are shown. Bonferroni adjusted alpha level of .017 was in diagnostic assessment of cause and effect and main idea and supporting details subskills in the narrative genre.

4.1.2.3. Investigation of the Third Sub-question, the results of MANOVA in the argumentative genre show no statistically significant difference between the groups, $F(3, 60) = 2.56$; Wilks' Lambda = .87, $p = .064$; partial eta squared = .12. The results show no significant difference between the groups in the argumentative genre. MANOVA in descriptive genre shows a statistically

significant difference between the groups, $F(3, 60) = 3.21$; Wilks' Lambda = .85, $p = .030$; partial eta squared = .14. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of main idea and supporting details in descriptive genre. In the expository genre, the results of MANOVA show no statistically significant difference between the groups, $F(3, 60) = .70$; Wilks' Lambda = .96, $p = .555$; partial eta squared = .03. MANOVA in the narrative genre shows a statistically significant difference between the groups, $F(3, 60) = 5.32$; Wilks' Lambda = .77, $p = .003$; partial eta squared = .22. Bonferroni adjusted alpha level of .017 was only in diagnostic assessment of cause and effect subskill in the narrative genre.

4.2. Discussion

This study was performed to examine the effects of EFL learners' diagnostic assessments on reading comprehension while probing the rating accuracy of their assessments in various reading genres. The application of diagnostic assessment in the present research revealed the learners' reading strengths and weaknesses in four genres; while some reading subskills were difficult to learn and diagnose for the learners, they eventually made improvements in reading comprehension.

In recent years, diagnostic assessment of reading comprehension has received increased research interest, particularly on the part of cognitive diagnostic assessment researchers (e.g., Chen & Chen, 2016; Jang 2009; Toprak & Cakir, 2020); however, little research attention has been given to techniques through which diagnostic data could be gathered from the assessment stakeholders. In this regard, while the importance of research on teacher diagnosis was noticed more (eg., Finkbeiner & Schluer, 2017; Shefelbine & Shiel, 1990), conducting a diagnostic assessment by learners has not been fully acknowledged in the related literature. This study accounted the learner involvement in the process of diagnostic assessment and confirmed that the participants' reading comprehension in both groups improved considerably. There are two plausible explanations for this result. The first is that diagnostic assessment through learner involvement could substantially increase the participants' awareness and provided them with necessary insights about the nature and causes of their weaknesses. The second one, which poses adequate empirical evidence in the literature, is that involving learners in the assessment process can promote their autonomy and self-regulatory learning (Butler, 2018; Butler & Lee, 2010; Paris & Paris, 2001), which can ultimately help them manage the learning process more actively and enhance their language skills.

Furthermore, exploring the rating accuracy of the learners' diagnostic assessments in different reading genres showed no significant difference between the groups; however, the learners in the diagnostic self-assessment group were

more accurate in diagnosis. The subtle superiority of the self-assessment group may indicate that training the learners for applying diagnostic assessment helped them improve the ability to recognize and trace their own reading difficulties better, hence they might get more accurate in this process. In line with the purposes of this research, Yang (2021) in a study investigated the validity or the accuracy of three diagnostic instruments for assessing L2 learners' reading comprehension and proved that all the instruments enjoyed assessment effectiveness. Furthermore, the findings show that the two groups were accurate in assessing almost all subskills except for the main idea/supporting details in the descriptive genre, and cause/effect in the narrative genre. It can be concluded that these two subskills were more challenging for the learners. This result is quite compatible with a cognitive diagnostic assessment study which confirmed that the subskills of cause/effect and main idea/supporting details were the least mastered subskills by the learners (Javidanmehr & Anani Sarab, 2019). In the same vein, Ravand (2015) also proved that the second most difficult subskill in his study was the main idea for the test takers. To be cautious in generalizing the results, it is worth noting that, unlike the present study which developed diagnostic tests and applied constant assessment of the learners' performance,

the above cognitive diagnostic assessment studies have mainly relied on retrofitting of existing non-diagnostic tests for diagnostic purposes rather than on designing a set of tests with diagnostic designs right from the outset.

The learners' inaccuracy in diagnostic assessment of the subskills in descriptive and narrative genres may also be attributed to the challenging nature of the two genres. Contrary to this finding, some other studies (Carrell & Connor, 1991; DuBravac & Dalle, 2002; Sahin, 2013) proved that narrative and descriptive texts were easier to comprehend. One suggested reason for this claim is related to the kinds of processes which are employed by texts with different genres; expository texts, for example, apply individual item processing, whereas narrative texts apply relational processing; thus, texts in the narrative genre are thought to encompass considerable related concepts which cause key propositions to be rehearsed frequently for better comprehension (Yoshida, 2012). The results also revealed that the two groups were never inaccurate in assessing fact/opinion subskill in four genres. Owing to the fact that diagnostic assessment throughout the course improved the reading comprehension of the learners, this result supports previous findings which traced a remarkable correlation between learners' mastery in distinguishing fact/opinion in texts and their reading comprehension (Ghahraki & Sharifian, 2005).

5. Conclusion and Implications

What motivated the current study was twofold: first, the lack of research in the area of diagnostic reading assessment which accounts for various stakeholders' involvement, and second, the lack of diagnostic tests and scales in this domain. Therefore, the present study was arranged to develop a number of diagnostic tests and a checklist to examine the effect of EFL learners' diagnostic assessments on reading comprehension.

Although in this research the two groups did not differ considerably in promoting their reading comprehension or in the accuracy of their assessments, the application of diagnostic self-and peer- assessment along with instructor-assessment throughout the course helped the learners find awareness regarding their weak points in reading comprehension and eventually they had improvement in this skill. Admitting that diagnostic assessment is a demanding job even for professionals, contrary to the expectations, the participants of this study were quite successful in performing this task. These findings can offer evidence for implementing diagnostic assessment via learner engagement in order to provide EFL students with suitable training, feedback, and appropriate remedial instruction.

The application of diagnostic assessment in EFL instruction bears both theoretical and pedagogical implications. According to the findings, developing purpose-built diagnostic tests and dividing the labor of diagnosis among the learners could substantially empower the participants with a diagnostic vision upon which they could find awareness regarding their own reading strengths and weaknesses. Since this advocates the principles of diagnostic assessment theory (Alderson et al., 2015), there is scope to advance the theoretical aspect of diagnostic assessment. This research has some pedagogical implications for language teaching and assessment. EFL teachers and administrators can apply diagnostic information to improve the EFL reading curricula in general and develop appropriate lesson plans based on learners' weaknesses in particular (Kim, 2015). Test developers can alter diagnostic information to some practical standards, which can contribute to modifying the teaching and learning pedagogy (Javidanmeh & Anani Sarab, 2019).

Although the results of the present research were supported empirically, they must be treated with caution for some reasons. First, the results are limited to the specific data employed in the current study; for locating the reading difficulties of different learners, all the required diagnostic steps should be taken. Second, the findings of this study would be more generalizable if they were confirmed in replication, focusing on various subskill types at different proficiency levels. Moreover, the way diagnostic information ought to be used, remains an open

question in SFL teaching contexts (Liu, 2014); therefore, subsequent research is required to investigate the actual application of diagnostic information in instruction.

Acknowledgments

We would like to thank the participants who agreed to take part in this study.

References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.
- Alderson, J. C. (2010). Cognitive diagnosis and Q-Matrices in language assessment”: A commentary. *Language Assessment Quarterly*, 7(10), 96-103. <https://doi.org/10.1080/15434300903426748>
- Alderson, J. C., Clapham, C. M., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Alderson, J. C., & Huhta, A. (2005). The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing*, 22(3), 301- 320. <https://doi.org/10.1191/0265532205lt310oa>
- Alderson, J. C., & Huhta, A. (2011). Can research into the diagnostic testing of reading in a second or foreign language contribute to SLA research? In L. Roberts, G. Pallotti, & C. Bettoni (Eds.), *EUROSLA Yearbook* (pp. 30–52). Cambridge University Press. <https://doi.org/10.1075/eurosla.11.04ald>
- Alderson, J. C., Brunfaut, T., & Harding, L. (2015). Towards a theory of diagnosis in second and Foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics*, 36(2), 236- 260. <https://doi.org/10.1093/applin/amt046>
- Alexander, L. G. (1968). *For and against*. Longman Group Ltd
- Aryadoust, V. (2017). An integrated cognitive theory of comprehension. *International Journal of Listening*, 33(2), 71-100. <https://doi.org/10.1080/10904018.2017.1397519>
- Ashton, K. (2014). Using self-assessment to compare learners’ reading proficiency in a multilingual assessment framework, *System* 42, 105–119. <https://doi.org/10.1016/j.system.2013.11.006>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: A comparative study of their effect on writing performance and rating accuracy. *IJAL*, 13(1), 23- 45.
- Butler, Y. G. (2018). The role of context in young learners’ processes for responding

- to self-assessment items. *The Modern Language Journal*, 102(1), 1-20. <https://doi.org/10.1111/modl.12459>
- Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27(1) 5–31. <https://doi.org/10.1177/0265532209346370>
- Carrell, P. L., & Connor, U. (1991). Reading and writing descriptive and persuasive texts. *Modern Language Journal*, 75(3), 314-324. <https://doi.org/10.1111/j.1540-4781.1991.tbo5361.x>
- Chen, H., & Chen, J. (2016). Retrofitting non-cognitive-diagnostic reading assessment under the generalized DINA model framework. *Language Assessment Quarterly*, 13 (3), 218-230. <https://doi.org/10.1080/15434303.2016.1210610>
- Chen, H., Cai, Y., & Torre, J. (2023). Investigating second language (L2) reading subskill associations: A cognitive diagnosis approach. *Language assessment quarterly*, 20(2), 166-189. <https://doi.org/10.1080/15434303.2022.2140050>
- Davis, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, 9(3), 185-197.
- Denis, D. J. (2021). *Applied univariate, bivariate, and multivariate statistics: Understanding statistics for social and natural scientists, with applications in SPSS and R*. <https://onlinelibrary.wiley.com>
- DuBravac, S., & Dalle, M. (2002). Reader question formation as a tool for measuring comprehension: Narrative and expository textual inferences in a second language. *Journal of Research in Reading*, 25(2): 217-231. <https://doi.org/10.1111/1467-9817.00170>
- Elahi, M. S. (2016). Assessing and improving general English university students' main sub-skills of reading comprehension: A case of university of Bojnord. *Sino-US English Teaching*, 13 (4), 245-260. <https://doi.org/10.17265/1539-8072/2016.04.002>
- Esfandiari, R., & Jafari, H. (2021). Morphological complexity across descriptive expository, and narrative text types in Iranian lower- intermediate language learners. *Issue in Language Teaching*, 10(1), 237-267. <https://doi.org/10.22054/ilt.2021.59736.580>
- Farhady, H. & Daftarifard, P. (2006). On the scalability of the components of the

- reading comprehension ability: A progress report. In H. Farhady (Ed.), *Twenty-five years of living with applied linguistics: Collection of articles* (pp. 189-204). Rahnama publisher.
- Field, A. (2018). *Discovering statistics using IBM SPSS, statistics for statistics*. (5th ed.). SAGE Publications.
- Finkbeiner, C., & Schluer, J. (2017). Developing prospective teachers' diagnostic skills through collaborative video analysis: Focus on L2 reading. *Language Awareness, 26*(4), 282-303. <https://doi.org/10.1080/09658416.2017.1415916>
- Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading, 10*(3), 323–330. https://doi.org/10.1207/s1532799xssr1003_7
- George, D., & Mallery, P. (2020). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.
- Ghahraki, S., & Sharifian, F. (2005). The relationship between overall reading comprehension and determination of fact/ opinion in L2. *The Reading Matrix, 5*(1), 36-46.
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*(4), 126-135. <https://doi.org/10.1080/19388076709556976>
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Gray, M. J. (1987). Comprehension: Process or product? *Reading Horizons: A Journal of Literacy and Language Arts, 27* (2), 146-152.
- Han, C., (2018). A longitudinal quantitative investigation into the concurrent validity of self and peer assessment applied to English- Chinese bi-directional interpretation in an undergraduate interpreting course. *Studies in Educational Evaluation, 1*-10. <https://doi.org/10.1016/j.stueduc.2018.01.001>
- Han, C., & Riazi, M. (2018). The accuracy of student self-assessments of English-Chinese bidirectional interpretation: a longitudinal quantitative study. *Assessment and Evaluation in Higher Education, 43*(3), 386-398. <https://doi.org/10.1080/02602938.2017.1353062>
- Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on

- diagnostic principles. *Language Testing*, 32(3), 317- 336. <https://doi.org/10.1177/0265532214564505>
- Jang, E. E. (2005). *A validity narrative: Effects of reading skills diagnosis on teaching and learning in the context of NG TOEFL*. (Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign). <https://www.researchgate.net/publication/33746641>
- Jang, E. E. (2009). Demystifying a Q-Matrix for making diagnostic inferences about L2 reading skills. *Language Assessment Quarterly*, 6(3), 210-238. <https://doi.org/10.1080/15434300903071817>
- Javidanmehr, Z., & Anani Sarab, M. R. (2019). Retrofitting non- diagnostic reading comprehension assessment: Application of the G- DINA model to a high stakes reading comprehension test. *Language Assessment Quarterly*, 16(3), 294-311. <https://doi.org/10.1080/15434303.2019.1654479>
- Karakoc, A. I. (2019). Reading and listening comprehension Subskills: The match between theory, course books, and language proficiency tests. *Advances in Language and Literary Studies*, 10(4), 166-17. <https://doi.org/10.7575/aiac.all.v.10n.4p.166>
- Kim, A. (2015). Exploring ways to provide diagnostic feedback with an ESL placement test: Cognitive diagnostic assessment of L2 reading ability. *Language Testing*, 32(2) 227-258. <https://doi.org/10.1177/0265532214558457>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95 (2), 163–182. <https://doi.org/10.1037/0033-295x.95.2.163>
- Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes*, 39(2-3), 125-128. <https://doi.org/10.1080/0163853X.2005.9651676>
- Lee, Y.-W., & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, 6(3), 239-263. <https://doi.org/10.1080/15434300903079562>
- Liu, H. H.-T. (2014). The conceptualization and operationalization of diagnostic testing in second and foreign language assessment. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 14, (1), 1-12. <https://doi.org/10.7916/D84F23B7>

- Lu, L. (2018). An analysis of peer-assessment in Chinese as a second language classroom presentation. *Chinese Language Teaching Methodology and Technology*, 1(3), 18-30.
- Ma, W., & Winke, P. (2019). Self-assessment: How reliable is it in assessing oral proficiency over time? *Foreign Language Annals*, 52(1), 66-86. <https://doi.org/10.1111/flan.12379>
- Mackey, D., & Savage, A. (2010). *Read this 2*. Cambridge University Press.
- Myers, S. S. (1991) Performance in reading comprehension -product or process? *Educational Review*, 43(3), 257-272. <https://doi.org/10.1080/0013191910430302>
- Nikmard, F., & Tavassoli, K. (2020). The effect of diagnostic assessment on EFL learners' performance on selective and productive reading tasks. *Journal of Modern Research in English Language Studies* 7(1), 79-104. <https://doi.org/10.30479/jmrels.2019.11562.1437>
- Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2nd ed.). Macmillan Heinemann.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13. <https://doi.org/10.1177/026553228900600103>
- Paleczek, L., Seifert, S., Schwab, S., & Gasteiger-Klicpera, B. (2015). Assessing reading and spelling abilities from three different angles—correlations between test scores, teachers' assessment and children's self-assessments in L1 and L2 children. *Procedia - Social and Behavioral Sciences*, 174, 2200 – 2210. <https://doi.org/10.1016/j.sbspro.2015.01.876>
- Paris, S.G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychology*, 36(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4
- Poulaki, S., Dowlatabadi, H., Ahmadian, M., & Yazdani, H. (2020). Dynamic assessment: A diagnostic tool to capture academic second language reading difficulties in the IELTS context. *Journal of Modern Research in English Language Studies* 7(4), 55-87. <https://doi.org/10.30479/jmrels.2020.11998.1494>
- Ravand, H. (2016). Application of a cognitive diagnostic model to a high-stakes reading comprehension test. *Journal of Psychoeducational Assessment*,

- 34(8), 782–799. [https://doi.org/ 10.1177/0734282915623053](https://doi.org/10.1177/0734282915623053)
- Rost, D. H. (1993). Assessing different components of reading comprehension: Fact or fiction? *Language Testing*, 10(1), 79–92. [https://doi.org/ 10.1177/026553229301000105](https://doi.org/10.1177/026553229301000105)
- Şahin, A. (2013). The effect of text types on reading comprehension. *Mevlana International Journal of Education*, 3(2), 57-67. [https://doi.org/ 10.13054/mije.13.27.3.2](https://doi.org/10.13054/mije.13.27.3.2)
- Sawaki, Y., Kim, H. J., & Gentile, C. (2009). Q-matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly*, 6(3), 190–209. <https://doi.org/10.1080/15434300902801917>
- Shelfbline, J., & Shiel, G. (1990). Pre service teachers' schemata for a diagnostic framework in Reading. *Reading Research and Instruction*, 30(1), 30-43. [https://doi.org/ 10.1080/19388079009558031](https://doi.org/10.1080/19388079009558031)
- Song, M.Y. (2008). Do divisible subskills exist in second language (L2) comprehension? A structural equation modeling approach. *Language Testing*, 25(4), 435–464. <https://doi.org/10.1177/0265532208094272>
- Spearritt, D (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly*, 8(1), 92-111. [https://doi.org/ 10.1002/j.2333-8504.1972.tb00192.x](https://doi.org/10.1002/j.2333-8504.1972.tb00192.x)
- Tengberg, M. (2018): Validation of sub-constructs in reading comprehension tests using teachers' classification of cognitive targets. *Language Assessment Quarterly*, 15(2), 169-182. <https://doi.org/10.1080/15434303.2018.1448820>
- Toledo, P. F. (2005). Genre analysis and reading of English as a foreign language: Genre schemata beyond text typologies. *Journal of Pragmatics*, 37, 1059–1079. [https://doi.org/ 10.1016/j.pragma.2005.01.002](https://doi.org/10.1016/j.pragma.2005.01.002)
- Toprak, T. E., & Cakir, A. (2020). Examining the L2 reading comprehension ability of adult ELLs: Developing a diagnostic test within the cognitive diagnostic assessment framework. *Language Testing*, 1–26. [https://doi.org/ 10.1177/0265532220941470](https://doi.org/10.1177/0265532220941470)
- Urmston, A., Raquel, M., & Tsang, C. (2013). Diagnostic testing of Hong Kong tertiary students' English language proficiency: The development and validation of DELTA. *Hong Kong Journal of Applied Linguistics* 14(2), 60–82.

- Vacca, R. T. (1980). A study of holistic and subskill instructional approaches to reading comprehension. *Journal of Reading*, 23(6), 512–518.
- Yang, S. (2021). Constructing diagnostic reading assessment instruments for low-level Chinese as second language learners. *Reading in a Foreign Language*, 33(2), 212–237.
- Yin, J. (2018). A review on researches of “genre-based teaching approaches” in recent 20 years in China. *International Journal of Secondary Education*, 6(1), 16-23. [https://doi.org/ 10.11648/j.ijssedu.20180601.14](https://doi.org/10.11648/j.ijssedu.20180601.14)
- Yoshida, M. (2012). The interplay of processing task, text type, and proficiency in L2 reading. *Reading in a Foreign Language*, 24 (1), 1-29. [https://doi.org/ 10125/66671](https://doi.org/10125/66671)
- Zhou, L., & Siriyothin, P. (2011). Effects of text types on advanced EFL learners’ reading comprehension. *Journal of Language and Culture*, 30(2), 45-66.

Appendix A: Diagnostic Reading Comprehension Checklist

Reading Comprehension Diagnostic Assessment Checklist		
Specify the questions that target the subskills of <i>Main idea and supporting details</i> in the test:		
Diagnostic Instructor Score:	Diagnostic Self / Peer Score:	
Description	Scoring	Reading Comprehension Ability level
The student can distinguish <i>main idea/ supporting idea</i> perfectly	Five questions answered correctly	Excellent 5
The student is good at distinguishing <i>main idea/ supporting idea</i>	Four questions answered correctly	Good 4
The student can distinguish <i>main idea/ supporting idea</i> to some extent	Three questions answered correctly	Fair 3
The student understands the basic meaning of the text	Two questions answered correctly	Poor 2
The student is unable to comprehend the basic meaning of the text	One/no question(s) answered correctly	Non-reader 1 or 0
Reading Comprehension Instructor Score:	Reading Comprehension Self / Peer Score:	

Appendix B: Table of Specifications for Diagnostic Reading Comprehension Tests

Text Topic	Objective	Type of Genre	Type of Test Items	Cognitive Level	No of Items	Total Points
Miracle on the Hudson	Identifying Cause / Effect	Narrative	Multiple – choice Short answer	Analyzing	10	10 points
Gentle Giant	Identifying Cause / Effect	Descriptive	Multiple – choice Short answer	Analyzing	10	10 points
Death by Internet	Identifying Cause / Effect	Expository	Multiple – choice Short answer	Analyzing	10	10 points
Fashion	Identifying Cause / Effect	Argumentative	Multiple – choice Short answer	Analyzing	10	10 points
Certain Death	Identifying Fact / Opinion	Narrative	Multiple – choice Short answer	Analyzing	10	10 points
Kiwi	Identifying Fact / Opinion	Descriptive	Multiple – choice Short answer	Analyzing	10	10 points
Power of the Mind	Identifying Fact / Opinion	Expository	Multiple – choice Short answer	Analyzing	10	10 points
The Second-class Citizen	Identifying Fact / Opinion	Argumentative	Multiple – choice Short answer	Analyzing	10	10 points
Fighting Disease with Disease	Identifying Main idea/ Supporting details	Narrative	Multiple – choice Short answer	Analyzing	10	10 points
The Fifth Taste	Identifying Main idea / Supporting details	Descriptive	Multiple – choice Short answer	Analyzing	10	10 points
Eat Less, Live Longer	Identifying Main idea/ Supporting details	Expository	Multiple – choice Short answer	Analyzing	10	10 points
Smoking	Identifying Main idea / Supporting details	Argumentative	Multiple – choice Short answer	Analyzing	10	10 points