



## Effect of a Cognitive Diagnosis Model-Based Feedback in an English Writing Class: Focusing on Korean English Learners\*

Sookyung Cho (Hankuk University of Foreign Studies)

Received: 24 August 2023  
Revised: 28 September 2023  
Accepted: 6 October 2023

**Cho, Sookyung. (2023). Effect of a cognitive diagnosis model-based feedback in an English writing class: Focusing on Korean English learners. *Modern English Education*, 24, 171-185.**

### Keywords

cognitive diagnosis model, formative assessment, L2 writing, teacher feedback  
인지진단모형, 형성평가, 제2 언어 작문, 교사 피드백

### Sookyung Cho

Professor  
Department of English Linguistics and Language Technology  
Hankuk University of Foreign Studies  
[sookyungcho@hufs.ac.kr](mailto:sookyungcho@hufs.ac.kr)

\*This work was supported by Hankuk University of Foreign Studies Research Fund of 2023.

### Abstract

This study aimed to investigate the effects of cognitive diagnosis model (CDM)-based teacher feedback on student writing ability and their perceptions. CDMs have been applied to various tests of language proficiency due to their ability to provide fine-grained information about a student's language proficiency. However, it has rarely been studied how this information can be incorporated into language instruction. Therefore, this study administered a multiple-choice writing exam to Korean English learners, analyzed the test results using a CDM, and used this information to evaluate student writing ability and provide feedback on their writing. The effects were measured in three ways: a comparison of the CDM analyses conducted at the beginning and end of the semester, a qualitative analysis of two focal students' writing, and an analysis of student questionnaires. The results revealed that most students increased their mastery level at the end of the semester. Also, the CDM-based feedback brought positive changes to the writing of the focal students. Finally, students had positive perceptions of the CDM-based feedback. These findings imply that CDM-based feedback can contribute to English as a Foreign Language (EFL) students' improvement in their writing abilities by adjusting it to their needs.

## INTRODUCTION

As formative assessment has taken a prominent role in educational discourse and practices (Wiliam, 2001a, 2001b), the importance of feedback has been re-emphasized in second language writing research. In formative assessment, teacher feedback plays a key role in students' acknowledging the gap between their current level and their goal and making their efforts in pursuing it. However, compared to one-on-one types of feedback, such as tutor- and peer-feedback, one of the biggest limitations of teacher feedback is its difficulty in tailoring feedback to each student's current level, especially in large-sized writing classes. For example, Truscott (1996) severely criticized teacher feedback on grammar correction, pointing out that it is not plausible for a teacher to know the correct level of each student in a large class. Furthermore, even if the teacher does know each student's level, giving consistent feedback to all of them throughout a semester can be challenging due to time and patience constraints.

These problems concerning teacher feedback can be alleviated by adopting cognitive diagnosis models (CDMs) for diagnosing students' writing abilities and tailoring teaching feedback to them. CDMs are a psychometric framework developed to collect, analyze, and report diagnostic data by scholars, such as Tatsuoka (1983) and Buck and Tatsuoka (1998). They are often used to extract fine-grained information from multiple-choice tests. For this reason, CDMs have not been as extensively applied to second language writing as they have been to listening and reading data, which are often tested in multiple-choice formats. Since Kim (2011), CDMs have been adopted in second language writing studies, implying the potential to provide writing instructors with detailed information about their student writers. However, until now, how CDM analysis can be incorporated into writing instruction has not been explored. Therefore, this study aims to incorporate the multiple-choice writing exam developed by Cho (2022) with a group of Korean learners of English and provide teacher feedback based on the CDM analysis of the test results. By investigating the students' test performance and writing drafts as well as their perceptions, this study seeks to answer the following research question:

- 1) How does CDM-based teacher feedback affect students' mastery levels in key areas in writing?
- 2) How does CDM-based teacher feedback affect students' writing?
- 3) What do students think about CDM-based teacher feedback?

## LITERATURE REVIEW

### Formative Assessment and Teacher Feedback

In the second language writing classroom, the importance of teacher feedback has already been highlighted due to the process-oriented approach to teaching writing in the 1980s. By emphasizing the process second language writers are involved in while creating their text, the process-oriented approach urges the writing instructor to intervene in this process and provide help appropriate to their current levels, which is often manifested in the form of teacher feedback on student writing. In the framework of formative assessment, this teacher feedback has become more important in the 1990s. Scholars who tried to explore the ways assessment is incorporated into instruction started to place an emphasis on feedback (Black & Wiliam, 1990, 2009; Ramaprasad, 1983; Sadler, 1998a, 1998b; Wiliam, 2011b). For example, Ramaprasad defines feedback as "information about the gap between the actual level and the reference level" (p. 4). Elaborating on this definition, Sadler (1989a, 1989b) emphasized that feedback should be provided for the purpose of altering the gap between the actual level of a learner and the reference level. In order to make changes on an individual learner's learning, thus, three steps have been recommended as essential to formative feedback: 1) to diagnose where the student is; 2) to indicate where the student needs to get; and 3) to present what is needed to get there (Culham, 2018; Tuttle, 2008; Wiliam, 2011a).

To date, studies that have investigated the role of formative feedback in a writing class have often focused on the second and third steps: their instructional treatments. For example, a group of scholars suggested that teacher feedback should be accompanied by other types of feedback, such as peer feedback, self-feedback, or student-teacher conferencing, so that it will raise students' awareness about good writing (Black & Wiliam, 1998; Hattie & Timperley, 2007; Hyland, 2003). Another group of scholars has suggested that teacher feedback should deal with a wide range of topics, including content, organization, mechanics, or style; it should provide feedback on global errors first and later on local errors; and its form should be indirect for long-term benefits (Hyland & Hyland, 2006; Ferris, 2014; Shute, 2008). Lastly, these scholars unanimously argue for the necessity of multiple drafts in a writing class (Huot & Perry, 2009; Lee, 2007; Nicol & Macfarlane-Dick, 2006; Sadler, 1998a, 1998b; Shute, 2008). According to McGarrel and Verbeem (2007), the chance for students to revise their drafts based on teacher feedback "prompts writers to go beyond surface-level revisions and to rethink their intended meanings at a deeper level of engagement with their texts" (p. 231). That is, without this chance, the students would not be able to fill the gap between their current level and the reference goal, which is the final goal of formative feedback.

Compared to the extensive research on providing formative assessment, the initial step of diagnosing where the student is, is often overlooked or not given adequate attention. In order to provide feedback appropriate to a student's needs, it is just as important to accurately diagnose their initial level as it is to provide them with formative feedback and teach them how to incorporate it into their revised texts. However, accurately diagnosing a student's initial level is not easy due to several characteristics inherent in the method of assessment. Writing is often assessed in a direct way where human raters evaluate or score an essay written by a student. When evaluating a student's writing, they usually depend on a holistic scale that encompasses various areas of writing or on an analytic scale that is based on a few key areas, such as content,

organization, grammar, or style. Since this procedure requires expertise from well-trained human raters who have taught and assessed a particular type of writing, which is tested in a writing test, more of them are needed as the number of students increases. In addition to the issue of time and effort invested in this type of direct assessment, there have been concerns about human bias involved in a direct assessment of writing. Through a computational comparison of self-, peer-, and teacher-assessment of student writing, Matsuno (2009) shows that teacher feedback is more biased than the other types of feedback. In order to compensate for this kind of shortcoming of teacher feedback, this study adopts a cognitive diagnosis model (CDM) in diagnosing students' initial level before providing formative teacher feedback.

## Use of CDMs in L2 Writing Studies

CDMs are psychometric frameworks designed to collect, analyze, and report diagnostic data (Xin et al., 2022). They were originally developed in the 1970s to extract more fine-grained information from test scores, such as math and reading scores, based on the assumption that a test measures latent multivariate attributes of an examinee, not just their unidimensional single attribute. Studies that use CDMs often develop a Q-matrix and apply it to the test results. By determining whether the examinees have mastered the latent attributes through the Q-matrix, these studies can eventually identify each examinee's weaknesses and strengths.

As seen, since the original purpose of CDMs is to extract diagnostic information about individual examinees' test scores, they have been applied to L2 learners' reading, listening, vocabulary, or grammar tests because of the ease with which they can quantify the results. For example, Sawaki et al. (2009) developed Q-matrices to provide more fine-grained information to test-takers of the TOEFL iBT (Test of English as a Foreign Language Internet-based Test). They asked a group of content experts to develop Q-matrices on TOEFL iBT listening and reading items and a group of psychometricians to apply these Q-matrices to the test scores to determine their validity. Lee and Sawaki (2009) also applied three different models of CDMs—the general diagnostic model, the fusion model, and latent class analysis—to TOEFL reading and listening items to see whether they made any difference in interpreting test scores. The comparison of the three CDM models in Lee and Sawaki (2009) revealed that all three were valid tools for providing detailed information on the strengths and weaknesses of test takers, despite subtle differences between them. Similarly, Park and Cho (2011) applied the DINA (deterministic inputs, noisy and gate) model to a grammar test by developing a Q-matrix of six attributes critical to grammar knowledge: tense, verb usage, idiomatic expression, modifier, agreement, and preposition. In addition, Park and Cho (2017) compared two different CDMs—the DINA and the DINO (deterministic inputs, noisy or gate) models—by applying them to a test of Korean learners' English reading comprehension. They found that the mixed model of the DINA and the DINO was the best, as test takers were able to compensate for their weaknesses with other skills, even if they had not mastered certain skills required to answer a particular test item.

While CDMs have been extensively applied to the interpretation of receptive skills in L2 learners, they have rarely been used to analyze productive skills such as speaking and writing. These skills are mostly tested in a direct way and were thought to produce more detailed information than a mere test score, making them difficult to quantify, unlike multiple-choice tests. However, since Kim (2011) first applied a CDM to a writing test in 2011, several scholars have used CDMs to analyze L2 learners' writing skills. Kim (2011) developed a Q-matrix consisting of 35 concrete, fine-grained descriptors and applied it to TOEFL writing essays written by ESL (English as a Second Language) learners. Building on Kim's work with modifications, Xie (2016) and Shahsavari (2019) also used CDMs to analyze writing exams written by different groups of students.

When CDMs are applied to writing evaluation, however, their procedure is not easily conductible and is time-consuming. These studies have asked second language writers to write about a topic during a certain period of time and a large number of human raters to evaluate each student's writing based on a Q-matrix. For example, Kim (2011) asked ten ESL writing instructors who had teaching experience of 9.2 years on average and asked them to grade 480 essays in total. Xie (2016) asked ten experienced EFL (English as a foreign language) writing instructors who had teaching experience of 5 to 20 years to grade 472 essays.

In order to overcome these difficulties, Cho and Park (2021) and Cho (2022) developed a multiple-choice writing exam instead of an essay exam and applied the DINA to the test results. They found that the test is statistically valid and that furthermore, it is less time-consuming and more easily conductible than an essay exam. Using this multiple-choice test, this study tries to explore a way of incorporating CDM analysis into writing instruction, as it has rarely been studied how CDM analysis can be used in a real classroom, in spite of its often-made claim that the fine-grained information obtained through a CDM can help both the instructor and the students in a classroom

## METHOD

### Participants

At the beginning of the semester, 176 students (79 females and 97 males) from an English-related department at a university in Seoul, Korea, volunteered to take the multiple-choice writing test developed by Cho (2022).<sup>1</sup> Their test results were analyzed using the DINA model, one of the most commonly used CDMs.<sup>2</sup> Out of these 176 students, only nine freshmen (seven females and two males) were included in this study to examine the effect of CDM-based feedback on student writing. They were selected for two reasons: first, based on a placement test, they had been placed in a lower-level writing course taught by the author of this study, which allowed her to provide feedback based on the CDM analysis of their test results; secondly, the CDM analysis showed that they had not yet mastered certain areas of writing, so the feedback was expected to help them identify areas for improvement. Table 1 shows their CDM results.

**TABLE 1**  
*The Students' CDM Results*

Group	Student	CON	ORG	GRM	VOC	MCH
CON	Student 1	0.72	0.83	0.82	0.93	0.76
MCH	Student 2	0.18	0.54	0.48	0.81	0.36
CON	Student 3	0.19	0.18	0.17	0.64	0.50
ORG	Student 4	0.14	0.15	0.16	0.65	0.45
GRM						
ORG	Student 5	0.55	0.23	0.53	0.04	0.63
VOC	Student 6	1.00	0.07	1.00	0.08	1.00
	Student 7	1.00	0.80	1.00	0.82	1.00
VOC	Student 8	0.93	0.91	0.90	0.87	0.94
	Student 9	0.33	0.33	0.30	0.04	0.45

*Note.* CON=content; ORG=organization; GRM=grammar; VOC=vocabulary; MCH=mechanics; graded areas refer to the relatively weaker areas compared to the other areas.

In Table 1, the scores closer to 1.00 indicate a higher probability of the participant's having mastery of a specific domain. Despite the differences in their exact mastery levels of those five areas, the CDM analysis reveals some congruent patterns among students, although there was some variability across the individual scores: Student 1 and 2 were relatively weak in content and mechanics; Student 3 and 4 were weak in content, organization, and grammar; Student 5, 6, and 7 were weak in organization and vocabulary; and Student 8 and 9 were weak in vocabulary. None of the students had experience studying abroad or working on academic writing in English. The instructor for this study was the author, who has over 15 years of experience teaching English writing to Korean learners of English and has taught this particular writing class since 2018.

### Data Collection

As explained above, on the first day, all participants took a multiple-choice writing exam. To raise their awareness of the

<sup>1</sup> This test was used in this study because it was proved to be sufficiently valid and reliable in Cho (2022).

<sup>2</sup> The DINA model is a conjunctive and noncompensatory CDM based on an assumption the examinee should master all necessary attributes for a correct response to an item. It is one of the most commonly used CDMs, whose equation is as follows:

$$P(x_{ik} = 1 | \eta_{ij}) = g_j^{1-\eta_{ij}}(1 - s_j)^{\eta_{ij}}$$

In this equation,  $\eta_{ij}$  refers to a latent variable showing whether an examinee  $I$  ( $=1, 2, 3, \dots, I$ ) mastered the necessary attributes to correctly answer the item  $j$  ( $=1, 2, 3, \dots, J$ ). The equation controls the noise factors, such as the probability of the examinee's correctly answering the item, although s/he did not master all the required attributes ( $g_j$ ) and the probability of the examinee's incorrectly answering the item, although s/he mastered all the required attributes ( $s_j$ ). R 4.1.1 (R Core Team, 2021) was used for all the DINA analyses.

current writing skills, the instructor shared these test results with the students before beginning instruction in the form of a diagnostic sheet (see Appendix 2 for details). From the eighth week to the twelfth week, the participants wrote two major assignments, an argumentative essay and a cause-and-effect essay. For each assignment, they were asked to write two drafts, a first and final draft. One week after submitting the first draft, students received teacher feedback focused on their weak areas, based on the CDM analysis of their test results, as seen in Figure 1. For instance, Student 4 from Group 2 wrote an essay that argued that online courses cause many problems and was given the following feedback because the CDM analysis showed that she was weak in content, organization, and grammar:

Dear Student 4

As your peer reviewer Student 5 mentioned, you picked a very timely issue for your essay. If you work on the following things in your revision, your essay will be better developed.

Content & Organization

1. In your first and third paragraphs, the topic sentences do not match with the details. For example, in your first body paragraph, you talked about “a gap,” although your major focus is on decreased academic ability of the students in the online classes. Also, your third body paragraph is about high-quality education, but you deal with problems inherent in conducting practical courses in the online courses. You need to tailor your topic sentence to the details or vice versa.

Grammar

1. Some of your sentences can be simplified if you remove some empty or vague words. See my suggestions. Check the usage of the expression “it is difficult to do sth” vs. “something is difficult to do.”

**FIGURE 1**

*Sample Teacher Feedback on Student 4*

Within a week of receiving this kind of teacher feedback, all students revised their first draft based on the teacher's feedback and resubmitted their revised draft. The participants repeated the same procedure for the second writing assignments as well.

Finally, on the last of the course, the participants took the same multiple-choice test again. This was conducted to determine whether the formative feedback based on the CDM analysis made a difference in their mastery levels of key writing skills. Additionally, they also answered a question that asked about their experiences of receiving the CDM-based feedback.

## Data Analysis

For the first research question, how the CDM-based teacher feedback affects students' mastery level of each area, this study compares all participants' CDM analyses of the multiple-choice writing skills test, which they had taken in the beginning of and at the end of the semester respectively. In order to address the second and the third research questions, qualitative analyses were conducted. For the question of how the CDM-based teacher feedback affects their writing, two focal students—Student 2 and 3 were selected because their mastery level increased remarkably compared to their initial diagnosis results. To see how teacher feedback based on the CDM analysis has affected their writing, their first draft and final draft were compared and contrasted in detail with regard to both of their writing assignments. For the third research question, the answers the nine students wrote about their experiences of receiving this CDM-based feedback were analyzed line by line.

## RESULTS

### Comparison of CDM Tests Across Participants

Table 2 shows the CDM results of the two tests each student took on the first day and on the last day of the course.

**TABLE 2**  
*Comparison of CDM Results*

Group	Student	First Test Results					Second Test Results				
		CON	ORG	GRM	VOC	MCH	CON	ORG	GRM	VOC	MCH
CM	Student 1	0.72	0.83	0.82	0.93	0.76	1.00	0.98	1.00	0.98	1.00
	Student 2	0.18	0.54	0.48	0.81	0.36	0.99	0.99	0.99	1.00	0.99
COG	Student 3	0.19	0.18	0.17	0.64	0.50	0.99	0.99	0.99	0.99	0.99
	Student 4	0.14	0.15	0.16	0.65	0.45	0.20	0.19	0.22	0.57	0.47
OV	Student 5	0.55	0.23	0.53	0.04	0.63	0.95	0.51	0.95	0.53	0.96
	Student 6	1.00	0.07	1.00	0.08	1.00	0.95	0.03	0.95	0.00	0.95
	Student 7	1.00	0.80	1.00	0.82	1.00	1.00	1.00	1.00	1.00	1.00
V	Student 8	0.93	0.91	0.90	0.87	0.94	1.00	1.00	1.00	1.00	1.00
	Student 9	0.33	0.33	0.30	0.04	0.45	0.77	0.79	0.15	0.04	0.48

*Note.* CM=content & mechanics; COG=content, organization, & grammar; OV=organization & vocabulary; V=vocabulary.

After receiving teacher feedback based on the CDM analysis, most of the students, except for Student 6 and S9, showed an increase in their mastery level in the areas where they were diagnosed as weak through the CDM analysis. Moreover, five of them—Student 1, 2, 3, 7, and 8—approached near mastery level in all the other areas as well. The comparison of groups showed that the CM Group and COG Group received relatively more benefits from the CDM-based teacher feedback than the other groups. Specifically, both Student 1 and 2 in the CM Group increased their mastery level in all areas, including their weak areas of content and mechanics. Similarly, Student 3 and 4 in the COG Group increased their mastery level in all areas, although Student 4's increase was much smaller than Student 3's and her mastery level in grammar decreased slightly (from .65 to 0.57). On the other hand, for Groups OV and V, there is variation among the participants. In the OV Group, two of the participants—Student 5 and 7—increased in all areas, while Student 6 decreased in the mastery levels of the focused areas, organization and vocabulary. In Group V, Student 8 increased in her mastery level of vocabulary, but Student 9 remained at the same level.

To sum up, it appears that CDM-based feedback on content, organization, and mechanics is more effective than feedback on vocabulary. The difference in mastery levels across the categories may indicate that a large volume of English vocabulary is not easy to master through just a few instances of feedback.

## Students' Changes Across Drafts

As seen in the above, out of the participants who increased in their mastery levels between the beginning and the end of the semester, Student 2 and 3 show more drastic changes—they did not master any of the areas in the beginning, but arrived at the near mastery level across all the areas, including their weak areas. For this reason, this section tries to analyze how their writing drafts have changed throughout the semester with the help of the CDM-based teacher feedback.

In her first writing assignment, Student 2 wrote about the problems with the label "the MZ generation," which refers to two groups: Millennials (born 1981-1995) and Generation Z (born 1996-2005). In her first draft, she gave three reasons explaining why this labeling is problematic: 1) it may not account for individual differences, 2) it can cause generational conflicts, and 3) it is too broad. The initial CDM analysis diagnosed that Student 2 was weak in content and mechanics, so the instructor gave feedback, particularly focusing on content, as follows:

Dear Student 2
Content & Organization
1. Most of your supporting points look quite valid, but they need more evidence and explanations.
2. Compared to your first and second supporting points, your third one is not that clear. You should clarify what you want to say in your third body paragraph.

**FIGURE 2**  
*Feedback on Student 2's Argumentative Essay*

Incorporating this teacher feedback into her revision, Student 2 replaced her third body paragraph with a new paragraph as follows:

Lastly, labeling generations causes identity confusion. Alphabetical generation classification(XMZA) is not created by the social science methodology. The year of birth, which is the basis for the generation MZ classification, might be slightly different depending on the analysis agencies. For example, some people who were born in 1996 sometimes could be classified as the Millennial generation, and sometimes the generation Z. According to a Pew Research Center survey, in addition, quite a few people were not sure which generation they belong to (qtd. in Cohen). Therefore, this generational classification is but an illusion, and it has no real basis.

### FIGURE 3

#### *Student 2's Revised Argumentative Essay*

At the request of the teacher's feedback that the third supporting point was not clear in her first draft, Student 2 dropped the original point, which stated that the term MZ generation is too broad. Instead, she introduced a new topic, stating that the labeling causes confusion even among youths who are reported to belong to this generation due to its vagueness. Additionally, Student 2 followed the teacher's first feedback by using more evidence and added a survey result from Pew Research Center to support her argument.

The improvement in her content development was witnessed again in her second assignment, which she was supposed to submit two weeks after she completed the first writing assignment. In the second writing assignment, she was asked to write a cause-and-effect essay and she wrote about why vlogs are so popular these days in Korea. On this draft, the instructor only asked Student 2 to elaborate on some points as seen in Figure 3, which requires relatively smaller-scale changes, compared to the first feedback which required Student 2 to develop a new paragraph.

Dear Student 2,

I enjoy reading your essay. Your topic looks quite interesting to me. I hope you will work on the following things in your revision.

#### Topic & Elaboration

1. I'm curious why vlogs are more informational and profitable than other types of medias. Although it is not a comparison-and-contrast essay, you'd better explain whether the benefits you mentioned in your essay are really unique to vlogs.
2. Since a cause-and-effect essay is a kind of argumentative essay, it will be helpful if you introduce some disadvantages of vlogs in the introduction or in the conclusion. This will make you a more sensible writer.

### FIGURE 4

#### *Feedback on Student 2's Cause-and-Effect Essay*

In her second draft, Student 2 tried to incorporate this teacher feedback and turned in a better-organized and more coherent essay.

On the other hand, Student 3 has a high interest in sports, so he chose topics related to it for both of his writing assignments: Islam Makhchev, the professional fighter is unbeatable for an argumentative essay and the Qatar World Cup has reverse effects for a cause-and-effect essay. Because he was diagnosed as weak in content, organization, and grammar, the teacher focused on providing feedback on these areas to Student 3 as in Figure 5.

Dear Student 3

I can easily tell that you are quite into sports. I look forward to reading your complete essay. When you revise your draft, please work on the following things.

#### Content & Organization

1. It is not clear how Islam overcame his weaknesses in your second body paragraph. Please explain about them.

2. Work on your conclusion.

Grammar

1. I hope that you bear in mind that a countable noun in English requires an indefinite article when it is first mentioned in your text.

### FIGURE 5

*Feedback on Student 3's Argumentative Essay: Overall Comments*

Regarding his content and organization, the instructor pointed out that his second body paragraph, which deals with how Islam overcomes his weaknesses, is not clear. As for his grammar, the instructor not only explained the rule of the English indefinite article that he often missed in his draft, as seen in Figure 5, but also made direct corrections on his draft, as seen in Figure 6.

Maybe, you can be surprised by the title. What this title means is Islam Makhachev, the current champion of lightweight division in UFC, not a religion 'Islam'. UFC is an abbreviation of Ultimate Fighting Championship and a corporation getting benefit by MMA(Mixed Martial Arts) broadcasting. UFC divides weight divisions for preventing its fighters from injuries, and there is a lightweight division among them. Anyway, after Islam becoming a champ, many people said that anyone can beat Islam until his retirement because of his amazing performance in the latest game. However, I am sure Islam cannot be broken not only for the above reason but also for his physical strengths, overcoming ability and present situation in lightweight division in UFC.

[채교:] 2022-11-04 13:59  
You mean "no one"?

### FIGURE 6

*Feedback on Student 3's Argumentative Essay: Direct Correction*

In his final draft, Student 3 elaborated on his second body paragraph as requested in the teacher feedback, but he also added one more paragraph that explains why Islam is unbeatable by the other players. In addition to these changes in his content and organization, Figure 7 demonstrates that Student 3 made fewer mistakes in his use of the English indefinite articles in his final draft.

Nevertheless, a certain circle say that there are some fighters like Beneil Dariush, Mateusz Gamrot, and Michael Chandler who are estimated that they can destroy Islam. Mateusz was a former champ of KSW, the biggest mma company in Poland, and he had both titles in lightweight and featherweight. Although Mateusz has a great career, Beneil stopped him at the identical event which Islam stole the champ's belt. Beneil even planned to fight with Islam in February 26, but he got ankle injury before their fighting, so Beneil still has a chance to prove his possibility next time(Okamoto). ....

### FIGURE 7

*Student 3's Final Draft of Argumentative Essay*

The paragraph shown in Figure 7 is a new paragraph added to his final draft, so the instructor did not have a chance to correct his use of the indefinite articles; however, the new paragraph does not contain as many mistakes of 'a/n' as his paragraph from Figure 5 does.

Although his second writing assignment is not error-free or perfect in terms of content, organization, and grammar, the teacher's feedback on Student 3's second writing assignment only asks him to clarify his thesis. Despite his title "Reverse Effect of the Qatar World Cup," some of his points did not clearly demonstrate its negative effects, and thus the instructor points this out as in Figure 8.

Dear Student 3

Your topic is quite timely, since the whole world are so fascinated by the World Cup these days. I look forward to reading your completed essay.

Topic & Elaboration

1. It seems that you would like to deal with negative effects of the Qatar World Cup, but it is not clear on what they have negative effects. You'd better clarify it from your thesis statement. For example, your first supporting sentence is about its negative effect on players' health. How about the others?

### FIGURE 8

*Feedback on Student 3' Cause-and-Effect Essay*

After receiving this feedback, Student 3 attempted to focus on the negative effects of the Qatar World Cup in his final draft, although his attempt was not flawless. While elaborating on the negative effects, he added another paragraph that explains the negative effects of Qatar's national religion, which is not necessarily related to the Qatar World Cup.

### Students' Perceptions on CDA-Based Feedback

Out of nine participants, eight showed quite positive opinions on the CDM-based feedback. All of them claimed that the greatest advantage of the CDA-based feedback is to make them aware of their weaknesses with concrete scores and figures so that they could keep these weaknesses in mind while writing their drafts throughout the semester.

"I was able to identify which skills I lacked through the concrete scores."  
(Excerpt from Student 1)

"At the beginning of the semester, I found out that I was weak in vocabulary, so I was able to improve it through the teacher feedback that focused on vocabulary."  
(Excerpt from Student 7)

It may be quite natural that Student 1 and 7, whose mastery levels increased in their second multiple-choice test, have shown this kind of positive opinion on the CDM-based teacher feedback. However, it is quite surprising that even the students who did not show an increase in their mastery level in the second test also think positively of the CDM-based feedback. For instance, Student 9 remarks, "Through the diagnostic test, I came to know which areas I am weak in. Since I received teacher feedback focusing on these weak areas, I was able to know what to fix and how to fix them."

Moreover, a few students, such as Student 3 and 4, show much higher confidence in taking the second multiple-choice test through their one-semester experience of receiving the CDM-based feedback.

"I felt more confident in the second test compared to the first test. I'm not sure whether I was right in all the questions, but I came to think that now I know how to solve them."  
(Excerpt from Student 3)

"I felt the second test was much easier to answer than the first test, and now I could easily tell which areas each question dealt with. So I thought that I learned from the teacher feedback."  
(Excerpt from Student 4)

As seen in these two excerpts, Student 3 and 4 gained more confidence in answering questions in the multiple-choice writing test. This confidence was confirmed in their near-mastery levels in all five areas in the results of the second test. While this raised confidence may relate to their familiarity with the test by taking it twice, the 14-week interval between the tests is too long for them to remember the test items. Rather, it is more likely that they gained a better understanding of what each test item tests and how to handle the questions, rather than simply remembering the test items and answers, as seen in the excerpts from Student 3 and S4.

## DISCUSSION AND CONCLUSION

The investigation of the CDM teacher feedback given to the students during the semester reveals that students seem to have an increase in their mastery levels of all the key areas. In particular, those who were diagnosed as weak in content and mechanics or content, organization, and mechanics seem to benefit the most from the CDM-based teacher feedback, looking at their test results. This finding suggests that areas such as content, organization, grammar, and mechanics may be good candidates for a writing instructor to provide feedback on by guiding students on where to improve and how to fix their problems. On the other hand, students who were weak in vocabulary showed variation across the participants, implying that CDM-based feedback alone may not be enough to improve a student's vocabulary knowledge. This can be attributed to a specific characteristic of English vocabulary, which is too vast to be mastered within a short period.

Secondly, the in-depth analysis of the two focal students, Student 2 and 3, who marked outstanding improvement from their first test result to the second one, demonstrates how CDM-based teacher feedback can make changes in student writing. For example, Student 2, who was diagnosed as weak in content and mechanics, received teacher feedback focused on these areas and made efforts to overcome her weaknesses throughout multiple drafts. In the case of Student 3, who was diagnosed as weak in content, organization, and grammar, he received teacher feedback with an emphasis on these areas. For instance, after being notified of his mistakes in the use of an indefinite article in English, he paid more attention to its use and reduced the number of such mistakes in his second draft.

Finally, students' answers to the question of how they perceive CDM-based teacher feedback reveal that most of them think it makes them aware of their weaknesses and enables them to work on these areas throughout the semester. In particular, they mention that the CDM profile is helpful because it indicates the extent to which they have mastered each area using concrete figures. Even the student who did not improve significantly in her second test result argues for the benefits of the CDM-based feedback. Additionally, CDM-based feedback also boosts the students' confidence level in taking the test itself.

This study is limited in that it was conducted on a small scale based on a small number of participants, and that it only examined lower-level students. Future research should consider including more participants of diverse proficiency levels to ensure a more comprehensive understanding and applicability. These limitations notwithstanding, the findings of this study imply that they can be easily expanded to a larger number of different groups of participants across various contexts. The development of a multiple-choice writing test and the application of CDM analysis to it will make teacher feedback more formative by providing writing instructors with rich information about their students' exact levels of skills and knowledge, as well as making students aware of the gap between their current status and their reference point so that they can put more effort into improving their skills.

As seen above, the findings of this study corroborate the positive findings of many studies on the effectiveness of a teacher's formative feedback in a classroom setting (Black & Wiliam, 1998; Hattie & Timperley, 2007; Hyland, 2003). According to them, since every student has different needs, formative feedback requires different kinds of responses tailored to their needs (Wiliams, 2011b). In order to utilize teacher feedback for the purpose of formative assessment, two major concerns emerge in a second language writing class. First, it is often not technically and logistically plausible for a teacher to provide customized feedback to each writer in a class where he or she teaches a large number of students. Second, even if a teacher is willing to go through the time-consuming and highly labor-intensive procedure of tailoring his or her feedback to each student, he or she may face another problem of accurately diagnosing each writer's writing ability. A multiple-choice writing test, as used in this study, may alleviate the problem of handling a large number of students: the application of CDM to the test results can help both the instructor and the student writer to be aware of their current levels of English writing skills.

## References

- Black, P., & Wiliam, D. (1990). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice*, 5(1), 1-65.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Buck, G. & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: Examining attributes of a free response listening test. *Language Test*, 15(2), 119-157.
- Cho, Sookyung. (2022). Application of a cognitive diagnosis model to a multiple-choice English writing test. *Korean Journal of Education Research*, 60(7), 21-40.
- Cho, Sookyung, & Park, Chanho. (2021). Development of a multiple-choice writing examination validated with a cognitive diagnosis

- model. *Modern English Education*, 22(4), 12-23.
- Culham, R. (2018). *Teach writing well: How to assess writing, invigorate instruction, and rethink revision*. Stenhouse Publishers.
- Ferris, D. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, 19(1), 6-23.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Huot, B., & Perry, J. (2009). Toward a new understanding for classroom writing assessment. In R. Beard, D. Myhill, J. Riley & M. Nystrand (Eds.), *The SAGE handbook of writing development* (pp. 423-435). Sage.
- Hyland, F., & Hyland, K. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101.
- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- Kim, Youn-Hee. (2011). Diagnosing EAP writing ability using the reduced reparameterized unified model. *Language Testing*, 28(4), 509-541.
- Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *The Canadian Modern Language Review*, 64(1), 199-214.
- Lee, Yong-Won, & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, 6(3), 239-263.
- Matsuno, S. (2009). Self-, peer-, and teacher-assessment in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75-100.
- McGarrel, H., & Verbeem, J. (2007). Motivating draft revision through formative feedback. *ELT Journal*, 61(3), 228-236.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Park, Chanho, & Cho, Sookyung. (2011). Cognitive diagnostic assessment of English grammar for Korean EFL learners. *English Teaching*, 66(4), 101-117.
- Park, Chanho, & Cho, Sookyung. (2017). An exploratory analysis of compensatory cognitive diagnosis in EFL reading comprehension. *Korean Journal of Language and Linguistics*, 17(1), 85-104.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Ramaprasad, A. (1983). On the definition of feedback. *Systems Research and Behavioral Science* 28(1), 4-13.
- Sadler, D. R. (1989a). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Sadler, D. R. (1989b). Formative assessment: Revisiting the theory. Formative assessment and the design of instructional systems. *Education: Principles, Policy, and Practice*, 5(1), 77-84.
- Sawaki, Y., Kim, Hae-Jin, & Gentile, C. (2009). Q-matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessment. *Language Assessment Quarterly*, 6(3), 190-209.
- Shahsavari, Z. (2019). Diagnosing English learners' writing skills: A cognitive diagnostic modeling study. *Cogent Education*, 6(1), 1-19.
- Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20(4), 345-354.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-369.
- Tuttle, H. G. (2008). *Formative assessment: responding to your students*. <https://ebookcentral.proquest.com/lib/malmo/detail.action?docID=1595017#>
- William, D. (2011a). *Embedded formative assessment*. Solution Tree Press.
- William, D. (2011b). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14.
- Xie, Q. (2016). Diagnosing university students' academic writing in English: Is cognitive diagnostic modelling the way forward? *Educational Psychology*, 37(1), 26-47.
- Xin, T., Wang, C., Chen, P., & Liu, Y. (2022). Editorial: Cognitive diagnostic models: Methods for practical applications. *Frontiers in Psychology*, 13, 1-3.

## Appendix

### Multiple Choice Writing Exam

Read the passage and choose the answer to each question that most effectively improves the quality of writing in the passage and that makes the passage conform to the conventions of standard written English. Each question includes a “NO CHANGE” option. Choose that option if you think the best choice is to leave the relevant portion of the passage as it is.

[1] There are a number of ways to learn a foreign language. (#1) Students can study alone from a book, or go to a country where the language is spoken and learn it [2] while living there. [3] For the majority of learners, however, taking private or group language lessons is a natural choice. Some people believe that private lessons are a better way to acquire a language because they are more flexible and can meet students’ specific needs. (#2) In reality, however, group lessons are more effective for language learning because they allow students to learn from each other, [4] they reduce the stress of the learning environment and create a supportive atmosphere, all of which are [5] vital to successful language acquisition.(#3)

First, there is evidence that peer-learning, or learning from other students, is essential in foreign language acquisition. [6] In language classes, learners do pair or group work, practice negotiating meaning, start and control discussions, and produce a wide variety of expressions (Raymond 30). All of these activities [7] believed to efficiently facilitate the acquisition of a foreign language. [8] Some people may argue that students can practice the same skills more effectively with teachers in private lessons. However, research indicates that most of the conversation in private lessons [9] consists of teachers asking questions and students answering them (Horthorn 52). Since it is questionable that the [10] skills mentioned above obtained successfully simply by answering questions, private lessons seem to fail in providing an environment in which learners can acquire the language competently.

Another advantage of group lessons is that it can create a less stressful atmosphere, which is also desirable in foreign language learning. To begin with, the most important condition for language learning is that the learner should not be [11] bullied or exhausted. [12] Because those who are against group lessons may insist that [13] a teacher’s being able to provide constant attention and support to a student is a significant advantage in private lessons, this may not always be true. According to *Teaching a Language in Urban Areas*, group lessons are considerably [14] less stimulating or tiring for the average learner because there is no “non-stop spotlight of the instructor’s attention (Jordin 40). [15] This finding suggest that group lessons are a desirable option for people who do not learn well in stressful situation.

[16] In addition the feeling of support which a learner can receive from their classmates is a valuable benefit which cannot be gained in private lessons. According to Johnson and Swinton, peers tend to encourage each other [17] which becomes a “fabulous motivator.” Also, it is true that [18] friendship and a sense of support peers share with each other should not be underestimated in language learning. Learning a language can be a challenging experience for shy people. It also requires a great amount of endurance because language acquisition does not occur instantly. Knowing that there are others who are going through the same challenge can encourage students to continue their learning. [19] Group lessons also have the great advantage of allowing different groups of students to perform different tasks and activities according to their ability. [20]

1. The writer is considering revising this sentence. Should the writer do this?

- A) Yes, because it is too general to attract readers’ attention.
- B) Yes, because it is off the topic of the passage.
- C) No, because it well-informs the readers of the topic of this passage.
- D) No, because it is grammatically correct.

2.

- A) NO CHANGE
- B) while living there
- C) they’re living there
- D) while they lived there

3. For the sake of cohesion in this paragraph, this sentence should be placed

- A) where it is now.
- B) in (#1).
- C) in (#2).
- D) in (#3).

4.

- A) NO CHANGE
- B) they reduce the stress of the learning environment, and create
- C) reduce the stress of the learning environment, and create
- D) reduce the stress of the learning environment, and they create

5.

- A) NO CHANGE
- B) active
- C) strong
- D) insignificant

6. At this point, the writer is considering adding the following sentence. Should the writer make this addition here?

Studies suggest that students learn language mostly from their peers rather than from teachers (Raymond 25)

- A) Yes, it is necessary because it works as key evidence that supports the topic of this paragraph.
- B) Yes, it is necessary because it suggests overall benefits of peer learning over teacher-directed learning.
- C) No, it is unnecessary because it shifts the focus of this paragraph from foreign language learning to peer-learning.
- D) No, it is unnecessary because there is enough evidence provided to support the topic of this paragraph.

7.

- A) NO CHANGE
- B) are believed to
- C) believe
- D) believed in

8. The writer is considering deleting this sentence. Should the writer do this?

- A) Yes, because it does not provide a transition from the previous paragraph.
- B) Yes, because it contradicts the authors' opinion.
- C) No, because it works as a bridge from the previous sentence to the following sentence.
- D) No, because it effectively supports the author's main argument in this paragraph.

9.

- A) NO CHANGE
- B) consist of
- C) is consisted of
- D) consists

10.

- A) NO CHANGE
- B) the skills mentioned above can be obtained successfully
- C) the skills are mentioned above obtained successfully
- D) the skills are mentioned above can be obtained successfully

11.

- A) NO CHANGE
- B) disturbed
- C) intimidated
- D) convinced

12.

- A) NO CHANGE
- B) However,
- C) While
- D) After

13.

- A) NO CHANGE
- B) a teacher is being able to provide constant attention
- C) a teacher can provide constant attention
- D) a teacher's ability to provide constant attention

14.

- A) NO CHANGE
- B) less daunting
- C) troublesome
- D) more relaxing

15. The writer is considering deleting this sentence. Should the writer do this?

- A) Yes, because it is off topic.
- B) Yes, because it is redundant.
- C) No, because it provides the writer's interpretation of the outside source.
- D) No, because it summarizes the main point of this paragraph.

16.

- A) NO CHANGE
- B) In addition,
- C) Nonetheless,
- D) In contrast,

17.

- A) NO CHANGE
- B) which
- C) of which
- D) what

18.

- A) NO CHANGE
- B) friendship and a sense of support peers shared with each other underestimated
- C) friendship and a sense of support peers underestimated should not be shared
- D) friendship and a sense of support peers share with each other are underestimated

19. The writer is considering deleting this sentence. Should the writer do this?

- A) Yes, because it is off topic.
- B) Yes, because it is redundant.
- C) No, because it is closely related to the topic of the paragraph.
- D) No, because it summarizes the main point of the paragraph.

20. The writer is considering adding a conclusion? Choose the sentence which does NOT fit in the conclusion.

- A) As seen here, learning in groups provides numerous opportunities to practice various skills which are essential in the acquisition of the target language.
- B) While working in a group, students can learn from others, be less stressed, and are situated in a more supportive environment.
- C) Moreover, the quality of learning in group lessons is no different from private lessons, but it is quite apparent that the former is more cost-effective than the latter.
- D) So, when you are looking to learn a new language, you should always consider learning in groups as it has many benefits over individual private tutoring

Sample Diagnosis Sheet

