

WRITING QUALITY OF NON-NATIVE MA THESIS: PERCEPTIONS AND PRACTICES OF EXAMINERS AND STUDENTS

Majid Fatahipour

*Islamic Azad University, Parand, Iran
majifata@piaau.ac.ir*

Majid Nemati

*Islamic Azad University, Parand, Iran
nematim@ut.ac.ir*

Received: 30 April 2016

Accepted: 7 June 2016

ABSTRACT

Amongst extensive research on thesis writing quality, few provide a detailed account of common writing challenges/errors as perceived by examiners vs. students. In this study, eight recent English theses marked-up by examiners and defended in Iranian universities were selected randomly among those thesis made available (n=45). Based on an appraisal of comments written by examiners, a set of common challenges were drawn, which was confirmed by observations made by the researcher in viva voce sessions and then inspired questions about thesis writing quality. When administered to twenty examiners and seventy students, the most recurring themes in their practice were confirmed. The results draw a more precise picture, than previously available, of what examiners actually expect of a thesis and found four major areas of challenges in thesis writing, classified under lack of clarity in explaining, thesis formatting, grammatical errors and organisational inconsistencies. Findings also discussed the most significant practices in writing each chapter and reveal differences in perceptions of students/examiners.

Keywords: *writing quality, thesis examination; examiner expectations; writing challenges; research training*

INTRODUCTION

Writing theses is an important showcase of excellence in all fields of study, and the expectations are high particularly in the context of Teaching English as a Foreign Language (TEFL) since many Iranian students are/will be either teachers of English or have their jobs related to English teaching. Therefore, postgraduate students and their supervisors and/or examiners of their writing are anticipated to be aware of high standards of writing. These concerns make the quality of dissertation writing in TEFL even more serious than other fields of study. In second language writing research, few studies have focused on exploring thesis writing in contrast to essay writing. There are at least two important reasons for the neglect of thesis text as a unit of research, one being less accessible than other publications (books and articles) and another, is the huge size of the thesis as unit of analysis.

Quality of Second Language Writing is an important expectation in university settings, especially in writing a postgraduate thesis. Examiners and postgraduate students writing their thesis have differing perceptions and expectations of writing quality. Their relationship is comparable to expert and novice writers, or native and non-native writers. Understanding the perceptions and practices of students is vital because it can set a localised plan of corrective action towards writing errors. The examiners' practices are also important and thus quality of writing is determined according to their perceptions and expectations. Thus, a double amount of challenge can be envisaged for English-related majors in non-English-speaking countries who should write in English as a second or foreign language, and their first lengthy piece of serious writing is their MA thesis in English. Moreover, students at postgraduate level are expected to produce similar levels of writing quality as native speakers because academic writing of dissertations or theses cannot be compromised because of non-native status of the writer. Moreover, second language (henceforth L2) writing quality is often assumed to be a key factor for success in career prospects as well as everyday work.

Despite such evidence of the importance of L2 writing, comparatively little research has been conducted in this area. It is given that any writing task in a second or foreign language can become very challenging if they are lengthy and there are high expectations. Writing a thesis is an inevitable writing task that meets both of the above-said criteria of difficulty. Moreover,

writing a thesis is of an urgent nature because postgraduate students should write one in a limited time and with guidance. In Iran, the MA students are only allowed 6 months to complete a thesis with a possibility of six months as extension. Most supervisors have many postgraduate students and extensive hours of teaching commitments which are higher than standard practices elsewhere due to increase in population at higher education. Thus, writing a MA thesis is one of the most challenging tasks non-native writers would face in their academic life. Therefore, as expected, many EFL/ESL writers commit errors in writing and finding, and treating those errors from the examiners' point of view could be a contribution to failure in postgraduate education. This study sheds some light on the differences between examiners and postgraduate students perceptions by clarifying their criteria for judgments of quality of writing.

Rationale for the study

The quality of writing reflects the quality of thesis itself. Most research in the area of thesis quality has been conducted in English as the native language of the students, e.g. in Australia, Mullins and Kiley (2002) studied a sample of 30 experienced examiners, or Holbrook et al. (2004) who focused on corpus of examiner reports. Their method was followed by Stracke and Kumar (2010) in a similar study. Examiners in Iran are not required to formally submit detailed reports before viva. Rather, they fill in checklists, authorise viva sessions to be held and then write most of their detailed comments on the margin of thesis apart from examining during viva.

This study situates itself in the Iranian context of booming higher education, especially the recent sharp rise in postgraduate and MA level student intake. In this context, as well as many similar contexts, the examiners comment according to the indicators of thesis quality, on which they also base their recommendations.

Research aims and objectives

The objectives of the study are divided into general and specific. The overall goal of the research is to find out the predictors of quality by

classifying and categorising the errors that affect the quality of L2 writing, dissertations in this case, from the point of view of examiners as well as students. Two major components of these specific objectives are the views expressed and actions conducted by the examiners and students. The views are aimed to be elicited from the observations of the examiners drawing on their knowledge as well as experience and the actions through actual comments and notes written by the examiners on the submitted drafts of theses and subsequently through corrections made by the students on the final version of theses. Being an examiner is said to be a big responsibility with little promise, according to Pearce (2005: 1). This study is thus ultimately aimed at improving understanding between students and their examiners in Pearce's (2005: 2) terms 'deflated' or 'dispirited'. A study in a similar context conducted by Sadeghi and Khajepasha (2015) revealed that theses mainly suffer from style and language problems, rather than, content and methodological and far less from organisational problems.

What this study does not aim to, is to emphasise the ubiquitous however insignificant instances in the practice of some examiners, such as what is called 'nit-picking' or focusing on trivial points. However, this cannot further be trivialised as Pearce contends that the thesis student is often faced with series of minor writing problems while the major expectation is that the student should be prepared to defend the thesis as a whole. Anyhow, the significance of style and presentation cannot be underestimated in the eyes of examiners. If examiners feel irritated about writing style, they may use it 'as an excuse not to engage with the substance of the text' (Pearce, 2005: 3). The writing errors are more easily targeted even in the case of a second common scenario explained by Pearce (2005: 4), when a jealous colleague is appointed as internal examiner. In any case, the common writing problems should be given the attention they deserve to avoid unwanted consequences.

Research Questions

Based on the suggestions made in literature and the need perceived in the local context, the following research questions are proposed.

1. What are the most common writing errors examiners find in EFL Iranian students dissertations?
2. Is the perception of the quality of L2 writing dissertations the same among students and examiners?

REVIEW OF LITERATURE

Studies have investigated the difference between perceptions of native and non-native English speakers (NNS) in evaluating students' writing (James, 1977; Hughes and Lascaratou, 1982; Santos, 1988; Kobayashi, 1992; Hinkel, 1994) some have also explored the difference between expert writings of native English speakers and nonnative English speakers (Connor-Linton, 1995b; Kobayashi and Rinnert, 1996; Hamp-Lyons and Zhang, 2001). However, few studies have focused on the perceptions of thesis examiners as they assess and comment on students as they write and respond. The perceptual difference between these two groups who need to understand each other closely is of crucial importance because if their understanding varies, the results might be unfavourable for writing dissertations. Hinkel (2004) argues that inadequate knowledge of second language grammar and vocabulary as well as the complexity of the task of writing dissertation itself are the reasons for the lower quality of theses written by non-native speakers as compared with native speakers in general even if highly advanced NNS writers improve their writing quality by decreasing the number of errors. Therefore, it is safe to assume that students alone may not be able to face this challenge. NNS students usually resort to several strategies to fill this gap, from attempting to learn useful writing patterns (Okamura, 2006), as well as obtaining feedback from their peers or supervisors (Tsui and Ng, 2000; Liu and Sadler, 2003; Burrough-Boenisch, 2003), avoid difficult structures by using simple language (Hinkel 2002). Limited research has been conducted in each area, and there is a gap in researching the comments written by the examiners on submitted theses.

There are linguistic features of quality of writing that were researched and suggested in previous research, although some predictors are non-linguistic. McNamara, Louwerse, McCarthy, and Graesser (2010) distinguished high quality from low quality texts using the linguistic criteria

and computational instruments. The three most important linguistic features of quality of writing are syntactic complexity, lexical diversity, and word frequency. In a study conducted in Japan, again the length of writing was correlated highly to writing quality (Mellor 2010). Despite the importance of linguistic features, other non-linguistic features are essential in writing quality of dissertations; such that Yu explains ‘the quality of a discourse, written or spoken, is defined and shaped by various linguistic features other than diversity in vocabulary (e.g., quality of handwriting, structural features of writing; good pronunciation or being fluent in speaking)’ (Yu, 2007: 80). The examiners, however, do not have time or tools to measure linguistic features of writing quality. Even if they could, they would not opt for an elaborate a statistical process of examining every single dissertation they encounter. They judge them intuitively.

Current practices of L2 writing are increasingly said to lack challenging environment to engage students who want to master L2 writing (Mirhassani, Samar, and Fattahipoor, 2006). The postgraduate students dealing with dissertation writing are those who do not learn much out of their classes, either, as their academic writing experience is extremely limited and are not as complex as it should be to meet the requirements of university and courses. When they are encouraged to improve their writing, there is only an emphasis on successful completion of writing rather than providing adequate support that students need to improve writing. Therefore, students may not receive necessary support for the university level writing. In such situations, students are left to teach themselves. The examiners are (or should be by law) detached from writing practice of students and comment on the product while in fact the supervisor or writing instructor should actively teach writing as a process and scaffold students in a step by step fashion (Atkinson 2003:10) and we suggest they can use insights from common errors, perceptions and practices to better achieve their goals. The comments on product may be repetitive and unclassified. If the errors students make are categorised locally, they can be a basis for teaching writing deliberately which in turn would guide them to improve the quality of writing. In short, the awareness of L2 writing is not easy to acquire from teaching in classroom alone and perceptions of writing quality diverge.

RESEARCH METHODS

The research method are designed and triangulated to capture perceptions in multiple ways. In order to make the method robust and decrease the possibility of estimation error, a corpus-driven content analysis, viva voce observations, and questionnaire data are used in this study. An analysis of written products and examiners' notes and comments were gathered and supported with observations to confirm the examiners' comments. In other words, both qualitative and quantitative methodologies are used to explore and confirm the results. In following sections, the participants in the study, instruments of data collection and the procedures for data analysis are explained further.

Participants and Instruments

To answer the first research question which is to find the most common writing errors EFL Iranian students have in writing dissertations in the perception of examiners, a pool of EFL dissertations with examiners' comments were collected as the corpus in this study. The theses were defended in four universities, namely University of Tehran, Shahid Beheshti University and Islamic Azad University (two Branches). The examiners' comments were usually written on the margins of the texts intended for examiners. The comments were explored and used as a basis to find the most common errors in writing dissertations from the examiners' points of view. The theses were all defended 2013 onwards, which indicate a relatively recent account. In total, eight full MA dissertations were randomly selected from the assessed theses in the English Departments of the universities. The criteria devised by Sadeghi and Khajepasha (2015: 362) were used initially as an instrument (See Appendix 1): All of the theses were written in English as a second or foreign in the field of Applied Linguistics within the English Department. The reason for the sampling, apart from the accessibility, is that the choice of a homogenous sample in terms of field of study, department, year and examiners, makes it easier to compare and draw implications from.

Preparing the questionnaire items

As mentioned previously, the list of common error patterns was used as guiding themes to obtain further data in a few viva sessions where the respective students and examiners were present. The data combined from the corpus analysis and semi-structured observations are summarised in (Table 5). It provided the tool to answer the second research question, where the points of divergence in the perception of examiners and students began to take shape. The next sections will further explain how the questionnaire was developed, validated, administered, scored and analysed.

Since self-administered surveys reduced the chance of bias introduced by the interviewer, hence the next phase of the study was conducted through questionnaires. A collection of potential items created the item pool, which other than some verbal creativity, were mostly based on two sources. First, qualitative, exploratory data obtained from informants, e.g. notes taken after viva voce meetings (n=20) where the first researcher was the examiner, and second, brainstorming with colleagues who were examiners in the universities mentioned previously and most significantly from the comments examiners had written in the theses. Few questions were obtained from experienced colleagues. Then, a reductive process began where the researchers omitted repetitive questions and edited wordings to make them clearer. A pilot test was done with a few colleagues where items with ambiguous wording were revised and tested.

Instrument validation

Reporting appropriate internal consistency is an absolute requirement and is the first index to be reported before indicating evidence from two types of validity. Thus, the questionnaire was administered two times with a few cases thought to be representative in the pilot phase. Test-retest reliability showed an internal consistency Cronbach $\alpha = 0.74$, which falls at an acceptable level (between $0.8 > \alpha \geq 0.7$) according to George and Mallery (2003).

Content validity

For the purpose of providing content validity evidence, the piloted version of the questionnaire was given to three experienced examiners to see whether they judged the items as valid and true representation of the perceptions and practices of examiners and students. They were encouraged to make comments and suggestions on the clarity of the wording, difficulties during completion, layout and style of the tool. Table 4 is the final version of questionnaire. Their comments were mostly to rectify style. A limitation of this method can be the mode of marking used by examiners in that the examiners employ the more traditional mode of writing in margins or commenting in a word processor or both. Johnson et al. (2012) confirms that the mode of marking did not affect marking accuracy.

Construct validity

In this study, the respondents' psychological features, age, gender and other personal or demographic features were used due to two reasons. First, sufficient information was available of profiles of students and examiners. Second, other methods were more suitable for longer questionnaires which elicit data from larger number of participants. Inevitably, more statistical methods can be used in future studies when the questionnaire items and sampling get more diverse and beyond the preliminary level of trial.

Concurrent validity

Parallel version questionnaire was used to collect some evidence on concurrent validity. Another version of the same questionnaire (See Appendix 1) was developed using SurveyMonkey templates. The results from administering it to similar respondents; students (n=10), examiners (n=5), show a strong positive correlation with the results obtained in the prototype main version. The correlation coefficient was at .80. Furthermore, the reliability was measured by Cronbach's Alpha and the results showed an alpha coefficient of .83 indicating a good level in the overall index of reliability (Clark & Watson, 1995).

Research Procedures

In sum, the study was conducted in two phases. Phase I of the study, the exploratory phase, is when the most common writing errors EFL Iranian students were found. This procedure contained a content analysis of written academic theses in 2013 and 2014. The intervening variables were controlled (e.g. year of termination, field of study, number of examiners, etc.) to reduce the possibility of error due to variability in examiners views and departmental practices and norms. The analysis was based on the criteria illustrated in Table 1.

In Phase II of the study, which is the confirmatory phase of the study, a quantitative approach was used to describe the perceptions of examiners and students of the writing quality. The corrected version of the finalised theses submitted to university library were also sought to find out comments of examiners which were promptly dealt with by students and those which were not. The themes drawn from viva observation were combined with the themes of error analysis conducted in phase I.

Table 1: Universal Framework for Thesis Writing

Preliminary pages	- Title page - Committee page - Dedication page - Acknowledgment - Abstract - Table of contents - List of Tables, Figures, abbreviations, etc.
Chapter 1: Introduction	- General background information on the study - Purpose and significance of the research - Research questions and hypothesis - Definition of key terms
Chapter 2: Review of the literature	- General review of relevant literature and report on theoretical frameworks relevant to the problem - Report on previous research relevant to the problem (major findings, how they were obtained and what can be learned from them) - Critique of previous research studies (problems in design, instrumentations and analysis techniques as well as their contribution to the literature) - Statement of the gap in the literature and the rationale for the study
Chapter 3: Method	- Research design - Participants - Research instruments - Procedures
Chapter 4: Results	- Statistical techniques for data analysis - Findings of the research
Chapter 5: Discussion and conclusion	- A re-statement of the research problem - A re-statement of the results - Discussion and interpretation of the findings in relation to previous research on the topic - Limitations - Implications - Suggestions for further research
References / Bibliography	Comprehensive list of all sources
Appendices	Questionnaires, tests, etc.

RESULTS

Some of the most common errors are found to be due to students' L1 Farsi influence. Eight theses were analysed for examiner comments (See Appendix for a sample of actual comments by examiners). Overall, the common errors are found at two levels; micro-level include unfamiliar grammatical elements like the use article (a/an/the), incomplete structures, etc. as well as macro-level issues with consistency and clarity of writing expression. The following areas were found as recurring areas of problem in the total thesis pool: lack of concern for APA style, incoherent writing, awkward writing style, acknowledgment, inappropriate headings, incomplete structures,

inconsistencies in presenting examples, inconsistencies in explaining treatment procedure (what is done for control group?), lack of clarity in presenting figures, lack of clarity in explaining the differences of treatment methods for groups, lack of clarity in explaining the differences of post-test with pre-test procedures and their results, lack of clarity in presenting discussions (needs revision), lack of clarity in theoretical framework.

The above points serve as a basis for the content analysis of the sampled dissertations (n=8) as commented by examiners and corrected by students. The means for each error category are presented in the first column in Table 2:

Table 2: The Frequency of Comments Written by Examiners in the Dissertation Corpus

Recurring problem areas found in theses	Mean	T1	T2	T3	T4	T5	T6	T7	T8
Lack of concern for APA style	1	1	0	0	1	1	2	1	2
Incoherent writing	1	1	1	0	0	1	1	2	2
Awkward writing style	1	0	0	1	0	1	2	2	2
No Acknowledgment	2	3	1	0	1	0	4	4	3
Inappropriate headings	2	3	1	2	1	2	1	2	3
Incomplete structures	4	2	3	3	4	4	5	5	6
Inconsistencies in presenting examples	3	3	3	2	3	2	3	4	4
Inconsistencies in treatment procedure	2	1	2	1	1	2	3	3	3
Lack of clarity in presenting Figures	2	2	0	0	2	2	4	4	2
Lack of clarity in explaining the methods	1	1	0	1	0	0	2	2	2
Lack of clarity in procedures and results	2	1	2	1	2	2	3	3	2
Lack of clarity in presenting Discussions	1	1	0	0	1	2	1	1	2
Lack of clarity in theoretical framework	1	1	0	1	2	0	1	2	1

Table 2 depicts the frequency of comments written by examiners in the corpus. The first column is the mean of statistics from eight theses in each column. The data were sampled from available theses drafts. Thus, there will be no claim on generalisability. Incomplete structures are found to be the most frequently occurring problem since in average there are 4 instances in each thesis. Inconsistency in presenting examples, with 3 instances on average, is the next frequent error that affects writing quality.

After analysing the submitted version of the theses marked by the examiners, the final version submitted were accessed through the library and reviewed again to analyse which comments were taken into account or discarded. Since students do not comment on the examiners' comments, we could not have a detailed explanation of why they revised in the way they did. The final revised version served as an indicator of their performance which might not always be in line with their perceptions. Observations made in some of the viva sessions served to triangulate the data and draw out the common themes. Table 2 shows the number of errors corrected by the students in each category in the dissertations corpus.

Table 3 categorises the comments of examiners corrected by students in the corpus. All the inconsistencies and lacks of problems were listed in Table 2 are re-grouped into four main problem areas, namely Lack of clarity - in explaining or presenting; Thesis formatting; Grammatical errors; and Inconsistencies in presenting with examples and details for each subcategory. Two raters were given the marked and final copies of the thesis to judge how many comments were actually implemented. Comparing Table 2 and Table 3 shows that students could only partially correct their theses based on the comments of examiners.

Table 3: The Comments of Examiners Corrected by Students in the Corpus

Main problems	Examples/details	
Lack of clarity - in explaining/ presenting	the differences of/between/ among	Treatment methods for groups Post-test with pre-test procedures and their results
	theoretical framework	
	Discussions (needs revision)	
	Figures	
Thesis formatting	Writing style	Incoherent/awkward
	Special sections	e.g. Acknowledgment
	Categorization	e.g. Inappropriate headings
	Lack of concern for APA style	
Grammatical errors	Micro	Article use (the/a/an)
	Macro	Incomplete structures
Inconsistencies in presenting	Presentation	Margins
	Content	Examples

The information obtained from the last two tables prompted the development of a questionnaire applicable to a wider audience or data points. Notably in Table 4, at least one choice out of each multiple choice item in the questionnaire was related directly and explicitly to writing quality. The choice regarding writing quality could be spotted by the students if they perceived them as more important than other thesis quality features. This served to avoid the research artefact of responses being made necessarily focused on a writing quality at the expense of other stylistic and organisational themes that examiners usually focused. The data obtained from the students and examiners served as a triangulation of research data with corpus analysis. The same common themes were the basis of the questionnaire where the resulting mean and standard deviations confirmed the results obtained from the analysis. Table 3 shows the questions developed based on the feedback received from the student participants. The themes that compose item stems and responses were drawn from a content analysis of comments written by examiners.

Table 4: The guiding questions used and the themes drawn from examiners

What examiners most probably focus on/ask examinees in an English defense session in Iran?		a	b	c	d
1	In Chapter 1	Research Question	Limitations of the study	Key terms	
2	In Chapter 2	Proper citation	Coverage of topics	Coverage of local studies	Other
3	In Chapter 3	Randomizing participants	Low number of participants	Prejudiced writing	Other
4	In Chapter 4	Monotonous writing	Lack of skill in APA style	Redundancy in Figures & Tables	Other
5	In Chapter 5	Mistaking the Results with Conclusions	Wrong or inaccurate Implications	Little real argumentative writing or Discussion	Other
6	Which Chapter is generally more focused on in Defence Session of Thesis?	Chapter 1	Chapter 2	Chapter 3	Chapter 4
7	What are the most frequent grammatical errors in Iranian students' dissertation writing?	Passive Constructions	Incomplete Structures	Improper use of Tenses	Article use (the/a/an)

Over 70 students and 20 examiners completed the questionnaires. The results are depicted in Table 5, with a mean shown for students and examiners separately. The percentage clearly shows a mismatch between the perceptions of examiners and students in a meaningful way, yet the difference is not statistically significant. Paired sample t-test also concluded that there is no statistically significant difference between the two groups. The difference could be due to chance. The questionnaire results shown in Table 5 compare teacher-student perceptions in the most significant writing points in each chapter of the theses. The significance computed is two-tailed and the p-value is 0.005. The result of the t-test showed 0.9 which is above 0.05 and therefore difference is not statistically significant.

Table 5: The questionnaire results comparing examiner-student perceptions in the most significant points in each chapter of the theses

Student and examiners perception		Students' perception (n=70)				Examiners' perception (n=20)			
Item choices		a	b	c	d	a	b	c	d
1	In Chapter 1	27	16	19	8	10	4	4	2
2	In Chapter 2	34	17	13	6	6	3	10	1
3	In Chapter 3	31	30	9	1	4	10	3	3
4	In Chapter 4	33	29	7	1	5	10	4	1
5	In Chapter 5	29	17	10	4	11	4	5	0
6	Which Chapter is generally more focused on in Defense Session of Thesis?	11	15	18	26	17	1	1	1
7	What are the most frequent grammatical errors in Iranian students' dissertation writing?	24	20	7	19	12	3	3	2
Total Mean		27.00	20.57	11.86	9.29	9.29	5.00	4.29	1.43
Percentage		39%	29%	17%	13%	46%	25%	21%	7%

As per the second research question, the following data analysis was done based on the results obtained in Table 5. Pearson product-moment Correlation coefficient, shows that in at least three out of the four choices (a, c, and d) made by examiners and students, the correlation is negative. It indicates that the perceptions between the two groups diverge. In other words, the inter-correlations between the perception of examiners and students were done between the counterpart choices, e.g. mean scores of each choice of examiners and students (choice a with choice a; choice b and choice b and so on). The results show that the only positive significant correlation is between choice c of each group and the rest of the correlations are negative. This results in a negative correlation as total average of all choices, which indicates a wide divergence between the examiners and students. In the discussion section, we explore how redemptive action is necessary to bring their perceptions closer.

DISCUSSION

The Universal Framework for Thesis Writing (Paltridge and Starfield, 2007) was considered to be familiar for both students and examiners (Table 1). This assumption, although true, turns out to be insufficient for ensuring mutual understanding between thesis examiners and students. One should acknowledge, perhaps no universal framework for thesis writing can be considered as a baseline since perceptions and understandings of students need to be checked to see if they matched closely enough with those of the examiners. Another point which was revealed from the findings is about the conduct of the examiners themselves. They claim or believe they use Universal Framework for Thesis Writing as a baseline. However, in personal communication with one of the examiners, it was revealed that they rely more on their intuition and previous experience of examining than any fixed guidelines. This is also apparent from their comments given on thesis drafts. Also the comments are supported with data from the viva voce sessions. Examiners of MA theses in Iran (where foreign status of English is considered a hindrance than help) would not expect a critical appraisal of the degree and the depth that Holbrook et al. (2007) would expect in a PhD thesis in Australian context.

The similarities and differences were the basis of themes taken from the content analysis data and final reduction to four main areas of lack of clarity - in explaining or presenting, thesis formatting, grammatical errors, and inconsistencies in presenting. The sub-areas are only the closest match (Table 3) and could well be re-categorised in another context. Areas of challenges in writing are presented in Table 3 as a synthesis of views held by students and examiners.

For example, lack of clarity, in explaining or presenting ideas consists of comments on students' writing regarding their examples or treatment methods for groups, and post-test with pre-test procedures and their results. Examiners commented that there were also other areas where explanations were not sufficient or clear such as the theoretical framework of the study, discussion sections or chapters in general as well as the figures which were not appropriately captioned or explained in text which needed further revisions. The problems relating to thesis formatting such as writing styles which were frequently incoherent or awkward, beside that, the writing of

sections such as Acknowledgments were also not appropriately written. This shows how attention to such style and language matters can improve the whole image of the thesis in the eyes of examiners.

Another point to discuss is that examiners commented on how the students categorise their theses since they observed frequent cases of poor headings, repetitive or converging sections. This emphasises the importance of organisation and can be interpreted as students' insufficient skill to diagnose their content, which in turn relates to their lack of experience. A suggestion is to train students to practise of thesis writing through tasks that break thesis into smaller pieces and consider this as transferrable skill.

Several issues relating to the use of APA and grammar can be related to the students themselves since these are normally taught in Iranian MA taught modules/courses. The lack of concern for APA style was evident when APA was explicitly mentioned as the norm for referencing. Perhaps, students need more hands-on practice with APA in the relevant course modules such as Essay Writing and Research Seminar. Micro-level as well as macro-level grammatical errors were also the issue that need attention. Interestingly, article use and incomplete structures were the most frequent grammatical errors at the micro and macro level respectively since both elements are not found in L1 Persian of the writers. Thus, it is very challenging for the thesis writers to implement those items correctly.

Inconsistency in presentation of the thesis was the last frequent issue found in content analysis. The formatting of some theses were inconsistent despite the stipulation of the rules. This type of problems can only be due to hasty preparations. The second aspect of inconsistencies in presenting the content in which examples or exemplifications were found inconsistent.

CONCLUSION AND IMPLICATIONS

The contribution of this research lies in not only explaining the benefits of the knowing the perceptual difference between students and examiners but also in throwing some light on the concept of quality of L2 writing, i.e. writing of thesis. Many courses and resources provide help for them in this process, but none has categorised the errors in terms of frequency and

significance from the actual comments given by the examiners or attempted to bridge the perceptual difference between writers and their examiners. The current study addresses this need and provides implications for the practice of writing and assessing EFL theses. Ultimately, thesis examiners are the ones who set the standard of what is acceptable as a thesis (or dissertation) and the award, thus understanding their perceptions of quality of writing a thesis is paramount. Hence, making difference in perceptions sufficiently clear is a contribution since this enables the students to focus more on particular points of conflict and also to find the most common errors from examiners' point of view lucid so that awareness of such issues may lead to increase in quality of writing MA theses in the described context. Since high standards of writing are mandatory for evaluation of theses, thus, responsible action should be pursued to address this issue that can ease the challenge in the practice of writing and assessing EFL theses at a large scale. Similar to the insights gained from a related study of Fook and Sidhu (2009), the findings of such awareness-raising studies empowers the faculty to be more receptive to the needs of students which in turn also enable them to address the critical problems more effectively. It is suggested that general essay writing courses as in-service or pre-service education are not sufficient and specifically designed courses on how to improve L2 writing is useful for L2 writers. These courses should work on assimilating the perceptions of examiners and students (who function as L2 writers and examiners) or at least bridge the gap and create a common ground for understanding the criteria for writing and assessing. It would also be extremely useful if these courses also focus on the most common errors in L2 writing. These the errors should be treated in action while writing up theses. Teachers should also be vigilant to find out the most common patterns of error emerging in writing dissertations and be prepared to take responsible action or intervention in correcting the students' errors and improving their practice. It may involve, at times, less direct error-correction and more awareness-raising about the differences between L2 and L1 writing and urging the students to produce multiple writings to give them the opportunity to excel themselves.

The results of this study can be compared to studies conducted in this area but with different research instruments and on various majors that span beyond English major. The research can be replicated to see if the same results are produced. Examiners may have reservations and may not want to give away the so-called 'tricks of the trade', so a more

viable recommendation is to hold workshops in which the process of writing and common writing errors are discussed generally from different perspectives, e.g. the potential examiners and writers take turn to discuss the most significant challenges in writing a dissertation. Such practices require specific skills.

REFERENCES

- Alderson, J.C., Clapham, C. & Wall, D. (1995). *Language test construction and evaluation*. Ernst Klett Sprachen.
- Atkinson, D. (2003). L2 writing in the post-process era: Introduction. *Journal of Second Language Writing*, 12(1), 3 – 15.
- Brown, J.D. (2000). What is construct validity? *Shiken: JALT Testing & Evaluation SIG Newsletter*, 4(2): 8 - 12. [ISSN 1881-5537]
- Burrough-Boenisch, J. (2003). Shapers of published NNS research articles. *Journal of Second Language Writing*, 12, 223-243.
- Clark, L. A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Connor-Linton, J. (1995) Crosscultural comparison of writing standards: American ESL and Japanese EFL. *World Englishes*, 14, 99–115.
- Fook, C.Y. & Sidhu, G.K. (2009). Integrating information and Communication Technology (ICT) into university teaching and learning: A case study. *Asian Journal University of Education*, 5, 71-94.
- George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th Ed.). Boston: Allyn & Bacon.
- Golding, C., Sharmini, S. & Lazarovitch, A. (2014). What examiners do: What students should know. *Assessment and Evaluation in Higher Education*, 39(5), 563–576.
- Hamp-Lyons, L. & Zhang, B.W. (2001). World Englishes: Issues in and from academic writing assessment. In Flowerdew, L. & Peacock, M., editors, *Research perspectives on English for academic purposes*. Cambridge: Cambridge University Press, 101–16.

- Hinkel, E. (1994). Native and nonnative speakers' pragmatic interpretations of English texts. *TESOL Quarterly*, 28, 353–76.
- Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hinkel, E. (2002). *Second Language Writers' Texts: Linguistics and Rhetorical Features*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Holbrooke, A., Bourke, S., Fairbairn, H. & Lovat, T., (2007). Examiner comment on the literature review in Ph.D. theses. *Studies in Higher Education*, 32(3), 337-356.
- Holbrooke, A., Bourke, S., Lovat, T. & Dally, K. (2004). Investigating PhD thesis examination reports. *International Journal of Educational Research*, 41, 98–120.
- Hughes, A. and Lascaratou, C. (1982) Competing criteria for error gravity. *ELT Journal*, 26, 175–82.
- James, C. (1977). Judgments of error gravities. *ELT Journal*, 31, 116–24.
- Johnson, M., Hopkin, R., Shiell, H. & Bell, J.F. (2012). Extended essay marking on screen: is examiner marking accuracy influenced by marking mode? *Educational Research and Evaluation*, 18(2), 107–124.
- Kobayashi, H. & Rinnert, C. (1996). Factors affecting composition evaluation in an EFL context: cultural rhetorical pattern and readers' background. *Language Learning*, 46, 397–437.
- Kobayashi, T. (1992). Native and nonnative reactions to ESL compositions. *TESOL Quarterly*, 26, 81–112.
- Liu, J. & Sadler R. (2003). The effect and affect of peer review in electronic versus traditional modes on l2 writing. *Journal of English for Academic Purposes*, 2(3), 193-227.
- Okamura, A. (2006). Two types of strategies used by Japanese scientists when writing research articles in English. *System*, 34(1), 68-79.
- McNamara, D.S., Louwerse, M.M., McCarthy, P.M. & Graesser, A.C. (2010). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes*, 47, 292–330.

- Mellor. (2010). Essay Length, Lexical Diversity and Automatic Essay Scoring. *Memoirs of the Osaka Institute of Technology*, 55(2), 1-44.
- Mirhassani, S.A., Samar, R.G. & Fattahipoor, M. (2006). The role of writing pedagogy in vocabulary improvement. *Indian Journal of Applied Linguistics*, 32, 91-106.
- Mullins, G. & Kiley, M. (2002). It's a PhD, not a Nobel Prize: How experienced examiners assess research theses. *Studies in higher Education*, 27(4), 369–386.
- Pearce, L. (2005). *How to Examine a Thesis*. Open University Press: McGraw-Hill Education.
- Paltridge, S. & Starfield, B. (2007). *Thesis and dissertation writing in a second language: A handbook for supervisors*. London: Routledge.
- Sadeghi, Karim & Khajepasha, Arash Shirzad. (2015). Thesis writing challenges for non-native MA students. *Research in Post-Compulsory Education*, 20(3), 357-373, DOI: 10.1080/13596748.2015.1063808
- Santos, T. (1988). Professors' reactions to the academic writing of nonnative speaking students. *TESOL Quarterly*, 22, 69–90.
- Stracke, E. & Kumar, V. (2010). Feedback and self-regulated learning: Insights from supervisors' and PhD examiners' reports. *Reflective Practice*, 11(1), 19–32.
- Tsui, A. & Ng, M. (2000). Do secondary L2 Writers Benefit from Peer Comments? *Journal of Second Language Writing*, 9(2), 147 – 170.
- Yu, G. (2007). Lexical diversity in MELAB writing and speaking task performances. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 5, 79-116.

APPENDIX 1

A detailed initial framework (adopted from Sadeghi and Khajepasha, 2015) and description of problem types and their examples in a sample of theses

Thesis #1

Problem type	Examples
Language	Padding is frequent; e.g. differential effects of different instructional treatments...
Style	Writing style is not academic particularly in 'Acknowledgements' section.
Content	The link between contents of initial sections in chapter 1 is weak.
Methodological	Treatment procedure is not clear.
Organisational	Guidelines reserve Chapter 5 for 'Discussions' but they are mixed with 'Results' in Chapter 4.

Thesis #2

Problem type	Examples
Language	Rosy language (gilding the lilly) is observed; e.g. excellent teachers, very experienced
Style	Figures and Tables are not in APA style.
Content	In the section on 'Significance of the study', the gap in literature is mentioned without mentioning the importance and novelty of research.
Methodological	Control group does not receive the unmarked instruction (should not be deprived).
Organisational	Guidelines reserve Chapter 5 for 'Discussions' but they are mixed with 'Results' in Chapter 4.

Thesis #3

Problem type	Examples
Language	The use of the Articles is mixed (the/a/an)
Style	Writing style is prejudiced
Content	Local studies are not covered
Methodological	Sufficient attempt is not made to randomise participants
Organisational	Conclusions are written mistakenly in 'Results' section.