The Relationship between Peer Assessment and the Cognition Hypothesis

Mona Khabiri

Department of English, Sanat Sq., Simae Iran St.

Islamic Azad University, Tehran Central Branch, Tehran, Iran

E-mail: monakhabiri@yahoo.com

Soroush Sabbaghan (Corresponding Author)

Department of English, Hesarak, Science and Research Branch

Islamic Azad University, Tehran, Iran

Tel: 98-912-105-6615   E-mail: inproxima@yahoo.com

Sahar Sabbaghan

Department of English, Hesarak, Science and Research Branch

Islamic Azad University, Tehran, Iran

E-mail: sahar_sabbaghan@yahoo.com

Abstract

It is believed that peer assessment equips learners with a skill set withheld from them by teacher assessments that enhances language learning. However, the benefits of peer assessment are limited to how well learners can conduct peer assessment tasks. Therefore, improving the efficacy of peer assessment is essential. One way to increase the consistency of peer assessment is to increase learner attention during the assessment task. The Cognition Hypothesis states that L2 learners engaged in complex tasks pay attention to more complex linguistic structures; as a result, learning increases (Robinson, 2001a, 2001b, 2005). The purpose of this study was to investigate whether complex tasks, as outlined by the Cognition Hypothesis, improve the accuracy of peer assessment. Thirty female EFL learners conducted three speaking tasks. Each task had a different level of complexity, and participants were assessed by their peers using a rating scale. The results indicated that the absolute mean deviations for the items on the rating scale decreased as task complexity increased. In other words, the findings showed that as task complexity increased, there was more agreement among the assessors. This indicatedthat peer assessment wasmore accurate and consistent for more complex tasks.

Keywords: The Cognition Hypothesis, peer assessment, task complexity, EFL speaking assessment

Introduction

In any teaching environment, assessment is critical. In the last two decades, there have been conceptual shifts in the practice of assessment. These shifts have moved toward the involvement of the learner in the assessment practice (Boud, 1995). Peer assessment, in which learners assess the work of other learners, is a form of learning that allows learners to provide feedback on each other's work.

Numerous studies have supported the claim that peer assessment is beneficial for learning (see Ballantyne, Hughes, & Mylonas, 2002; Boud, 1990). Additional studies have suggested that peer assessment promotes reflective thinking through observation of other learners' performances, which in turn allows learners to understand the requirements of a classroom task (see Falchikov, 1986; Topping, 1998). Moreover, Birdsong and Sharplin (1986) demonstrate that peer assessment contributes to higher order reasoning. Peer assessment could also promote self-learning (Oldfield, Mark, & Macalpine, 1995) and deep learning (Entwhistle, 1987; Gibbs, 1992). Kwan and Leung (1996) suggest that peer assessment encourages cooperative group work. If students are taught through peer assessment as instruction tasks, satisfaction with the class increases (Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). In sum, there is little evidence that peer assessment elicits negative reactions in the learning process (see Cheng & Warren, 1997 for negative reaction).

The benefits of peer assessment in the EFL/ESL context is limited to the extent to which learners can implement peer assessment practices. One method of increasing peer assessment consistency is to train the learners. In the foreign language context, studies (Berg, 1999; Stanley, 1992) have shown that training learners in conducting peer assessment increases learning efficacy. However, McGroarty and Zhu (1997) found that training learners for peer assessment does not impact learners'final grades.

Increasing learner focus and attention during peer assessment could be another way to improve peer assessment practices. The Cognition Hypothesis states that requiring L2 learners to engage in complex tasks facilitates L2 learning by promoting interaction, focus on form, and attention to more complex linguistic structures (Robinson, 2001a, 2001b, 2005). The question is whethercomplex tasks that increase attention and focus, which facilitate learning, also increase attention and focus in peer assessment.

Robinson (2001a, 2001b, 2003, 2005)distinguishes three sources of cognitive demands in a language task: (a) task complexity, which refers to the cognitive factors that relate to how a task is designed; (b) task conditions, which refers to the interactional factors relating to participation (e.g., one-way vs. two-way); and (c) task difficulty, which refers to affective and learner ability variables (e.g., motivation). Within these sources of cognitive demands, Robinson identifies two dimensions for task complexity: resource-directing and resource-dispersing, as described in Table 1.

According to Table 1, Robinson's (Robinson, 2001a, 2001b, 2003, 2005, 2007a) two dimensions for task complexity includes cognitive/conceptual dimension (i.e., resource-directing) and performative/procedural dimension (i.e., resource-dispersing). Resource-directing variables require more attention, working memory, and cognitive functions that help learners to focus on linguistic forms. These variables are: [± few elements], [± here and now], and [± no reasoning demand]. As Table 1 shows, a less complex narration task requires [+ few elements], [+ here and now], and [+ noreasoning demand], but a more complex task requires [- few elements], [- here and now], and [- no reasoning demand].

Resource directing variables are those that necessitate the use of attentional and memory resources but do not direct learners to any particular linguistic forms (Robinson, 2001b, 2005). Increasing task complexity using resource directing components therefore attracts thelearner's attention to many non-linguistic areas of the L2. Examples of resource-dispersing factors include: [± planning], [± single task], and [± prior knowledge]. Low complexity conditions would consist of [+ planning], [+ single task], and [+ prior knowledge], but high complexity conditions would consist of [- planning], [- single task], and [- prior knowledge].

Many studies have tested Robinson's Cognition Hypothesis. For example, various degrees of complexity variables, such as [± no reasoning demand] (Nuevo, 2006), [± here and now] (Gilabert, 2005; Robinson, 1995; Robinson, Ting, & Urwin, 1995), [± single task] (Robinson, 2007b), and [± few elements] (Kuiken, Mos, & Vedder, 2005; Kuiken, Vedder, & Matters, 2007) have been investigated.

In sum, previous studies regarding the Cognition Hypothesis have focused on the influence of task complexity on L2 production. Most of these studies have concluded that complex tasks increase attention and focus on form, which in turn enhances L2 production. To date, no published study has investigated the effect of complex tasks on peer assessment. Given that peer assessment is beneficial to English language learning in the EFL/ESL context, improving this practice is essential. One way to do so is to increase learners' attention to peer assessment tasks. This may be accomplished by increasing task complexity. The purpose of this study is to investigate whether increasing task complexity increases the accuracyand consistency of peer assessment of L2 oral production.

Method

The participants of the study consisted of 30 female Iranian EFL learners. The participant selected for this study all took the Oxford Placement Test (Allen, 2004) and obtained a score between 120 and 134, which designated them as low intermediate users of English; this score range corresponds with ALTE (2009) B1 level. All participants were provided with a thorough explanation of the research, its purposes, and how the findings would be valuable to the field of English language teaching. All participants were free to leave the project at any time, and incentives were not provided for their participation.

Oral production tasks were selected because one of the directives of the study was to investigate consistency in peer assessment tasks within the limits imposed by practical classroom considerations. The three speaking tasks in the study were designed to be either simple or more complex by adding and/or removing resource-directing and resource-dispersing variables. The first and simplest task (Task 1) was a descriptive narration. The three selected topics were: (a) describe a great vacation, (b) describe a great roommate, and (c) describe a great restaurant. These topics were selected because the participants had previously carried out these tasks in their EFL courses. The distribution of the resource-directing and resource-dispersingvariables, as described in Table 2, makes this task less complex.

The topics in Table 2 require the learner to describe a person, an object, or an event. Therefore, the 'few elements' of the resource-directing variables isgiven a plus because the learner was required to describe only one object/person/event. Furthermore, descriptive tasks do not require reasoning, so the 'no-reasoning-demands'

variable is also given a plus. However, since the task requires a description of a person, event, or object in the past without a mutually shared context, a minus is given to the 'hereandnow' variable.

In the category of resource-depleting variables, a plus is given to 'planning' because the researchers allowed the participants to work in groups. Furthermore, a plus was given to the 'single task' variable because the participant only described the topic and was not required to answer any questions during the task. Finally, a plus was given to the 'prior-knowledge' variable because participants had at one time completed a task with a similar topic.

The second task (Task 2) was to make a persuasive speech on three topics. The topics included: (a) persuade someone to learn English, (b) persuade someone to buy a used car, and (c) persuade someone to lose weight. These topics were selected because they were novel topics for the participants. Table 3 describes how Task 2 is more complex than Task 1.

According to Table 3, the task complexity variable layout for Tasks 1 and 2 was similar except for two variables. Because the topics for the second task were persuasive and required reasoning, a minus was given to the 'no-reasoning-demands' variable in the resource-directing category. Also, because the topics were new to the participants, a minus was given to the 'prior-knowledge' variable. However, these topics do refer to events happening at the moment. For this reason, the 'here and now' variable was given a plus. In sum, because there is one less variable in Task 2 than in Task 1, it is assumed that Task 2 is more complex than Task 1.

The final task (Task 3) was a debate. The topics for this task included: (a) discuss the pros and cons of the quality of life in Iran and in other countries, (b) choose between two perfumes and decide which one to buy, and (c) decide whether it is better to be married or single. As with Task 2, these topics were new to the participants and had not been debated in their EFL courses.

The arrangement of variables for this task is almost identical to the arrangement of variables in Task 2, with one exception. During the course of the debate, the participants were asked to challenge and question the speaker. Therefore, the speaker not only had to persuade the other participants, but she also had to answer questions and remark on the comments of other participants. In other words, the speaker had to perform two tasks simultaneously. For this reason, the 'single task' variable was given a minus. Information on the level of the task complexity for Task 3 is provided in Table 4.

A modified version of Yamashiro and Johnson's (1997) rating scale was used to assess the performances of the speakers (see the Appendix). Yamashiro and Johnson assert that their rating scale can be used for peer assessment and self-assessment of public speaking skills. The rating scale is composed of four categories: (a) voice control, (b) body language, (c) content of oral presentation, and (d) effectiveness.

The category of voice control was further divided into the four sections of projection, pace, intonation, and dictation. Projection refersto the loudness of a speaker's voice. Pace indicates the rate of speaking. Intonation refersto the use of proper pitch patterns and pauses, and dictation pertainedto speaking clearly without mumbling or using an interfering accent.

The category called body language was divided into three sections. These sections were posture, eye contact, and gesture. Posture signifies standing up straight and looking relaxed. Eye contact refersto how much the speaker looksat the audience. Gesture indicates the speaker's use of suitable gestures and avoidance of distracting ones.

The category called content of oral presentation "has obvious parallels with academic essay writing" (Yamashiro & Johnson, 1997, p. 1). This category isdivided into three sections: introduction, body, and conclusion. Introduction refersto the speaker's inclusion of a thesis statement and attention getting devices. Body deals with the speaker's use of academic writing structures and transitions. Finally, conclusion refersto the speaker's inclusion of a restatement, or summation, and a closing statement.

The final category, effectiveness, is furtherdivided into the three categories of language use, vocabulary, and purpose. The original rating scale for this category includes a subsection called topic. However, since the topics were given to the speakers, this subcategory was omitted in this study. Language use refersto the use of grammatically correct sentences. Vocabulary referred to the speaker's use of words appropriate for the audience. Purpose was the degree to which a speaker was successful in completing the task that they were given during oral production.

The speaker's performance in the areas outlined by the subsections was rated on a five-point Likert scale. Possible scores ranged from one (needs work) to five (very good). The total scores of all of the ratings represented the speaker's ability in the speaking task.

Before the participants carried out the oral tasks, the researchers met with them and thoroughly explained the background and purpose of the study. The researchers explained the concepts on the rating scale and provided examples and demonstrations of how to use the rating scale. Then, the 30 participants were divided into three groups of 10, and each group was assigned to a separate class. This division of the participants was imposed by the language schoolwhere the study was being conducted.

The participants and the researchers met three times a week and data collection occurred over several weeks. Tasks were conducted in an order based on the level of difficulty. In other words, Task 1 was done first, then Task 2, and finally Task 3.

The procedure for Tasks 1 and 2 was similar. At the start of both Task 1 and Task 2, the participants were randomly put into groups of threeand four. Each group was given one of the topics described in Table 2 and Table 3. The group members were encouraged to discuss their topic. They were given a total of 25 minutes for this purpose. Then, a member from each group was randomly selected to give a presentation on the topic. During the presentation, asking questions or making comments was not allowed. After each presentation, all participants, with the exception of the speaker, were asked to assess the speaker's performance using the rating scale. The assessment sheets were collected after a member from each group had presented a topic.

The procedure for Task 3 had minor differences. For example, for the debate on the perfume topic, first, props (paper strips scented with different perfumes) were given to the group that debated the choice of the perfume (during the presentation, the paper strips were distributed to all participants). Second, during the presentation, the participants were encouraged to ask the speaker questions and make comments during herspeech. When the debate was over, the participants were asked to assess the speaker.

Each participant conducted three speaking tasks and was assessed by her peers. Therefore, each participant received three sets of scores, each set corresponding to a speaking task of a different level of complexity. To investigate whether participants grew more vigilant during the peer assessment task, the degree of agreement among peers for every subsection of the rating scale was calculated. This was accomplished by calculating the absolute mean deviation (AMD) of the scores. A small absolute mean deviation indicated that scores in a subsectionwere similar. For instance, an AMD of zero would show that all participants gave the same score for a particular subsection of the rating scale. Therefore, the AMD wasan indicator of the degree of agreement among participants who hadassessed a particular speaker.

The AMD of the scores awarded to each participant for each of the three tasks was calculated. Then, the Friedman test (a non-parametric repeated measures comparison test) was used to compare the scores. The Friedman test indicated whether AMD distributions from different tasks were statistically different. In other words, the test indicated whether the amount of agreement among the assessors was statistically significant. An ANOVA was not used because a Levene's homogeneity test revealed that the variances of the scores were significantly different.

Results

The AMD for each of the 13 items on the peer assessment rating scale was calculated. As mentioned before, the participants in the study were divided into three groups of 10. Therefore, in every class, nine peersassessed the performance of eachparticipantfor each of the three tasks. An example of this calculation is displayed in Tables 5, 6, and 7, which show the scores given by the nine participants to Student 2 oneach item of the rating scale for Task 1, Task 2, and Task 3, respectively.

To compare the AMDs of the three tasks for each speaker, the Friedman test was employed. Table 8 displays the results of the Friedman test.

Table 8 shows that the absolute mean deviations are significantly different for the three levels of task complexity for each participant except for student 4 ($p_{\text{all participantsexcept for student 4}} < .05$).

Table 9 displays the averages of the absolute mean deviations for each of the three levels of task complexity for each participant. Figure 1 displays the graphic representation of Table 9. As is displayed in Figure 1, the average of the AMD for each participant decreases as the complexity level rises.

Discussion

The results indicated that the AMDs of peer assigned scores decreased as task complexity increased. In other words, the consistency of peer assessment increased for more complex tasks. As mentioned before, small AMDs are an indication of a high degree of agreement among peer assessors.

The Cognition Hypotheses could explain this phenomenon, i.e., the AMDs decreased as the complexity of tasks increased and the reason could have been due to the fact that more complex tasks require more attention and

awareness. This increase in attention and awareness allowed the learners to be more accurate in their assessments. Thus, the results of this study support the claims of the Cognition Hypothesis.

Motivation is another factor that might have affected the results of the study. Studies conducted outside the field of foreign language learning (Campbell, 1988; Kernan, Bruning, & Miller-Guhde, 1994) have revealed the connection between performance motivation and task complexity. Within the field of language learning, studies have shown a connection between motivation, achievement, and effort (Chambers, 1998; Dornyei, 2002; Dörnyei, 1994; Williams & Bruden, 1997; Williams, Burden, & Al-Baharna, 2002). Some of these studies have demonstratedthat cognitively difficult tasks increase the desire for achievement; people therefore put more effort into these tasks, which in turn results in higher degrees of achievement. It may be the case that, in this study, more cognitively complex tasks increased motivation, which in turn increased the learners' precision in their assessments.

The practice effect might also have had a role in the outcome of the study. Practice effects occur when a participant in an experiment is able to perform a task and then perform it again at some later time. Generally, the practice effect allows the participants to become better at performing the task. In the data-gathering phase of the study, each participant assessed nine peers three times over several weeks. Therefore, the participants might have gradually gained expertise in assessing their peers with the rating scale.

In sum, several different factors might have influenced the outcome of the study. However, the researchers believe that an increase in learner attention and awareness, as predicted in the Cognition Hypothesis, resulted in the increased accuracy of peer assessors. As mentioned before, motivation could have affected the outcome of the study, but there are not at present any published studies that examine the relationship between task complexity, as defined by the Cognition Hypothesis, and motivation. Therefore, the effects of motivation on assessment were difficult to define in our study. Also, although the practice effect might have influenced the outcome, because the study was conducted over several weeks and because assessment did not take place every day (it occurred every fourth day), the practice effects should have diminished. Therefore, it is highly likely based on the cognition hypothesis (Robinson, 2005) that the increase in task complexity explains the increase in the precision of the peer assessments.

References

Allen, D. (2004). Oxford Placement Test. Oxford: Oxford University Press.

ALTE. (2009). Overall General Ability. Retrieved 7th June, 2009, from http://alte.org/can_do/general.php

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education, 27*(5), 427-441.

Berg, E. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing, 8*(3), 215-241.

Birdsong, T., & Sharplin, W. (1986). Peer Evaluation Enhances Students. *Highway One, 9*(1), 23-28.

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education, 15*(1), 101-111.

Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.

Campbell, D. (1988). Task complexity: A review and analysis. *Academy of Management Review, 13*(1), 40-52.

Chambers, G. (1998). Pupils' perceptions of the foreign language learning experience. *Language Teaching Research, 2*(3), 231.

Cheng, W., & Warren, M. (1997). Having second thoughts: Student perceptions before and after a peer assessment exercise. *Studies in Higher Education, 22*(2), 233-239.

Dornyei, Z. (2002). The motivational basis of language learning tasks. *Individual Differences and Instructed Language Learning, 2*, 137-158.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *Modern Language Journal, 78*(3), 273-284.

Entwhistle, N. (1987). A model of the teaching process. In J. T. E. Richardson, M. W. Eyesynck & D. W. Piper (Eds.), *Student learning: research in education and cognitive psychology*. Guildford: Open University Press.

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment & Evaluation in Higher Education, 11*(2), 146-166.

Gibbs, G. (1992). *Improving the Quality of Student Learning*. Bristol: Technical & Educational Services Ltd.

Gilabert, R. (2005). Task complexity and L2 narrative oral production. Unpublished Ph. D. dissertation. Universitat de Barcelona.

Kernan, M., Bruning, N., & Miller-Guhde, L. (1994). Individual and group performance: Effects of task complexity and information. *Human Performance, 7*(4), 273-289.

Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System, 37*(2), 254-268.

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen (Ed.), *Eurosla Yearbook* (Vol. 5, pp. 195-222): John Benjamins Publishing Company.

Kuiken, F., Vedder, I., & Matters, M. (2007). Cognitive task complexity and linguistic performance in French L2 writing. In M. P. Garcı´a-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 117–135). Clevedon: Multilingual Matters.

Kwan, K., & Leung, R. (1996). Tutor versus peer group assessment of student performance in a simulation training exercise. *Assessment & Evaluation in Higher Education, 21*(3), 205-214.

McGroarty, M., & Zhu, W. (1997). Triangulation in classroom research: A study of peer revision. *Language Learning, 47*(1), 1-43.

Nuevo, A. (2006). Task complexity and interaction: L2 learning opportunities and development. Unpublished Doctoral Dissertation. Georgetown University.

Oldfield, K., Mark, J., & Macalpine, K. (1995). Peer and Self-assessment at Tertiary Level--an experiential report. *Assessment & Evaluation in Higher Education, 20*(1), 125-132.

Robinson, P. (1995). Attention, memory, and the noticing hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). New York: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 32.

Robinson, P. (2007a). Criteria for classifying and sequencing pedagogic tasks. In M. d. P. G. Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 7-26). New York: Multilingual Matters.

Robinson, P. (2007b). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL, 45*(3), 193.

Robinson, P., Ting, S., & Urwin, J. (1995). Investigating second language task complexity. *RELC Journal, 26*(2), 62.

Sluijsmans, D., Brand-Gruwel, S., & van Merriënboer, J. (2002). Peer Assessment Training in Teacher Education: effects on performance and perceptions. *Assessment & Evaluation in Higher Education, 27*(5), 443-454.

Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing, 1*(3), 217-233.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249.

Williams, M., & Bruden, R. (1997). *Psychology for language teachers*. Cambridge: Cambridge University Press.

Williams, M., Burden, R., & Al-Baharna, S. (2002). Making sense of success and failure: The role of the individual in motivation theory. In Z. Dornyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (2 ed., pp. 171-184). Manoa: University of Hawai'i Press.

Yamashiro, A., & Johnson, J. (1997). Public speaking in EFL: Elements for course design. *The Language Teacher, 21*(4), 13-17.

Appendix

Public Speaking Class Peer Rating Sheet	
Speakers Name:_____ Presentation topic:_____ **Score scale: 5** (very good) **4** (good) **3** (average) **2** (weak) **1** (poor) Circle a number for each category, and then consider the numbers you chose to decide an overall score for the presentation.	
Voice Control **Rating**	
1. **Projection** (loud/soft)	
2. **Pace** (speech rate; fast/slow)	
3. **Intonation** (patterns, pauses)	
4. **Diction** (clear speaking)	
Body Language	
1. **Posture** (standing straight, relaxed)	
2. **Eye contact**	
3. **Gestures** (well used, not distracting)	
Contents of Presentation	
1. **Introduction** (grabs attention, has main points)	
2. **Body** (focused on main ideas, has transitions)	
3. **Conclusion** (summary of main points, closing statement)	
Effectiveness	
1. **Language use** (clear, correct sentences/slide information)	
2. **Vocabulary** (words well-chosen and used)	
3. **Purpose** (informative, teaches about topic)	
Total	

Table 1.Robinson's task complexity dimensions

Cognitive factors	Example
Resource-directing +/- few elements +/- no reasoning demands +/- here & now	Fewervs. more pictures to narrate Pictures presented in order of narration vs. not in order of narration Pictures present during narration vs. not present during narration
Resource-dispersing +/- planning +/- single task +/- prior knowledge	Narration with vs. without planning time Narrate a picture vs. narrate a picture and write a story Familiar vs. not familiar with the story plot

Note: Adopted from Kim (2009)

Table 2. Complexity variables oftask 1 (descriptive narration)

Topic	Resource-directing	resource-dispersing
describe a great vacation	+ few elements + no reasoning demands - here and now	+ planning + single task + prior knowledge
describe a great roommate	+ few elements + no reasoning demands - here and now	+ planning + single task + prior knowledge
describe a great restaurant	+ few elements + no reasoning demands - here and now	+ planning + single task + prior knowledge

Table 3.Complexity variables of task 2 (persuasive speech)

Topic	Resource-directing	resource-dispersing
persuade someone to learn English	+few elements - no reasoning demands + here and now	+ planning + single task - prior knowledge
persuade someone to buy a used car	+ few elements - no reasoning demands + here and now	+ planning + single task - prior knowledge
persuade someone to lose weight	+ few elements - no reasoning demands + here and now	+ planning + single task - prior knowledge

Table 4.Complexity variables oftask 3 (debate)

Topic	Resource-directing	Resource-depleting
quality of life in Iran and in other countries	+ few elements - no reasoning demands + here & now	+ planning - single task - prior knowledge
which of two perfumes would you buy	+ few elements - no reasoning demands + here & now	+ planning - single task - prior knowledge
benefit of being single or married	+ few elements - no reasoning demands + here & now	+ planning - single task - prior knowledge

Table 5. Peer scores and absolute mean deviation for student 2 on task 1

Topics	P1	P2	P3	P4	P5	P6	P7	P8	P9	mean	AMD
Projection	4	3	4	3	5	4	3	4	3	3.67	0.59
Pace	4	3	4	3	5	3	3	3	4	3.56	0.62
Intonation	5	5	5	2	5	4	4	5	5	4.44	0.74
Diction	3	2	4	4	4	5	4	3	4	3.67	0.67
Posture	2	5	2	4	5	2	2	3	3	3.11	1.04
Eye contact	3	3	3	4	5	3	5	3	3	3.56	0.74
Gestures	4	5	4	3	4	3	4	4	3	3.78	0.52
Introduction	2	4	3	3	3	3	3	3	3	3.00	0.22
Body	3	5	3	5	3	2	4	3	3	3.44	0.81
Conclusion	2	3	4	3	5	4	3	4	5	3.67	0.81
Language use	4	3	4	3	5	3	4	3	4	3.67	0.59
Vocabulary	4	3	4	3	3	5	3	4	3	3.56	0.62
Purpose	4	5	4	3	3	4	3	4	3	3.67	0.59

Note: p = peer; AMD = Absolute mean deviation.

Table 6.Peer scores and absolute mean deviation for student 2 ontask 2

Topics	P1	P2	P3	P4	P5	P6	P7	P8	P9	mean	AMD
Projection	3	3	4	3	3	4	3	3	3	3.22	0.35
Pace	3	3	4	3	3	3	3	2	3	3.00	0.22
Intonation	4	4	5	4	5	4	4	5	4	4.33	0.44
Diction	3	3	4	4	3	5	3	3	3	3.44	0.59
Posture	3	3	2	3	5	2	3	3	3	3.00	0.44
Eye contact	5	5	3	4	5	5	5	3	5	4.44	0.74
Gestures	4	4	4	4	4	3	4	4	3	3.78	0.35
Introduction	3	4	4	4	3	3	4	3	4	3.56	0.49
Body	3	5	3	2	2	2	4	2	3	2.89	0.79
Conclusion	3	3	3	3	5	4	3	2	2	3.11	0.62
Language use	3	4	4	3	4	3	4	4	4	3.67	0.44
Vocabulary	4	4	4	3	4	5	4	4	4	4.00	0.22
Purpose	4	4	4	3	4	4	3	4	3	3.67	0.44

Note: p = peer; AMD = Absolute mean deviation.

Table 7.Peer scores and absolute mean deviation for student 2 ontask 3

Topics	P1	P2	P3	P4	P5	P6	P7	P8	P9	mean	AMD
Projection	4	4	4	4	3	4	3	4	4	3.78	0.35
Pace	3	3	3	3	3	3	3	2	3	2.89	0.20
Intonation	3	3	3	4	3	4	4	3	4	3.44	0.49
Diction	4	3	4	4	4	4	4	4	3	3.78	0.35
Posture	5	4	4	4	5	4	4	5	5	4.44	0.49
Eye contact	5	5	5	4	5	5	5	5	5	4.89	0.20
Gestures	4	5	5	5	4	5	4	4	5	4.56	0.49
Introduction	4	4	4	4	3	3	4	4	3	3.67	0.44
Body	3	3	3	3	3	3	3	2	3	2.89	0.20
Conclusion	3	3	3	3	3	4	3	3	3	3.11	0.20
Language use	3	3	4	3	4	3	4	4	4	3.56	0.49
Vocabulary	4	4	3	3	4	3	4	4	3	3.56	0.49
Purpose	4	3	4	3	3	4	3	4	3	3.44	0.49

Table 8.Results of the Friedman test

Number	Name	df	Chi Sq	Asymp Sig
1	Student 1	13	14.88	0.001
2	Student 2	13	9.385	0.009
3	Student 3	13	12.923	0.002
4	Student 4	13	2.667	0.264
5	Student 5	13	23.50	0.000008
6	Student 6	13	12.745	0.002
7	Student 7	13	15.52	0.0004
8	Student 8	13	9.542	0.008
9	Student 9	13	5.216	0.074
10	Student 10	13	18.863	0.00008
11	Student 11	13	6.157	0.046
12	Student 12	13	14.56	0.001
13	Student 13	13	8.167	0.017
14	Student 14	13	16.12	0.00032
15	Student 15	13	11.692	0.003
16	Student 16	13	15.864	0.00036
17	Student 17	13	17.077	0.0002
18	Student 18	13	10.36	0.006
19	Student 19	13	14.217	0.001
20	Student 20	13	7.791	0.027
21	Student 21	13	19.792	0.00005
22	Student 22	13	10.500	0.005
23	Student 23	13	13.216	0.001
24	Student 24	13	19.889	0.00005
25	Student 25	13	22.615	0.00005
26	Student 26	13	13.161	0.001
27	Student 27	13	19.538	0.00006
28	Student 28	13	10.792	0.005
29	Student 29	13	16.442	0.0027
30	Student 30	13	8.51	0.015

Note: Chi sq = Chi Square; df = Degree of Freedom; Asymp Sig = Significant Value.

Table 9.Averages of absolute mean deviations for the three tasks

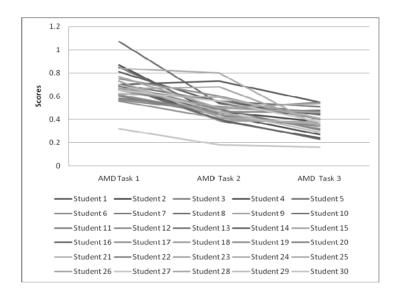| Name	AMD Task 1	AMD Task 2	AMD Task 3	Name	AMD Task 1	AMD Task 2	AMD Task 3
Student 1	0.73	0.39	0.24	Student 16	0.69	0.40	0.23
Student 2	0.66	0.47	0.38	Student 17	0.75	0.60	0.40
Student 3	0.70	0.49	0.31	Student 18	0.70	0.44	0.35
Student 4	0.87	0.44	0.27	Student 19	0.61	0.46	0.32
Student 5	1.07	0.54	0.41	Student 20	0.58	0.49	0.55
Student 6	0.85	0.46	0.48	Student 21	0.62	0.59	0.41
Student 7	0.66	0.56	0.45	Student 22	0.65	0.51	0.47
Student 8	0.62	0.59	0.40	Student 23	0.67	0.50	0.46
Student 9	0.56	0.41	0.36	Student 24	0.73	0.42	0.36
Student 10	0.70	0.73	0.55	Student 25	0.84	0.80	0.34
Student 11	0.60	0.44	0.34	Student 26	0.66	0.68	0.36
Student 12	0.57	0.47	0.46	Student 27	0.65	0.56	0.38
Student 13	0.70	0.56	0.51	Student 28	0.77	0.50	0.54
Student 14	0.81	0.55	0.44	Student 29	0.70	0.47	0.29
Student 15	0.69	0.39	0.38	Student 30	0.32	0.18	0.16

Figure 1. Mean of AMDs for tasks 1, 2, and 3