Reciprocal Contribution of Writing Attributes to One Another

Leila Ghaderpanahi[1], Sayyed Mohammad Alavi[2], Hossein Karami[3]

[1]Ph.D. Candidate, English Department, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, Email: leilaghaderpanahi991@yahoo.com

[2]Corresponding Author, Professor, English Department, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, Email: smalavi@ut.ac.ir

[3]Assistant Professor, English Department, Faculty of Foreign Languages and Literatures, University of Tehran, Tehran, Iran, Email: hkarami@ut.ac.ir

Abstract

Formative writing assessment can help writing instructors to explore weaknesses and strengths of language learners' writing performances. The current research aimed to explore firstly writing attributes and secondly examine their reciprocal contribution to one another. To achieve such an objective, the participants (N=200) were asked to write about two different topics. One writing sample before treatment was considered as the pre-test and the other after the treatment was considered as the post-test writing sample. Having scrutinized the pre-writing samples, five raters extracted the writing attributes which appeared in pre-test and post-test writing samples. Results indicate that there is a statistically significant difference among the participants' performances in terms of using writing attributes. The results can be advantageous for both instructors and syllabus designers to provide pedagogical materials which identify particular frailties and notify them about the more troublesome points to concentrate on in the classroom so as to arrange effective education.

Keywords: predictive validity, reciprocal contribution, writing ability, writing assessment, writing attributes

1. Introduction

Reviewing the attempts in clarifying the writing attributes within a wide variety of educational domains shows that using writing attributes has a solid conceptual foundation, though its implementation to the domain of language evaluation stands relatively untried (Alderson et al., 2014). Using writing attributes is expected to equip both learners and instructors with ample feedback required to establish prospective targets and overwhelm the frailties in various levels of learning via scaffolding and mediating. It seems that assessing the utilization of writing attributes can provide the learners and instructors with the type of awareness they require to arrange their reciprocal and personal objectives for upcoming development from a Zone of Proximal Development to another via sufficient scaffolding and mediating (Lantolf, 2000).

Although assessment of the utilization of writing attributes is not much implemented in actual pedagogic settings, they have been broadly investigated and consistently accentuated by a lot of language instructors (Ross, 1998). Therefore, writing attributes particularly appear to be appropriate to be incorporated in any language learning curriculum. Rupp et al. (2012) argue that the validity of any explanation and resolution about students expressed based on a data analysis concerns the plan, execution and the utilization of the writing attributes in the real-life framework.

It seems essential to continue scrutinizing the various constituents of writing ability so as to comprehend this paramount skill and assist learners in learning it. Therefore, if troublesome modules of the writing skill are diagnosed throughout the course of a semester, they can be ameliorated by allocating proper tasks in that module. It seems that in the literature related to writing assessment, researchers have rarely offered feedback on skills and attributes pertaining to writing ability that can conceivably be employed to boost learning.

Criticism resides in the principal aim of academic tests which present a quantitative assessment of a learner's general ability and proficiency as compared to other learners' in the normative group. This kind of norm-referenced testing has been employed broadly for ranking and selecting learners for different academic resolutions. These standards of evaluation just present general facts regarding learners' expertise and their capability to perform on a test and are not able to provide comprehensive information concerning learners' weaknesses and strengths which would probably assist them in enhancing their writing skill or that can even help the instructor in pedagogical preparation. Alderson (2010) expresses his dismay about the lack of writing attributes utilization assessments. Indeed, so far most of the researches have been done on large-scale evaluations rather than on writing attributes utilization. He maintains that far more researches have been devoted to employing proficiency tests than utilizing writing attributes in the form of formative or classroom-based evaluations. Researchers propose that writing attributes utilization assessment can have a crucial role in developing the instructional value of evaluation (Leighton & Gierl, 2007; Alderson, 2010; Rupp et al., 2012; de la Torre, 2009; Jang, 2005). Hence, the current study aims to employ writing attributes as a possible means to strengthen the students' writing abilities.

2. Literature Review

Writing ability is an essential tool for transferring information and development in education in the typical pedagogical milieu in particular and quotidian life in general. Thus, the notion of writing ability has been the pivot of study throughout the last few years in education, psychology and applied linguistics (Urquhart & Weir, 1998). Despite the considerable investigation on writing ability, there is controversy about how foreign language writing ability is defined and how its process can be examined and described. Alderson et al. (2015) believed that the productive skill of writing is not that hard to evaluate as the errors can be observed in learners' writings. However, there is a more desperate necessity for comprehending students' difficulties. Apparently, learners have not been provided with adequate feedback which can be employed for enhancements in writing ability.

Writing ability can be evaluated in terms of product and process approaches. As Alderson (2000) states, the writing product indicates the final result of writing and the writing process designates what occurs between the writer and the writing. The attributes of writing are examined on the basis of the product approach (Urquhart & Weir, 1998). Writing not only requires mechanical organization but also necessitates purposeful management (Kim, 2011). Hence, concentrating on the product or writing skills, the processes are not contemplated; however, considering the processes of writing for achieving the product is essential.

Examinations can present a total score of writing ability that fulfils the exam's immediate aim; however, they cannot be indicators of the writing development process. Merely providing an overall mark for an exam cannot offer the required data about learners' precise weaknesses and strengths. A comprehensive score account of the students based on writing attributes utilization, incorporating their presentation instead can be utilized to both develop writing ability and direct education (Snow & Lohman, 1989).

Although copious studies have been done on writing attributes for distinguishing learners' weaknesses and strengths, it appears that only a few experimental researches have concentrated on employing writing attributes for improvement of learners' writing abilities. The current study aims to employ writing attributes assessment at the university level to evaluate students' learning contributing to the development of writing abilities which therefore directs instructors teaching.

2.1. Writing Assessment

University instructors' classroom writing assessments seem to play a crucial role in their teaching and their students' learning. Therefore, instructors are suggested to consistently evaluate their students so as to help teaching and learning development (Fulcher & Davidson, 2007). Assessment reminds students of the pressure of exams which they must confront anxiously. Particularly in classrooms where competition rather than diagnosis is dominated, it requires the instructors' and learners' second thoughts to benefit from ongoing assessment of the writing attributes utilization rather than concentration on the final mark (Mercer et al., 2004).

Over the years writing has been evaluated by two propositions, i.e., direct assessment and indirect assessment. Direct assessments are those in which a sample of a learner's writing is taken under supervised conditions and consequently assessed by one or more examiners. Indirect assessments are those in which estimates of feasible skills in writing are made via consideration of particular sorts of knowledge about writing such as grammar and sentence structure. Therefore, direct assessments are associated with writing samples and indirect assessments with multiple-choice questions. Thus, in the current study, the direct assessment will be applied.

Apparently, classroom writing assessment is limited to the instructors' feedback given to the papers the learners are assigned to write independently and minimal diagnostic instruction and determined mediation are presented to the learners in the distressing need of learning writing skills particularly in a pedagogic setting. Diversely, the contemporary humanistic cooperative outlook to pedagogy suggests a kind of involvement in the problem of assisting learners in developing a more dynamic and accountable task for their own learning evaluation (Dewey, 2012). Despite the tendency to detect alternative types of evaluation to boost the validity and reliability of assessments, it seems that writing attributes utilization assessment has not been much implemented by instructors and learners at different levels (Taras, 2002). In Iran, for example, both instructors and learners appear to have trivial former experience of such sorts of options in assessment in the language classrooms as the assessment has customarily been the instructors' exclusive privilege and responsibility.

3. Research Questions

For the purpose of the above-mentioned investigation and comparison, the following research questions are posed:

1. Is there any statistically significant difference among the participants' performances on the facility in the use of language, inappropriate choice of words, accumulation of errors in sentence structure and serious errors in usage?

2. Is there any difference in the contribution of writing attributes to one another?

4. Method

4.1. Participants

200 undergraduate Iranian university EFL students participated in this study. The students attended the EFL class (17 sessions, 34 hours). The University of Tehran was selected as the research context for the study. The materials and the textbook chosen were authentic. The authentic text refers to a piece of language which is created by a native speaker of the language in which it is produced. The participants were randomly chosen. The participants' writings assigned during the course were rated by their instructors.

The examinees were classified into three distinct writing mastery ranks according to their attribute utilization as it is believed that more skilled writers tend to employ more writing attributes. Moreover, these attributes were classified into three groups as "more frequent, relatively frequent, and less frequent". For manageability reasons, the least frequent attributes in the writing samples were removed from the analysis. However, the less frequent attributes that could have a contribution to the more frequent attributes were kept.

4.2. Instruments

The participants of this study selected based on the feasibility of research in this area were asked to write about one general and one academic topic. A pre-test and post-test were used in this study. The subjects of the pre-test and post-test were decided on the basis of the subjects the participants had studied in their textbooks during the course. A treatment package was also developed based on the general writing diagnostic domains and diagnostic features which required the students to deliver a summary of the texts in their books to their instructors each session which consequently were corrected and returned to the students to read the feedback and ask their questions.

4.3. Procedure

The randomly selected students took a pre-test for their writing ability to be evaluated before they undergo the experiment. The conducted treatment was during one term of 17 weeks (each week one session of 90-minute classroom practice). The students underwent writing attributes utilization practices and experienced process writing method of teaching in which their instructors were engaged in giving feedback to the students. The students' writings were collected and rated by their instructors. Finally, the participants took a post-test and the results of their performances indicated the difficulty of the writing attributes.

4.4. Frequency of Attributes Used in Writing Samples

For attaining data concerning the allocation of the writing feedback, the total participants' (N=200) performance on the pre and post-tests was analyzed. The lowest and highest mean of the four major attributes in the pre-test were 48.25 (serious errors in usage) and 65.25 (accumulation of errors in sentence structure) respectively. Besides, the lowest and highest mean of the four major attributes in the post-test were 80.32 (serious errors in usage) and 88.70 (facility in the use of language) respectively. The results show the improvement of students' writing attributes performances as a result of being exposed to writing attributes and reject the null hypothesis. Table 1 shows the mean and the standard deviation of four attributes examined in this study.

Table 1

Descriptive Statistics of Four Attributes in Pre/Post Tests

		Pre test				Post test			
			Mean		SD		Mean		SD
		Value	Rank	Value	Rank	Value	Rank	Value	Rank
A1	facility in the use of language	62.25	2	18.90	3	88.70	1	9.03	4
A2	inappropriate choice of words	60.79	3	17.55	4	82.33	3	9.32	3
A3	accumulation of errors in sentence structure	65.25	1	23.49	1	85.33	2	10.90	1
A4	serious errors in usage	48.25	4	19.18	2	80.32	4	10.38	2

Attribute 4, serious errors in usage, was the least frequent attribute in the pre and post-tests. While Attribute 2, inappropriate choice of words, was the most frequent one in the pre-test and Attribute1, facility in the use of language, was the most frequent one in the post-test. The lowest and highest standard deviation of the four major attributes in the pre-test were 17.55 (inappropriate choice of words) and 23.49 (accumulation of errors in sentence structure) respectively. Besides, the lowest and highest standard deviation of the four major attributes in the post-test were 9.03 (facility in the use of language) and 10.90 (accumulation of errors in sentence structure) respectively. The results not only indicate lower standard deviations in the post-test but also smaller differences among attributes compared to the ones in the pre-test demonstrating the homogeneity of scores in the post-test. Consequently, the results show that there is a statistically significant difference among the participants' performances in terms of the above mentioned four major writing attributes.

Comparison of the major four attributes scores also shows that the attribute "serious errors in usage" in the pre-test is significantly different form other pre-test attributes and the attributes "facility in the use of language, inappropriate choice of words and accumulation of errors in sentence structure" are not significantly different from each other. Besides, the attribute "facility in the use of language" in the post-test is significantly different form other post-test attributes and the attributes "inappropriate choice of words, accumulation of errors in sentence structure and serious errors in usage" are not significantly different from each other. The differences in the subsets (Appendix A) is presented in table A-1.

As the participants of this study were not homogenous, the univariate analysis of variance showed that administration of pre-test and presentation of the feedback to the students had a significant effect on the treatment and post-test results. The results show a significant difference among pre A1 to A4 (A1: facility in the use of language, A2: inappropriate choice of words, A3: accumulation of errors in sentence structure, A4: serious errors in usage), the four attributes, ($f_{(1, 791)}=846.55$, p<0.000). The eta squared is high enough to allow for generalizing the findings. However, the difference in the performance of male and female is not statistically significant ($f_{(1, 791)}=1.70$, p=0.19) and the eta squared is low and does not allow for generalizing the findings. The interaction among A1 to A4 (A1: facility in the use of language, A2: inappropriate choice of words, A3: accumulation of errors in sentence structure, A4: serious errors in usage) in post-test and sex is not significant ($f_{(3, 791)}=1.94$, p=0.12); the eta squared is low and does not

allow for generalizing the findings. The univariate analysis of variance (Appendix A) is presented in table A-2. Analysis for looking at pre-test as a covariate of post-test scores is presented in table 2.

Table 2

Descriptive Statistics

fac.sex. four times	fac.AlpostA1.2.A4	Mean		Std. Deviation		N
		Value	Rank	Value	Rank	
female	Post A1	89.5556	1	7.99876	4	135
	Post A2	81.2346	3	9.48063	2	135
	Post A3	85.0617	2	10.82694	1	135
	Post A4	79.0476	4	8.99284	3	135
male	Post A1	86.9231	1	10.74172	3	65
	Post A2	84.6154	3	8.61891	4	65
	Post A3	85.8974	2	11.11445	2	65
	Post A4	82.9670	1	12.48036	1	65

4.5. Writing Attributes

The writing ability and its attributes were determined on the basis of a language ability blueprint as it was a case of language use in which via writing assignments, the learners' writing ability interrelated with their performance (Bachman & Palmer, 1996; Purpura, 2004). The attributes designated at least by five raters were regarded as crucial. Writing ability includes language knowledge affected by strategic competence for language use to emerge which is essential for student's writing ability interconnection with the written assignment (Bachman & Palmer, 1996). As Hartz (2002) suggests, the attributes that are employed by fewer than three learners do not present analytically consequential data and hence can be integrated with analogous attributes or removed. The nine attained main general attributes were facility in the use of language, inappropriate choice of words, accumulation of errors in sentence structure, serious errors in usage, unity, style, conventions, cohesion and coherence and organization. By virtue of the complex essence of writing, copious writing attributes are engaged in learners' productions (Urquhart & Weir, 1998; Alderson, 2000), as occurred in the current research. The list of writing attributes elicited from the pre-test and post-test papers as demonstrated in Table 3 was designated with regard to the previous literature, and the rater's judgment.

Table 3

Attributes of Writing Ability

	Writing Attributes
A1	facility in the use of language
A2	inappropriate choice of words
A3	accumulation of errors in sentence structure
A4	serious errors in usage
A5	unity
A6	style
A7	conventions
A8	cohesion and coherence
A9	organization

Having standardized the scores of the participants in these four major groups of attributes, a univariate analysis was performed to find the statistically significant difference in the existence of these attributes in the pre and post writing samples of male and female language learners. The tests of between-subjects effects in the pre and post-tests of the four mentioned attributes (Appendix A) is presented in table A-3. The results show a significant difference among A1 to A4 (A1: facility in the use of language, A2: inappropriate choice of words, A3: accumulation of errors in sentence structure, A4: serious errors in usage), the four attributes, ($f_{(7, 1584)} = 147.88$, $p < 0.000$). The eta squared is high enough to allow for generalizing the findings. The difference in the performance of males and females is also statistically significant ($f_{(1, 1584)} = 3.96$, $p = 0.04$); however, the eta squared is low and does not allow for generalizing the findings. The interaction among A1 to A4 (A1: facility in the use of language, A2: inappropriate choice of words, A3: accumulation of errors in sentence structure, A4: serious errors in usage) in pre and post-tests and sex is also significant ($f_{(7, 1584)} = 4.85$, $p < 0.000$); however, the eta squared is low and does not allow for generalizing the findings.

5. Results

5.1. Attributes in Writing Assessment

Quantifiable abilities of foreign language writing in diagnostic methods have been widely investigated so far. There are various construct explanations but the suggested replicas are not completely similar. Writing assessment has been mostly done in two types: holistic and analytical. "In analytic writing scripts are rated on several aspects of writing criteria rather than given a single score. Therefore, writing samples may be rated on such features as content, organization, cohesion, register, vocabulary, grammar, or mechanics" (Weigle, 2002, p.114). This practice assists in producing useful diagnostic input concerning examinees' writing abilities, which is the crucial advantage of analytic plans (Gamaroff, 2000). On a holistic scale, diversely, a single mark is assigned to the whole written texts. "The underlying assumption is that in holistic marking raters will respond to a text in the same way if a set of marking benchmarks are to guide them in marking" (Weigle, 2002, p.72). Among the writing attributes assessed in the writing assessment of EFL students in academic contexts, 30 attributes were detected in the participants pre-test and post-test writings for the purpose of this study.

5.2. Writing Attribute Performances

The attributes were classified into nine general groups, i.e., facility in the use of language, inappropriate choice of words, accumulation of errors in sentence structure, serious errors in usage, unity, style, conventions, cohesion and coherence and organization. 21 of the 30 attributes were assigned to the four attributes of 1. facility in the use of language, 2. inappropriate choice of words, 3. accumulation of

errors in sentence structure and 4. serious errors in usage. Table 2 shows the categories of general and specific attributes.

Table 2

General and Specific Attributes

No.	General attributes	Specific attributes
1	facility in the use of language	redundancy, layout, auxiliary verb, adverb, negative choice
2	inappropriate choice of words	verb choice, vocabulary choice, subject choice, gerund, relative adjective, adjective
3	accumulation of errors in sentence structure	subject verb agreement, active passive choice, infinitive
4	serious errors in usage	past participle, tense, article, pronoun, part of speech, relative pronoun, connector

5.3. Predictive Power of Writing Attributes

A multiple regression analysis was performed to explore the predictive power of writing attributes. The results of the pre-test analysis (table A-4) show that the attribute "facility in the use of language" has the predictive power to the attributes "inappropriate choice of words and serious errors in usage" (see appendix A).

Facility in the use of language=52.47 + (0.12) (inappropriate choice of words)

– (0.28) (accumulation of errors in sentence structure)

+ (0.34) (serious errors in usage)

Table A-5 shows that in the pre-test, the attribute "inappropriate choice of words" has the predictive power to the attributes "accumulation of errors in sentence structure, serious errors in usage and facility in the use of language" (see appendix A).

Inappropriate choice of words=52.87 + (0.27) (accumulation of errors in sentence structure)

+ (0.35) (serious errors in usage)

+ (0.10) (facility in the use of language)

Table A-6 shows that in the pre-test, the attribute "accumulation of errors in sentence structure" has the predictive power to the attributes "inappropriate choice of words and serious errors in usage" (see appendix A).

Accumulation of errors in sentence structure=42.22 - (0.24) (facility in the use of language)

+ (0.28) (inappropriate choice of words)

+ (0.31) (serious errors in usage

Table A-7 shows that in the pre-test, the attribute "serious errors in usage" has the predictive power to the attributes "facility in the use of language, inappropriate choice of words and accumulation of errors in sentence structure" (see appendix A).

Serious errors in usage=-5.16 + (0.26) (facility in the use of language)

+ (0.33) (inappropriate choice of words)

+ (0.27) (accumulation of errors in sentence structure)

The results of the post-test analysis (table A-8) show that the attribute "facility in the use of language" has the predictive power to the attribute "inappropriate choice of words" (see appendix A).

Facility in the use of language= 64.96 + (0.47) (inappropriate choice of words)

- (0.16) (accumulation of errors in sentence structure)

- (0.03) (serious errors in usage)

Table A-9 shows that in the pre-test, the attribute "accumulation of errors in sentence structure" has the predictive power to the attributes "inappropriate choice of words and serious errors in usage" (see appendix A).

Inappropriate choice of words=8.20 + (0.35) (facility in the use of language)

  + (0.34) (accumulation of errors in sentence structure)

  + (0.23) (serious errors in usage)

Table A-10 shows that in the post-test, the attribute "accumulation of errors in sentence structure" has the predictive power to the attributes "inappropriate choice of words and serious errors in usage" (see appendix A).

Accumulation of errors in sentence structure=41.07 - (0.14) (facility in the use of language)

+ (0.40) (inappropriate choice of words)

+ (0.24) (serious errors in usage)

Table A-11 shows that in the post-test, the attribute "serious errors in usage" has the predictive power to the attributes "inappropriate choice of words and accumulation of errors in sentence structure" (see appendix A).

Serious errors in usage=35.90 - (0.03) (facility in the use of language)

  + (0.29) (inappropriate choice of words)

  + (0.26) (accumulation of errors in sentence structure)

The results showed that the attribute "serious errors in usage" has a predictive contribution to all attributes but a negative contribution to the attribute "facility in the use of language".

6. Discussion and Conclusion

Writing plays an important role in the academic life of learners. Of the four major areas of communication skills and language development- listening, speaking, reading, and writing- the one that seems to be the most difficult to practice is writing (Jaramillo & Medina, 2011). Production skills are believed to be the last ones to be developed and the most challenging in terms of accuracy. When it is time for students to learn to write, they still have to listen so that they gain knowledge and information to follow directions, have spoken about different subjects and read a lot of texts to be able to gather enough input for producing an acceptable amount of output. Based on suggested regulations that are specific in any discipline, university students need to be able to write academic articles to share their achievements with their counterparts in the academic world. Issues related to writing in general and writing skills, in particular, have been widely explored. However, an examination of writing from the perspective of writing attributes utilization appears to be an unexplored problem in writing research.

Two facets of writing were examined in the current study. First was considering the writing attributes that are required for improving the writing ability. The raters pinpointed different attributes by investigating the writing attributes derived from the pre-test and post-test writing feedback. Second, the examinees' performance in the writings were analyzed for diagnostic aims and the scores were examined.

The results of the research indicate that four writing attributes "facility in the use of language, inappropriate choice of words, accumulation of errors in sentence structure and serious errors in usage" can strongly differentiate between professionals and non-professionals. The frequency of writing attributes in the pre-test and post-test was used to determine a recurring pattern among the writing attributes.

The participants' performances on the attributes employed in the writings were assessed. Roussos et al. (2007, p. 293) suggest, "A key issue for mastery/non-mastery of diagnostic models is whether the proportion of examinees estimated as masters on each skill is relatively congruent with the user's expectations." Having investigated the associations among the examinee's presentations on the writings, it was designated that the masters of attributes surpassed the non-masters of attributes suggesting considerable comparability between the determined pattern and detected information, showing a powerful diagnostic strength of the pattern.

"Inappropriate choice of words" was regarded as the hardest attribute with .455 and "facility in the use of language" with .630 the easiest attribute that is in conjunction with the view that language use related to recognizing word meaning requires lower-level organization (Alderson, 2000). In general, according to the results of the present study, the participants performed successfully on the four specified attributes (i.e., facility in the use of language, inappropriate choice of words, accumulation of errors in sentence structure and serious errors in usage) by virtue of the nature of the attribute that necessitates reasonable diagnostic processing.

The current research likewise evaluated the weaknesses and strengths of participants in three distinct writing mastery classifications: beginner, intermediate, and advanced. Results depicted that each category displayed different proficiency models of the writing attributes. The beginners had very low attribute proficiency possibilities; less than almost 20% of the beginners had mastered each attribute. The intermediates presented better than the beginners and indicated a broad span of writing attribute proficiency possibilities, varying from almost 21% to 79%. The advanced participants had very high proficiency possibilities of every writing attribute alternating more than 80%. Comparing the attributes among the three mastery classifications and the total category, the attribute "facility in the use of language" had the highest proficiency among all categories and the attribute "inappropriate choice of words" had the lowest mastery among all classifications.

Consequently, on the basis of the three writing mastery categories' attribute proficiency models, it was feasible to deduce their weaknesses and strengths in writing. The beginners and intermediates' proficiency of attributes were disproportionate. Namely, they had high proficiency over specific attributes and lower proficiency over the rest, distinctly indicating their weaknesses and strengths. The advanced participants displayed high proficiency possibilities across almost every attribute; in other words, they mastered approximately every attribute and did not seem to have particular weaknesses in writing.

The comprehensive score accounts of the current research appear worthwhile in assisting learning on the part of the student and in instructor empowerment and syllabus development on the part of the instructor. With thorough accounts of exam results, instructors can be well informed of learners' troublesome points and concentrate on them in lesson planning and learning material provision. As a result of practicing writing attributes, the instructors can alter, substitute and facilitate the difficult points for the students progressively. Therefore, a task bank could be established for writing attributes in this study and those in future.

6.1. Final Remarks

The current research presented some suggestions for instructors and professionals, both theoretically and didactically. The most important is the didactic indication for evaluation aims in order to employ writing attributes to measure writing mastery. The necessity for using exams on the basis of diagnostic blueprints is getting progressively essential in the domain of evaluation and language assessment. While presenting diagnostic feedback to the students which involves attribute proficiency possibility for them, instructors can benefit from the data to improve and modify the writing material and the type of assessment they use in their classrooms.

Since the development of every study encounters some limitations, the current research may undergo some as well. As designating the attributes requires a profound apprehension of the nature of diagnostic expertise, the complication of the writing ability does not permit a comprehensive mastery of its diagnostic procedures (Lee & Sawaki, 2009). Furthermore, there is the absence of concurrence on the expertise constituents of writing (Alderson, 2000). Therefore, even though a large number of attributes might be recognized, not all attributes could be used. Therefore, the aim is not to recognize every attribute which can be included in the writing but to analyze the main attributes necessary to successfully accomplish the tasks. As a result, the writing attributes are not comprehensive but are particularly associated with the writing exam. Simultaneously, the attributes which were recognized for the current research were not all applicable. Hence, investigation of extra attributes to be employed is recommended for future studies. Employing writing attributes demands more precise studies and doing more research can develop the accuracy of results.

One limitation of the study is that in the current research, five content specialists rated the writing attributes that is adequate for the aim of the research, but having a large number of specialists can probably make the attributes more applicable (Jang, 2005). Another limitation is that because of the limited number of participants in the study and their particular learning situations, generalizability is limited. The students studied in the research are limited to 200 Iranian university EFL students and all the students in the current study are currently enrolled in intensive English for Specific Purpose programs at a university.

Though this research was a trial of writing attributes utilization, it displays the necessity for progressive study in the domain of the Iranian context. Therefore, now is the time to concentrate on planning and improving academic evaluations on the basis of writing attributes utilization blueprint. Conducting such a venture requires the collaboration of different specialists from various fields. By pursuing such a goal, academic evaluations will be more pedagogically-aligned and more applicable to the needs of present-day classrooms.

References

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Alderson, J. C. (2010). Cognitive diagnosis and q-matrices in language assessment: A commentary. *Language Assessment Quarterly*, *7*(1), 96-103.

Alderson, J. C., Brunfaut, T., & Harding, L. (2014). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics*, *36*(2), 236-260.

Alderson, J. C., Haapakangas, E. L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2015). *The diagnosis of reading in a second or foreign language*. Routledge.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Routledge.

de la Torre, J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, *33*(3), 163-183.

Dewey, J. (2012). *Democracy and Education*. Pensylvania: Electronic Classics Series. DIALANG. www.lancaster.ac.uk/researchenterprise/dialang/about.htm/

DiBello, L., & Stout, W. (2008). Arpeggio documentation and analyst manual. *Chicago: Applied informative assessment research enterprises (AIARE)—LLC*.

Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.

Gamaroff, R. (2000). Rater reliability in language assessment: The bug of all bears. *System*, *28*(1), 31-53.

Hartz, S. M. (2002). A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 63(2-B), 864.

Jang, E. E. (2005). *A validity narrative: Effects of reading skills diagnosis on teaching and learning in the context of NG-TOEFL* [Unpublished doctoral dissertation]. University of Illinois at Urbana Champaign. Available from ProQuest Dissertations and Theses database. (AAT 3182288)

Jaramillo Urrutia, L., & Medina Gutiérrez, A. S. (2011). Adolescents' awareness of environmental care: Experiences when writing short descriptive texts in English. *Profile: Issues in Teachers' Professional Development*, *13*(1), 11-30.

Kim, H. S. J. (2011). *Diagnosing examinees' attributes-mastery using the Bayesian inference for binomial proportion: A new method for cognitive diagnostic assessment* [Unpublished doctoral dissertation]. Georgia Institute of Technology.

Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 1-26). Oxford University Press

Lee, Y. W., & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, *6*(3), 239-263.

Leighton, J. P., & Gierl, M. J. (2007). *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press.

Mercer, N., Dawes, L., Wegerif, R. & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Journal*, *30*(3), 359–377.

Purpura, J. E. (2004). *Assessing grammar*. John Wiley & Sons, Inc.

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, *15*(1), 1-11.

Roussos, L. A., DiBello, L. V., Stout, W. F., Hartz, S. M., Henson, R. A., & Templin, J. H. (2007). The fusion model skills diagnostic system. In J. Leighton & M. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 275–318). Cambridge University Press.

Rupp, A. A., Templin, J., & Henson, R. A. (2012). *Diagnostic measurement: Theory, methods, and applications*. Guilford Press.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational measurement* (pp. 263–331). Macmillan Publishing Co, Inc; American Council on Education.

Taras, M. (2002).Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education*, *27*(6), 501. http://dx.doi.org/10.1080/0260293022000020273

Urquhart, S., & Weir, C. J. (1998). *Reading in a second language: Process, product and practice.* Longman.

Weigle, S.C. (2002). *Assessing writing*. Cambridge University Press.

APPENDIX A: Writing Attribute Performances

Table A-1. Homogeneous Subsets

fac.Alpr.posA1.2.A4	N	Subset			
		1	2	3	4
Pre A4 (serious errors in usage)	200	48.2500			
Pre A2 (inappropriate choice of words)	200		60.7917		
Pre A1 (facility in the use of language)	200		62.2500		
Pre A3 (accumulation of errors in sentence structure)	200		65.2500		
Post A4 (serious errors in usage)	200			80.3214	
Post A2 (inappropriate choice of words)	200			82.3333	
Post A3 (accumulation of errors in sentence structure)	200			85.3333	
Post A1(facility in the use of language)	200				88.7000
Sig.		1.000	.318	.171	.701

Means for groups in homogeneous subsets are displayed.

Based on observed means.

The error term is Mean Square (Error) = 243.052.

a. Uses Harmonic Mean Sample Size = 200.000.   b. Alpha = .05.

Table A-2. Univariate Analysis of Variance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	49418.830[a]	8	6177.354	131.074	.000	.570
Intercept	307306.128	1	307306.128	6520.579	.000	.892
AlpreA1.2.A4	39896.834	1	39896.834	846.551	.000	.517
Fac.sex.4	80.466	1	80.466	1.707	.192	.002
fac.AlpostA1.2.A4	3313.560	3	1104.520	23.436	.000	.082
Fac.sex.4 * fac.AlpostA1.2.A4	274.757	3	91.586	1.943	.121	.007
Error	37278.767	791	47.129			
Total	5754641.270	800				
Corrected Total	86697.596	799				

Table A-3. Tests of Between-Subjects Effects in Pre and Post-tests

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	301640.180[a]	15	20109.345	82.737	.000	.439
Intercept	7266969.584	1	7266969.584	29898.884	.000	.950
A1 to A4 in pre and post	251605.352	7	35943.622	147.885	.000	.395
sex	964.138	1	964.138	3.967	.047	.002
A1 to A4 in pre and post * sex	8266.705	7	1180.958	4.859	.000	.021
Error	384993.622	1584	243.052			
Total	8901442.800	1600				
Corrected Total	686633.802	1599				

a. R Squared = .439 (Adjusted R Squared = .434)

A1: facility in the use of language, A2: inappropriate choice of words, A3: accumulation of errors in sentence structure, A4: serious errors in usage

Table A-4. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	52.473	4.919		10.667	.000
	PreA2. Inappropriate choice of words	.135	.084	.126	1.611	.109
	PreA3. Accumulation of errors in sentence structure	-.226	.059	-.281	-3.818	.000
	PreA4. Serious errors in usage	.338	.076	.343	4.430	.000

a. Dependent Variable: PreA1.facility in the use of language

Table A-5. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	25.877	4.887		5.295	.000
	PreA3. Accumulation of errors in sentence structure	.204	.050	.273	4.098	.000
	PreA4. Serious errors in usage	.323	.064	.353	5.077	.000
	PreA1.facility in the use of language	.097	.060	.104	1.611	.109

a. Dependent Variable: PreA2. Inappropriate choice of words

Table A-6. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	42.221	6.526		6.470	.000
	PreA1.facility in the use of language	-.306	.080	-.246	-3.818	.000
	PreA2. Inappropriate choice of words	.386	.094	.289	4.098	.000
	PreA4. Serious errors in usage	.385	.089	.314	4.328	.000

a. Dependent Variable: PreA3. Accumulation of errors in sentence structure

Table A-7. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-5.169	5.504		-.939	.349
	PreA1.facility in the use of language	.269	.061	.265	4.430	.000
	PreA2. inappropriate choice of words	.360	.071	.330	5.077	.000
	PreA3. accumulation of errors in sentence structure	.227	.052	.278	4.328	.000

a. Dependent Variable: PreA4. serious errors in usage

Table A-8. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	64.962	6.178		10.515	.000
	PostA2. inappropriate choice of words	.459	.073	.474	6.299	.000
	PostA3. accumulation of errors in sentence structure	-.134	.062	-.161	-2.143	.033
	PostA4. serious errors in usage	-.034	.064	-.039	-.527	.599

a. Dependent Variable: PostA1. facility in the use of language

Table A-9. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.205	6.875		1.193	.234
	PostA1. facility in the use of language	.366	.058	.355	6.299	.000
	PostA3. accumulation of errors in sentence structure	.293	.052	.343	5.616	.000
	PostA4. serious errors in usage	.207	.055	.230	3.759	.000

a. Dependent Variable: PostA2. inappropriate choice of words

Table A-10. Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	41.071	8.251		4.978	.000
	PostA1. Facility in the use of language	-.172	.080	-.142	-2.143	.033
	PostA2. Inappropriate choice of words	.473	.084	.404	5.616	.000
	PostA4. Serious errors in usage	.256	.070	.244	3.665	.000

a. Dependent Variable: PostA3. Accumulation of errors in sentence structure

Table A-11. Coefficientsa

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	35.904	8.270		4.341	.000
	PostA1. facility in the use of language	-.042	.080	-.037	-.527	.599
	PostA2. inappropriate choice of words	.325	.087	.292	3.759	.000
	PostA3. accumulation of errors in sentence structure	.250	.068	.263	3.665	.000