



Content list available at <http://ijltr.urmia.ac.ir>

*Iranian Journal
of
Language Teaching Research*
ORIGINAL ARTICLE



Urmia University

Classroom-based Diagnostic Assessment Practices of EFL Instructors

Hossein Farhady ^{a,*}, Merve Selcuk ^b

^a *Yeditepe University, Turkey*

^b *Altinbas University, Turkey*

ABSTRACT

Diagnostic assessment (DA) in language classrooms is reported to have a substantial role in providing immediate and quality feedback for remedial work that would improve learning (Alderson, Brunfaut & Harding; 2014; Hattie & Timperley, 2007; Huhta, 2008; Jang & Wagner, 2013). However, applying DA in class remains a challenge for many practitioners for at least two reasons: First, there is no clear conceptualization of DA in the field of second language acquisition (Alderson, 2005; Alderson et al., 2015; Knoch & Macqueen, 2017). Despite advancements in teacher education (TE) programs, they mostly provide only a handful of courses to equip teacher candidates with necessary classroom assessment knowledge (Popham, 2011). To address the issue, this study attempted to explore DA practices and procedures of experienced university EFL instructors with the hope of developing a sound and empirically supported framework for classroom-based DA in the EFL context. The data came from 17 instructors with an MA in TEFL and related fields with more than 5 years of teaching experience. They received scenarios presenting certain problems, asked to explain what procedures and processes they would employ to diagnose the problem, and what they would do as remedial instruction. Based on the findings, we propose a data-driven framework for classroom-based DA. The paper also provides a discussion of the need to incorporate theoretical and practical dimensions of DA in pre-service and in-service TE programs.

Keywords: diagnostic assessment; classroom-based assessment; teacher education

© Urmia University Press

ARTICLE HISTORY

Received: 25 Mar. 2021

Revised version received: 4 May 2022


Accepted: 5 June 2022

Available online: 1 July 2022

* Corresponding author: Yeditepe University, Istanbul, Turkey

Email address: hossein.farhady@yeditepe.edu.tr

© Urmia University Press

 10.30466/ijltr.2022.121184

Introduction

Teachers usually assume responsibility for both enhancing and measuring learning in classroom settings (Rea-Dickins, 2007) by implementing different methods of teaching and assessment. Despite some concerns about teacher-based assessments (Brindley, 2001; Rea-Dickins & Gardner, 2000), classroom-based assessment, either summative or formative, provides an opportunity to shape and improve learning (Huhta, 2008) by blending teaching with assessment.

Teachers use a variety of summative assessment instruments in the classroom to ensure accountability. Nevertheless, though such tests may yield useful information about the attainment level of students, they may not be sufficient in producing detailed and necessary information about individual student's progress. Therefore, teachers often conduct a formative assessment to obtain information about student progress so that they can modify instructional approaches, teaching materials, and academic support in the course of instruction. One of the purposes of formative tests is usually diagnostic that "is used to discover a learner's specific strengths or weaknesses" (ALTE, 1998, p. 142). Of course, there stand few diagnostic language tests such as online diagnostic assessment system (DIALANG), Diagnostic English Language Needs Assessment (DELNA), and Diagnostic English Language Tracking Assessment (DELTA). Yet, teachers need more tailor-made assessment tools to diagnose students' learning problems and to intervene in learning where problems arise (Edelenbos & Kubanek-German, 2004). More importantly, such teacher-made tests enable teachers to "guide learners on where they need to improve and give feedback to learners" (Alderson, 2005).

Foundations of Diagnostic Assessment (DA)

The major purpose of DA is to provide specific information about individual weaknesses and proceed with remedial teaching (Alderson, 2005). Simply put, "diagnosis is the teacher's concern of what has gone wrong" (Kunnan & Jang, 2009, p. 6) to identify students' weaknesses. Teachers could accomplish this objective through formal tests or informal assessments for taking future action whether by providing additional teaching or implementing other strategies (ALTE, 1998; Lee, 2015). Further, the major objective of DA is to provide information on a test taker's strengths and weaknesses to utilize the data to provide feedback for the improvement of learning (Jang & Wagner, 2013). By providing quality feedback and remedial instruction (Lee, 2015), DA serves to enhance learning and promote teaching practices. However, despite its significance, DA studies have received relatively little attention in second or foreign language teaching contexts. Alderson (2005) claims that DA is one of the most discussed topics in language testing but with no results regarding its implementation or even its definition. To date, there have been only a handful of recent studies in the second and foreign language (SFL) that have made efforts to address the need to operationalize DA and its principles in the SFL context (Nickmard & Tavassoli, 2020).

Teachers often perform DA by using assessment instruments or by on-site observation with human judgment. They may also carry out DA consciously or incidentally while student learning is in process. As Blood (2011) states, "informal diagnostic assessment occurs regularly in the form of student questioning, explanation, and the provision of written feedback on quizzes, tests, and written work" (p. 57). In addition to informal assessment (Brindley, 2001; Edelenbos & Kubanek, 2004; Hsu, 2005; Yang, 2008), any sort of formative classroom assessment instrument defined in the literature, from formal pen and paper tests to alternative assessment (Brindley, 2001; Brown & Hudson, 1998; Gattullo, 2000) could also perform diagnostic functions. These assessment procedures are usually structured in diverse forms like "observation, recycling of work, diagnostic testing, learner self-assessment, ad hoc tests" (Brindley, 2001, p.127), portfolios, teacher-constructed performance tasks (Yang, 2008, p.144), projects (Brown & Hudson, 1998; Işık, 2020),

homework and oral exams (Işık, 2020), collection of data samples, different forms of corrective feedback, informal questioning, correcting, judging, rewarding, clarifying, task criteria, metacognitive questioning, peer assessment, “oral exams, role play, musical presentations, student-teacher conferences, and performance tasks” (Brindley, 2001). All these possibilities boil down to the point that teachers need to use their judgment during informal assessments to interpret what they have obtained to assist student learning. However, it is important to note that what makes the data coming from any of the above-mentioned forms of assessment diagnostic is not the test but the user of the test who diagnoses and interprets the data (Alderson et al, 2014).

However, limited literature on the concept of DA reveals that a major requirement is a qualified diagnostician to make sound decisions to notice the weaknesses and take remedial action (Edelenbos & Kubanek-German, 2004; Huhta, 2008). While traditionally DA has often been performed by judgmental abilities of teachers regarding their students’ performances, (Bates & Nettelbeck, 2001; Feinberg & Shapiro, 2003; Helmke & Schrader, 1987), Edelenbos and Kubanek-German (2004) suggested shifting the focus from product to process. The researchers have usually identified teachers’ diagnostic assessment activities using arbitrary time sampling methods and then proposed the notion of diagnostic competence. Their definition of diagnostic competence, the prerequisite for DA, is "observational and interpretative competence shown by teachers during classroom teaching" (p. 277). By focusing on learner behavior rather than achievement, they operationalized DA as "the ability to interpret students' foreign language growth, to skillfully deal with the assessment material and to provide students with appropriate help in response to this diagnosis" (p. 260). Quoting Edelenbos and Kubanek-German who presuppose that diagnostic competence is the initial requirement of DA, Alderson et al., (2014) pushed it further and identified the components of DA by formulating the steps involved in the cycle of the DA. Through a qualitative study, they examined DA processes by collecting data from diverse fields of professions and vocations using semi-structured interviews and mapped out a framework for diagnosis in SFL assessment. They identified five tentative principles of DA as follows:

1. *It is not the test that diagnoses; it is the user of the test. This will often be a teacher, who will need to make an informed diagnosis (...), and then utilize a range of assessment tools as well as their informed judgment and the expertise of others, and finally, form a decision about the nature of a specific problem. (...)*
2. *(...) Instruments themselves should be designed to be user-friendly, targeted, discrete, and efficient to assist the teacher in making a diagnosis. (...) and should generate rich and detailed feedback for the test-taker.*
- 3 *(...) DA process should include diverse stakeholder views, (...)*
4. *(...) DA should ideally be embedded within a system that allows for all four diagnostic stages: (1) listening/observation, (2) initial assessment, (3) use of tools, tests, expert help, and (4) decision-making (...)*
5. *(...) DA should relate, if at all possible, to some future treatment. (...)* (Alderson et al., 2014, pp. 20-21)

The fourth principle that involves the stages of the process of DA and starts with listening/observation and leads to decision making is further explored in their subsequent research (Harding et al., 2015). It provides a working framework for how these stages might be followed by a teacher. For example, in the listening/observation stage, the teacher identifies that a student is experiencing particular difficulty in a specific skill and takes action to develop a deeper understanding of the student's problem. In the initial assessment stage, the teacher makes a

hypothesis about the problem based on the data obtained from stage 1 and utilizes certain instruments (a test, a task, etc.) to test the hypothesis in stage 3, and finally makes a decision. If the solution does not seem to be working, then the teacher goes back to stage 1 and starts over. Despite its practical appeal, no empirical evidence is available to support the functionality of the model (Harding et al., 2015). Figure 1 below is a schematic representation of these processes.

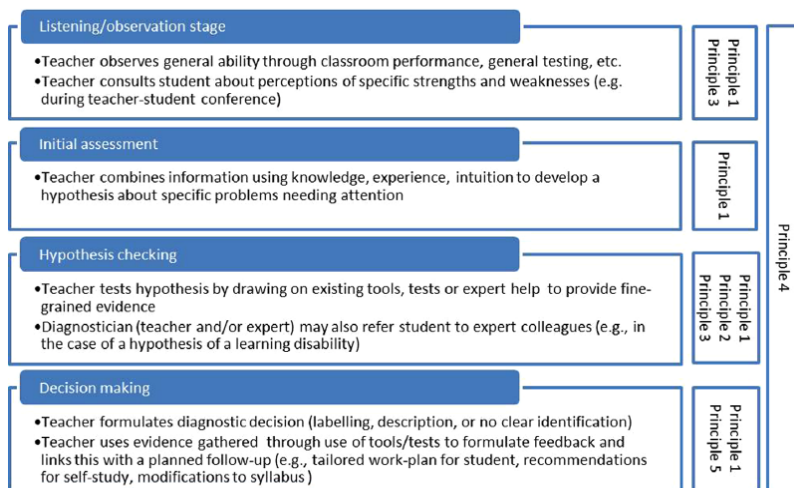


Figure 1. Diagnostic process (Harding et al., 2015, p. 319)

The motivation and rationale for this project stemmed from concerns about how teachers could perform DA in class since one of the principles of DA requires a skilled diagnostician. The findings of Alderson et al., (2014) and Harding et al., (2015) have been an incentive to design this study to discover how experienced teachers shape their DA practices or criteria in practice in SFL assessment. Based on the findings, then we could provide a workable framework to assist the teacher population. The context of this study is the tertiary level EFL instruction in Turkey and the focus is on the academic reading skills since it is the most significant skill at his level. Further, including other skills would have certainly been beyond the scope of a single study. More specifically, the present study sought to answer the following question:

What procedures do experienced EFL instructors follow in DA processes regarding reading skills?

DA and Reading Skills in EFL Contexts

Reading comprehension is an important skill in educational life, as every student needs to develop this ability to read texts in academic settings. As a result, there is a burgeoning literature on reading ability in diverse language-related domains such as psychology, education, and applied linguistics (Urquhart & Weir, 2014). Grabe and Stoller (2019) explain the importance of reading by emphasizing how L2 reading ability, especially in the case of English, is becoming increasingly vital, as most of the world experience multilingual settings. Reading comprehension is a complicated task requiring the orchestration of several components of linguistic knowledge such as vocabulary and grammar, and some cognitive skills such as referencing, inferencing, and so on.

However, despite the abundant literature on reading in EFL/ESL context, diagnostic assessment of reading has not received due attention. Only a handful of researchers explored the impact of DA on learners' achievement in the EFL context. For instance, Nickmard and Tavassoli (2020) investigated the impact of DA on EFL learners' performance on selective and productive reading comprehension skills. Diagnostic assessment group participants (experimental group) received feedback pursuing the steps of DA suggested by Harding, Alderson and Brunfaut (2015). The class teacher observed the weaknesses of learners using the KET (Key English Test) (listening/observation stage) and then conducted feedback sessions to identify sources of difficulties that the learners experienced (initial assessment). Next, the teacher checked the hypothesis using different tools (hypothesis checking). Finally, the learners received detailed feedback based on the set of assumptions the teacher made in the previous stage (decision-making). They reported a significant positive impact of the intervention on performing reading comprehension tasks. Besides, DA is not only beneficial in improving reading comprehension, but also in enhancing learning other skills. For example, Kazemi and Tavassoli (2020) using pre-and post-test design, investigated the impact of DA on the speaking skills of 133 adult English Language Learners. Participants who received diagnostic feedback for three formative speaking tests demonstrated significant improvement in their speaking ability level compared to that of the comparison group. In addition to reading and speaking skills, Mazloomi and Khabiri (2016) investigated the effectiveness of DA in developing writing skills and reported that appropriate feedback on students' diagnostic self-assessment of their writings improved their performance.

Method

This study aimed to explore what teachers do through the 4 stages of DA, namely, "listening/observation, initial assessment, use of tools, tests and/or expert help, and finally decision making" (Harding et al., 2015). We attempted the principles of the qualitative research method to obtain detailed and rich information about the above-mentioned research question.

Participants

A total of 17 EFL instructors, 12 females and 5 males, from 13 different school contexts responded to scenario-based questions. They contributed to this study following convenience sampling and all had considerable expertise and experience in EFL teaching at intensive pre-faculty language courses at the tertiary level in Istanbul, Turkey. These teachers shared the background characteristics of having at least 5 years of professional experience in English Language Teaching, Translation Studies, or English Language Literature and all holding at least an MA degree or teaching certificates in language teaching. The average age of the participants was 34 with a range of 28 to 48 years. The average teaching experience in years was 10.3 with a range of 5-22 years. Table 1 is a summary of information about the participants. All participants who agreed to assist us in conducting the study signed a consent form. The names of the participants referred to here are pseudonyms.

Table 1
Information on the Participants

Gender	Degree	Place of teaching	Years of experience	Major
F (12)	Ph.D. (8)	University Prep School (16)	btw 5-7 yrs. (7)	ELT (12)
M (5)	MA (7) Certificates DELTA/CELTA (2)	Other (1)	btw 8-14 yrs. (3) more than 15 years, (7)	Translation, ELL (4) Other, (1)

Instrumentation

We collected the data by running three scenarios with open-ended questions developed for this study. The structure of each scenario required participants to provide information on the four stages of DA suggested by Harding et.al., (2015) namely, listening/observation, initial assessment, hypothesis checking, and decision making. Using scenario-based instruments has two advantages over other data collection procedures such as observations and questionnaires. First, teacher decisions during instruction may not be apparent during observation since the student behavior that requires the diagnostic assessment process may not occur during the time of observation. Second, scenarios would provide opportunities for observing the developmental pattern of teacher reactions to the same problems on repeated occasions.

The scenarios included case descriptions of English language learners' problems that may result in low achievement in language learning (Table 2). Each scenario focused on a specific reading deficiency reported in the ESL/EFL reading literature. While the scenarios represent the listening/observation stage of DA, the follow-up questions are intended to elicit information on how they implemented other stages of DA processes. Each scenario is followed by the same seven open-ended questions. Two experienced EFL teachers reviewed the questions to prevent any language or content ambiguity. Table 3 presents the reading problems addressed in each scenario and open-ended questions following them. The participants needed to figure out the underlying reason behind the problems and suggest solutions by justifying their remedies. The previous studies did not indicate how the researchers deployed the data obtained in the final stage of DA in SFL contexts. That is why we have included this stage here as well to provide further information and suggest further research.

Table 2
Scenarios and Problems the Scenarios Illustrate

Scenarios	Problems they exemplify
Scenario 1:	The students do not apply reading strategies to solve reading comprehension problems.
Scenario 2:	The student cannot create a text-level representation.
Scenario 3:	The students cannot perform lower-level reading skills due to orthographic differences between their L1 and L2.

Table 3
Reading Problems Represented by Each Scenario and Open-ended Questions Following Them

Questions	stages
1. What would you do to obtain more in-depth information about the students' detected problem?	1. Listening/Observation
2. What would be your initial hypothesis about the errors or mistakes?	2. Initial assessment
3. How would you test your hypothesis? (What specific ways would you follow/what specific tasks would you use to make sure that your hypothesis is correct?)	3. Hypothesis checking/ use of tools, tests, expert help, class discussion
4a. What would you do after diagnosing the problem? (Assume that your first diagnosis is correct and note down all possible ways you could follow)	4. Decision-making
4b. What kind of intervention would you employ? (If you decided to intervene)	
4c. How would you check the effect of intervention?	
4d. How would you approach the students who did not benefit from the intervention?	

Procedure

Before delivering the questionnaire to the participants, the researchers gave them information about the nature and general aim of the study. They needed to provide as detailed written responses as they could to the problems addressed in the scenarios. They also needed to explain in detail what procedures they would employ to diagnose the problem and what they would do as remedial instruction. To avoid receiving hasty responses from the participants, they were given two weeks to respond and return the scenarios.

Data Analysis

To analyze the data, the researchers coded the answers to scenarios independently. The analyses drew on the theoretical framework of diagnostic assessment proposed by Alderson et al., (2014) to formulate initial coding categories. To this end, we identified three components of diagnostic assessment as the themes and explored classroom applications referring to these themes throughout the data. The themes included “Means of diagnosis (M)”, “Diagnostic procedures (D)”, and “Treatment (I)” or follow-up. To assure the accuracy of the codes, we checked the data several times and identified the keywords for the components. Then the researchers came together to discuss the disagreements to reach a consensus on the codes.

Results

Although the present study examined the implementation procedures of expert EFL instructors in the DA process regarding reading skills, there emerged some insightful findings of teachers' practices. First, the preliminary interpretation of these observations indicated that almost all participants consistently applied the five principles of DA identified by Alderson et al., (2014).

Second, the data revealed that teachers utilized multiple sources of information to develop a more comprehensive picture of the problem in the DA process during the stages of data collection for initial assessment and hypothesis checking. For problems presented in the scenarios, the informants offered multiple solutions to the same problem, which refers to the creativity of teachers. It also indicates how sensitive they are to the diverse needs of learners. The following quotations corresponding to themes regarding detection of students' problems illustrate how the participant teachers tried to understand learner problems by looking at the issue from different perspectives that led them to generate various solutions.

“The students' reluctance to read the texts and answer comprehension questions and their hesitancy in participating in the reading class could result from their lack of motivation which usually grows out of lack of confidence. Lack of enough vocabulary to understand a text and inability to express themselves in English might lead to this lack of confidence. Another reason for their reluctance in reading the texts or participating in the class might be just they do not find the texts interesting enough. To get more in-depth information about the causes of the problem, I would give the class a short questionnaire” Esin

“I would administer think-aloud protocols and one-on-one interviews.” Merve

“The first thing coming to my mind would be this: I would think that the student's reading fluency needs to be improved. Besides, s/he might lack the ability to use reading strategies effectively.” Merve

The analysis of respondents' answers to the scenarios reveals that they come up with innovative tasks, procedures, and solutions to the problems presented to them by paying attention to contextual and affective factors.

"I would assume that since Arabic has a different orthography, they have difficulty in reading in English. (...) Since their language is the source of the problem, they need more reading practice in English. So, I would assign them weekly readings on their interest." Esin

An additional finding emerging from scenario answers is that, if their plan does not work at the end of the DA cycle, they digress to one of the earlier stages of the cycle to readdress the problem.

"I would question my hypothesis and change my strategy if needed." Esin

"For those that did not show improvement after all the above activities were implemented, I would try different things. Again, assuming that they are pre-intermediate level, and the material is appropriate for them, I would try to come at the problem from a different direction. Perhaps that student would be more motivated if his partner in the classroom was more encouraging. Often, a new seating arrangement can have far-reaching effects on students. Perhaps his interests need to be addressed (and were not in the previous attempts)." Neville

The following tables summarize participants' solutions to the problems presented to them.

Stage 1. Listening/Observation First, the teachers explained how they would react to three problems presented to them. The following frequency table displays all answers of the participants dealing with each question.

Q1. What would you do to obtain more in-depth information about the student's detected problem?

For problems presented, the informants offered multiple solutions to the same problem, which refers to the creativity of teachers. Table 4 displays the frequency of answers to question 1.

Table 4
Frequency Table of the Participants' Responses to Question 1

<i>The frequency of solutions I would</i>	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
check the material appropriacy (proficiency level, vocabulary, content)	4	4	1
conduct needs analysis	1	1	1
raise students' awareness of the issue	1	-	-
carry out a class survey	1	-	-
ask open-ended questions about how they read the text	1	-	-
prepare a questionnaire based on predicted reasons & relevant literature (e.g. I find reading enjoyable/boring; I know many of the words/do not know many of the words; about the techniques they used; What would you like to read about, etc.)	3	-	-
prepare a diagnostic reading test based on predicted reasons and relevant literature	2	-	-
administer a market diagnostic reading test/placement test	2	2	6
prepare a diagnostic questionnaire based on observed weaknesses	-	-	2
do some trial and error with some different exercises, e.g., Find passages related to their interests, see if the problem persists	2	-	-
give students another reading test with similar exercises and conduct think-aloud protocols/retrospective protocols	2	4	-
interview students individually	7	1	11
give students more reading passages with similar tasks	-	-	-
give students another reading passage with similar tasks and	-	1	-

observe them, take notes to see whether they exhibit the same behavior consistently			
give students another reading test with similar tasks and have a class discussion based on the exercise	3	-	-
ask for an L1 translation of the reading passage	3	1	-
Design tasks (more texts and similar tasks)	-	1	-
Check other top-down skills	-	1	-
Can-do statement checklist	-	7	-
Read aloud	-	-	1
Do trial and error with some tasks	-	1	-
Getting background info about their experience in learning English (learning style, their experience in learning English, which English courses they took, which skills they focused on)	-	-	3
Use more literacy-focused teaching materials to understand what the level of difficulty is	-	-	2
Check students' papers to decide on strong and weak points	-	-	2
	-	-	-

Stage 2. Initial assessment

Next, the participants received some questions about their hypotheses for three scenarios and they provided the suggestions presented in Table 5.

Q2. What would be your initial hypothesis about the errors or mistakes?

Table 5
Initial Hypothesis about the Errors or Mistakes

	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Motivation	8		
Cannot fulfill higher-level processing	1		
Insufficient BICS knowledge (Vocabulary & grammar knowledge)	4		
Lack of reading strategies (guessing word meaning, the main idea)	9	7	
Lack of Reading skills			5
Being exam oriented (reading for the questions vs. reading for the main idea)	1		
Lack of content vocabulary/background knowledge/content is difficult	6	2	
Problems with lower-level skills Limited surface-level reading		7	
They are analytic learners		1	
No critical thinking		1	
Insufficient practice		1	
Mistakes/errors in orthographic control (letter-sound correspondence)			12
Open-ended questions may cause students to make more mistakes			1
Need advanced grammar practice			1
Educational background			1
Language proficiency			1
L1 intervention			7
Lack of reading practice			1

Stage 3. Hypothesis checking/use

In the third stage of DA, the participants were required to list all possible tools they would use to check the hypothesis they made in the second stage of DA. Table 6 shows the frequency of the responses the participants gave to the following question.

Q3. How would you test your hypothesis? (What specific ways, tools, tests, expert help, would you follow and what specific tasks would you use to make sure that your hypothesis is correct?)

Table 6
Frequency table of the answers to the stage of hypothesis testing

<i>Testing hypothesis</i>	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Task and class discussion /activity observation	3		
Think-aloud protocols	3	1	1
Interview individually	3	5	1
An experimental study	3		
Make a survey	4		
Give another similar task/text and observe if the problem persists	1		
Diagnostic Test (Reading speed test 2, Eye tracking 1, Vocab 1)		4	
Action research		1	
Awareness-raising activity		1	
More tasks/activities/questions regarding top-down skills		5	
Give him the easier level of passages with easier tasks and try to observe difficulties		1	
Give a similar task and try to observe the difficulties		1	1
An experimental study (Assuming they have vocab problems, give them an easy and hard text and compare the results + A reading text but with some technical unknown words and another easy task with the same questions to see the difference)		3	
Test-retest		1	
The same test in L1			1
Timed reading exercise			1
Give a text and observe students on task			1
Give the same task as a recording and compare the results			1
Action research			1
Observe students in small groups of 5-6			1
Read aloud			1
Slower reading			1
(Giving students a standardized test to understand their level in the previous stage), analyze the results of each item to understand what type of questions they are struggling with			1
Compare these students' performance with the other same proficiency level students' performance			1
Reading vocab and grammar tests			3
Observe to elicit their lang. prof. (in other skills as well)			2

Stage 4. Decision-making

In the final stage of the DA cycle, the participants needed to identify how they would fix the problem they detected in students' language proficiency and to state if they had any intervention plans. As part of the fourth stage of DA, they also needed to present their solutions for learners who, hypothetically, did not benefit from the intervention they provided. Table 7 presents their responses to the above-mentioned questions.

Q4. What would you do after diagnosing the problem? (Assume that your first diagnosis is correct and note down all possible ways you could follow)

Table 7
The Frequency of the Participants' Responses for Decision-making Stage

	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Instruction (private tutoring)	8		5
Practice & Tasks assignments	9	5	11
Instruction and guidance	6		
One-to-one meeting	3	5	
Design VAKOG-relevant instruction	1		
Make changes, modify the syllabus	4		
Instruction (private tutoring)		4	
Experimental design (two classes)		1	
Extensive reading		5	2
Strategy training		5	
Spelling exercise			5
Passages supported with audio recordings			1
Awareness-raising activities for reading skills			1
Apply Kolb's experiential learning cycle			1

Stage 4. Decision-making

Q5. (If you have decided to intervene) What kind of intervention would you employ?

Table 8
The Frequency of the Participants' Responses to Question 5

Intervention	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Adapt and design tasks & assignments and guide students during/after the exercises (observe)	6	5	2
Extensive reading (periodical, based on students' interests, go over them together for strategy training)	3	2	2
Strategy training	4	3	4
Grammar reinforcement	2		3
Talk to the student, convince him (of the importance of learning strategies) and make a tailored study plan for the student	4		
Give students a reading text in Spanish or Italian and have them find specific info (numbers, names) in a text written in a language they do not know just to show that they can do it!	1		
Give students a degree of choice for the materials and activities	3		
Scaffolding: Similar tasks with easier passages/easier tasks/easier shorter readings to help students gain motivation and confidence and then increase the level of difficulty, gradually difficult texts	1	3	
Give an easy passage in which some words are replaced by pseudowords and show the student she can still grasp the meaning without knowing the words		1	
Peer teaching (activities & strong ones help share their strategies with the weaker ones)		1	
Working on students' writing to integrate other skills in the comprehension process		1	
Read aloud exercises practice (seeing hearing together)			1
Sub-skills training			3
Practice to improve letter/word recognition, decoding skills, and develop literacy			3

Stage 4. Decision-making*Q6. How would you check the effect of the intervention?*

Table 9
The Frequency of the Participants' Responses to Question 6

<i>Checking the effect of an intervention</i>	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Informal assessment (group discussions, observation, asking questions, etc., periodically)	10	7	13
Post-test/re-test	6	9	8
Feedback from analysis of assignments, homework, classwork (reading text and activities)	1	4	5
Self-assessment checklists	1		
Peer assessment checklists	1		
A new text and focus groups on how they arrived at the answers	1		
Focused interviews on the effect of the intervention	3		
Reflection from learners		1	1
Design productive tasks (writing & speaking) with specific criteria and rubrics		3	
The same test in L1		1	
Interview individually		1	
Keep a teacher's diary			1
Compare post-test results with other international students' results			1
Think aloud			1
Aptitude tests			1

Stage 4. Decision-making*Q7. How would you approach the students who did not benefit from the intervention?*

Table 10
The Frequency of the Participants' Responses to Question 7

<i>Solutions for those who did not benefit from intervention</i>	number of answers for scenario 1	number of answers for scenario 2	number of answers for scenario 3
Tutorials remedial classes	1	3	1
Extensive reading	1		
Interview/negotiate with the students	6		8
Interview + a diagnostic reading test	1		
More extra materials, more practice	5	4	2
Reconsider my hypothesis, look at the issue from a different perspective, restart the cycle	1		4
Analyze their learning preferences	1		
Dig more into their background experience	1		
Questionnaire/Survey	2		
Give more time to the students	1		
Extensive reading		1	2
Extensive reading texts with audio			1
Interview/negotiate with the students / ask if they have other weaknesses I have not observed		4	
Find another way/ other intervention methods (go back to stage 3)		4	
Adopt a humanistic approach (Give more time to the students to talk to the student, motivate to study more, ask him to be patient, and keep working with him, encouragement)		3	6
Awareness-raising activities, tasks		1	
Lower-level material scaffolding		1	
Focus on the mistakes and think about the reasons (SLA theories)			1

Discussion

The findings of the current research demonstrated that experienced EFL teachers follow the procedures and stages outlined by Harding et al., (2015) for DA of reading skills. Therefore, we may safely assume that teachers come up with various solutions for each stage of DA. We may also assume that experienced teachers employ diagnosis, feedback, and guidance for future learning to overcome deficiencies as a whole cycle.

The findings also yielded information regarding the principles of the DA process exemplified below.

Principle 1. Alderson (2005) suggests that all tests could be used for diagnostic purposes but requires a skilled diagnostician.

Q3. "As I said, I would give them professionally written evaluations. These exams are proven on an international standard. They have had millions of dollars invested into their accurate assessing abilities, so they can be trusted to give me an accurate read on my students' levels. In addition, I would assess the students in a detailed way. I would not just get results on reading comprehension. Rather, I would get a detailed analysis of what types of questions they struggled with. Can they read for a purpose, main idea, details, and facts, work out the meaning of words they don't know, paraphrase?"
Nurcan

Like other participants of the study, this instructor uses a standardized test but utilizes or exploits the test results for diagnostic purposes only and uses the results in a problem-solving way. As stated by Jang and Wagner (2013), the purpose of giving diagnostic feedback is to indicate the discrepancy between the learner's current level of performance and the desired level of performance or goal. Types, amount, delivery, and impact of diagnostic feedback have been widely explored in ESL/EFL literature (Ferris, 2003; Ferris & Roberts, 2001; Hedgcock & Lefkowitz, 1996; Jang & Wagner, 2003) with promising findings on behalf of learning enhancement. A skilled teacher can use any kind of assessment tool to generate diagnostic feedback. As feedback gets more specific, the learner gets more benefit that ultimately leads to improvement (Goodman & Wood, 2004) and according to Jang and Wagner, the specificity of feedback makes it diagnostic.

Principle 2: As suggested by Harding et al., (2015), instructors design "targeted, purpose-built diagnostic tools, selected from a bank according to purpose" (p. 318). All teachers in our study stated that they design specific tools and instruction tailored activities for individual students. It is important to highlight that designing DA tools for reading comprehension is not easy unlike the ones for the assessment of grammar and vocabulary skills (Purpura, 2014). One reason may be the presence of controversial issues regarding the complex process of reading comprehension that leads to disagreements among scholars, especially on the micro-skills underlying reading skill (Alderson & Lukmani, 1989; Lumley, 1993; Selçuk, 2020).

Principle 3: "Diagnostic assessment process should include diverse stakeholder views, including learners' self-assessments" (Alderson et al., 2014, p. 22). Likewise, our findings reveal that the instructors use can-do statements, self-assessment forms, and interviews to identify learner weaknesses and involve them in the DA process. In this way, the finding is consistent with Oscarson (2014) who suggested the integration of DA and self-assessment tools.

"Often, students are aware of the reasons for the breakdown that is occurring in their reading comprehension." (Neville)

Principle 4: The informants followed all four diagnostic stages.

Principle 5: All participants offered several suggestions for future treatment as envisaged by Alderson et al. (2014)

Conclusion and Implications

DA is relatively a new research area awaiting investigation. As Harding, et al. (2015) stated, the framework, developed by exploring other professional fields, was untested in the field of language assessment. This study was a step in the direction to identify behavioral patterns of teachers while performing DA in the EFL context. More specifically, by collecting data by a scenario-based instrument, the present study revealed that experienced teachers were capable of identifying essential behavioral patterns they need to implement DA stages. Despite complexities inherent in language assessment, it is believed that teachers are capable of doing classroom assessment as Brindley (2001) states, "With experience, many teachers will become skilled judges and observers capable of evaluating the quality of language performances and making fine-grained diagnoses of learners' difficulties" (p. 127). However, studies indicate that little amount of class time is spent on diagnostic assessment (Edelenbos & Kubanek-German, 2004). Additionally, considering concerns relating to the lack of emphasis given to assessment in initial teacher education programs, it is necessary to put special emphasis on DA in pre-and in-service training programs. Finally, yet importantly, the lack of an agreed-upon framework for diagnostic assessment (Alderson, 2007) is a challenge for its practice in class. Hence, the findings described in the present study may raise awareness in the field of diagnostic assessment.

Scarce literature on the impact of DA on learning progress reveals that learning improves if DA is applied (Kazemi & Tavassoli, 2020; Mazloomi & Khabiri, 2016; Nikmard & Tavassoli, 2020). However, the findings of the present study suggest that experienced teachers develop a rich toolkit of diagnostic tools. Therefore, classroom teachers, especially novice ones, might be the best beneficiaries of the outcomes of studies on classroom DA. The findings of such studies could aid teachers and teacher candidates to be equipped with guidance and/or training on "how diagnosis might be appropriately conducted, what content diagnostic tests might have, what theoretical basis they might rest on, and how their use might be validated" (Alderson, 2005, p. 10). As teacher education programs have influence on pre-service teachers' beliefs about English Language teaching (Clark-Goff & Eslami, 2016) awareness raising tasks for DA might be incorporated in programs. The instruments to detect weaknesses and the remedies and solutions suggested here may help teachers, researchers, material developers, and teacher trainers.

The findings of this study have also pedagogical implications for teacher education or training programs. Workshops on helping teachers develop skills in implementing the processes of DA can help create more optimal learning environments. Finally, this study provided empirical evidence for the validity of the DA model suggested by Harding et al., (2015). Evidence favoring this DA model would be useful in improving the quality of teacher training and education programs.

References

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.

- Alderson, J. C. (2007). The challenge of (diagnostic) testing: Do we know what we are measuring? In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 21-39). University of Ottawa Press.
- Alderson, J. C., Brunfaut, T., & Harding, L. (2014). Issues in Language Testing Revisited. *Language Assessment Quarterly*, 11(2), 125-128.
- 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, 236–260.
- Alderson, J. C., & Lukmani, Y. (1989). Cognition and reading: Cognitive levels as embodied in test questions. *Reading in a Foreign Language*, 5(2), 253-270.
- ALTE (1998). Multilingual glossary of language testing terms. Cambridge University Press.
- Bates, C., & Nettelbeck, T. (2001). Primary school teachers' judgments of reading achievement. *Educational Psychology*, 21(2), 177-187.
- Blood, I. (2011). Diagnostic Second Language Assessment in the Classroom. Teacher's College, Columbia University Working Papers in TESOL and Applied Linguistics, 11(1).
- Brindley, G. (2001). Language assessment and professional development. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (eds.), *Experimenting with Uncertainty, Essays in Honor of Alan Davies, Studies in Language Testing 11* (pp. 126–136). Cambridge University Press,
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Clark-Goff, K., & Eslami, Z. (2016). Exploring Change in Preservice Teachers' Beliefs about English Language Learning and Teaching. *Iranian Journal of Language Teaching Research*, 4(3), 21-36.
- Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language Testing*, 21(3), 259-283.
- Feinberg, A. B., & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, 18(1), 52.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Routledge.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of second language writing*, 10(3), 161-184.
- Gattullo, F. (2000). Formative assessment in ELT primary classrooms: An Italian case study. *Language Testing*, 17(2), 278-288.
- Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading*. Routledge.

- Goodman, J. S., Wood, R. E., & Hendrick, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, 89(2), 248.
- 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.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Hedgcock, J., & Lefkowitz, N. (1996). Some input on input: Two analyses of student response to expert feedback in L2 writing. *The Modern Language Journal*, 80(3), 287-308.
- Helmke, A., & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgment accuracy on achievement. *Teaching and Teacher Education*, 3(2), 91-98.
- Hsu, Y. K. (2005). Developing and researching multiple assessments: An action research. National Science Council Individual Research Project (NSC 92-2411-003-031). Taipei: National Science Council.
- Huhta, A. (2008). 33 Diagnostic and Formative Assessment. *The handbook of educational linguistics*, 469.
- Jang, E. E., & Wagner, M. (2013). Diagnostic feedback in the classroom. *The companion to language assessment*, 2, 693-711.
- Işık, A. (2020). Do students feel that they are assessed properly?. *Iranian Journal of Language Teaching Research*, 8(1), 63-92.
- Lee, Y. W. (2015). Future of diagnostic language assessment. *Language Testing*, 32(3), 295-298.
- Lumley, T. (1993). The notion of subskills in reading comprehension tests: An EAP example. *Language Testing*, 10(3), 211-234.
- Kazemi, N., & Tavassoli, K. (2020). The Comparative Effect of Dynamic vs. Diagnostic Assessment on EFL Learners' Speaking ability. *Research in English Language Pedagogy*, 8(2), 223-241.
- Knoch, U., & Macqueen, S. (2017). Assessment in the L2 classroom. *The Routledge handbook of instructed second language acquisition*. Routledge.
- Kunnan, A. J., & Jang, E. E. (2009). 32 Diagnostic Feedback in Language Assessment. *The handbook of language teaching*. Wiley-Blackwell.
- Mazloomi, S., & Khabiri, M. (2016). Diagnostic assessment of writing through dynamic self-assessment. *International Journal of English Linguistics*, 6(6), 19-31.
- 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.

- Oscarson, M. (2014). Self-assessment in the classroom. In A.J. Kunnan (Ed.) *The Companion to Language Assessment*. Malden, MA: Wiley-Blackwell.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, 46(4), 265-273.
- Purpura, J. E. (2004). *Assessing grammar* (Vol. 8). Cambridge University Press.
- Rea-Dickins, P. (2007). Classroom-based assessment: Possibilities and pitfalls. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching: Part one* (pp. 505–520). New York: Springer.
- Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215-243.
- Selcuk, M. (2020). *The Correspondence between Test Developers' Presumptions and Test Takers' Perceptions* [Unpublished doctoral dissertation]. Yeditepe University.
- Urquhart, A. H., & Weir, C. J. (2014). *Reading in a second language: Process, product, and practice*. Routledge.
- Yang, T. L. (2008). EFL teachers' multiple classroom assessment practices of young language learners. *Studies in the Humanities and Social Sciences*, 9, 139-176.

Hossein Farhady received his PhD in Applied Linguistics from UCLA and has taught graduate courses in many universities, presented papers at professional conferences, conducted workshops, and published at national and international levels. His major areas of interest are research methodology, language assessment, and ESP. He is presently a professor of applied linguistics at Yeditepe University, Istanbul, Turkey.

Merve Selçuk is a METU graduate and holds a PhD degree in the field of English Language Education. She acts as the director of the School of Foreign Languages at Altınbaş University in Istanbul. She is a part-time lecturer at Bahçeşehir University ELT Department. She is also the co-founder and trainer of Language Assessment and Testing Academy.

APPENDIX**Classroom-based Diagnostic Assessment Practices of EFL Instructors**

Dear Respondent,

The purpose of this study is to explore experienced EFL instructors' diagnostic assessment (DA) practices related to reading. The information you provide will advance our understanding of teachers' classroom DA implementation. While answering the questions, please bear the specific problem you had experienced in mind and try to give as detailed answers as possible. Please remember that there are no single best but a variety of ways to implement the DA process to address the issue raised in the scenario.

The information will remain strictly confidential and will be used just for research purposes. We will also send a summary of our findings to participants upon request.

Name:

Email address:

Directions: Please check (✓) where applicable.

1. Gender Female Male
2. Highest academic degree

<input type="checkbox"/> Diploma	<input type="checkbox"/> BA/BS Degree (double)	<input type="checkbox"/> PhD Degree
<input type="checkbox"/> Associate Degree	<input type="checkbox"/> MA/MS Degree	
3. Place you teach/taught if you are not employed at present

<input type="checkbox"/> Primary School	<input type="checkbox"/> Senior Secondary School
<input type="checkbox"/> University Prep School	<input type="checkbox"/> Faculty
4. Years of teaching experience _____
5. Your field of study at university

BA	_____
MA	_____
PhD	_____
Other (CELTA/DELTA/etc.)	_____

Scenario I

In a pre-intermediate level intensive EFL class at the tertiary level, students complain that they neither enjoy reading nor like answering comprehension questions because they do not understand passages though the texts are not above their proficiency level. When you observe them for a week to find a way to overcome this problem, you notice that they do not highlight, underline or take notes while reading and they are not being active participants while reading. You also observe that they do not attempt to respond to a question when they do not understand the text. They fail to transform or manipulate certain tasks such as summarizing, paraphrasing, analyzing, and using context clues.

Scenario II

You are teaching reading skills in an advanced-level EFL class at the tertiary level. You observe that one of your students seems to be reading word by word and appears to look up every word in the dictionary. He understands single sentences but he cannot answer questions related to finding the main idea and supporting details. Nor can he perceive the author's attitude type of questions.

Scenario III

You are teaching a group of adult EFL learners in an intensive program. You feel that your international students from an Arabic country lack the skills to cope with reading comprehension tests. Most of them received low or failing grades in the mid-term because they performed poorly on the reading comprehension part of the test. They did much better on oral production and listening comprehension.

Questions for All Scenarios

1. What would you do to obtain more in-depth information about the student's detected problem?
2. What would be your initial hypothesis about the errors or mistakes?
3. How would you test your hypothesis? (What specific ways would you follow/what specific tasks would you use to make sure that your hypothesis is correct?)
4. What would you do after diagnosing the problem? (Assume that your first diagnosis is correct and note down all possible ways you could follow)
5. (If you have decided to intervene) What kind of intervention would you employ?
6. How would you check the effect of intervention?
7. How would you approach the students who did not benefit from the intervention?