True or False Questions in Translation Oriented Research Methodology Exams: The Reality between Myth and Truth

Zakia A. Deeb
Tripoli University
Tripoli – Libya
zakiadeeb@hotmail.com
z.deeb@uot.edu.ly

45

ABSTRACT

This paper investigates the way MA students handle True or False type of questions in a translation oriented research methodology final examination with regard to managing tricky segments, checking cues, retaining information and scoring marks. In the present study, the common belief among students that True/False questions, compared to other types of questions, are easy to handle and meant to secure marks is challenged. The evidence is drawn from an actual final examination using a True/False question involving ten statements (five True and five False) together with four other types, including an essay type question. While, in general, students managed to handle the question relatively well, in details they made unjustified and incorrect choices. This only proves the misconception among students that True or False type of questions offer an opportunity to score high marks that might not be possible with other types of questions. Quantitative data and, to a less extent, qualitative data collected in the study and analyzed show that this type of questions is neither straightforward to handle nor it guarantees scoring higher marks. On the contrary, results show that the True or False question scored the second least marks among the other questions. Moreover, none of the 23 students engaged in the study scored a full mark in this question while a number of students did manage to score it on a number of occasions in the other types of questions.

Keywords: Translation, Research methodology, True or False, Myth, Misconception

1. Introduction

There is a consensus among scholars that quality research requires both knowledge in the discipline and competence in research skills. Unlike before, high standards set recently for higher education research for qualifications led to better understanding of skills training requirements for research students. For example, frameworks such as the European Higher Education Area (Bologna process http://www.wg.aegee.org/ewg/bologna.htm) and the Dublin Descriptors (Clark & Mayer, 2008) leave no room for compromise with regard to quality in research. This led to the increasing demand for ensuring quality in research in higher education in general (Reisz 2008) and in translation studies in particular. For example, the League of European Research Universities (LERU) offers regular advice on good practice elements in doctoral training. LERU (2010) proposes principles of excellence in doctoral training and recommends that universities should rigorously follow those principles. In the context of enhancing translation

and interpreting research quality, efforts such as those initiated by Advancing Research in Translation & Interpreting Studies (ARTIS, n.d.), Translation Research Summer School (TRSS) and Centre for Translation Studies CETRA, let alone the vast number of conferences dedicated to exploring some aspects of Translation Studies as a research discipline and societies that foster research in translation and interpreting, are apparent. Moreover, a number of scholars addressed particular issues in translation studies that might be challenging in research, e.g. Gambier (2009, unpaginated) "challenges in research on audiovisual translation, García, (2009, unpaginated) "research on translation tools".

However, the fact that "(…) poor research is still frequent in Translation Studies", Gile, 2009: 36) is alarming. According to Gile this is "(…) attributable to lack of training, not to a lack of intellectual capability", (ibid.). Among the frequently found flaws, Gile lists the following: "non-representative samples, invalid research design, overgeneralization of findings, misrepresentation of views expressed in the literature, logical problems in inferencing or incorrect grasp of concepts imported from cognate disciplines", (Ibid).

In translation pedagogy and, maybe, in other disciplines too instructors recounting on their classroom experiences is not unusual. To mention only a couple from the most recent studies, the following are only two examples: (Pym, 2016, Berber and Kretschmann, 2016), and the present study also falls in this category. For the purpose of elaboration, it would probably be necessary to explain the expression "Translation Oriented Research Methodology" (TORM), which might not be evidently familiar, and contextualize the relevant course upon which the conclusions are drawn here. TORM is simply intended to help translation students develop their translation oriented research competence. This, however, differs to the "Translation Oriented Research Competence" referred to by (Öner: 2013). Öner in her study is using the expression to mean "(…) the ability to apply the required research knowledge for translation purposes", (Öner: 2013: 3160). TORM in the present study refers to a tailored course to train students researchers enrolled in a translation/ interpreting MA course to conduct research and write their dissertations. Students enrolled in the programme are granted the MA degree after successfully completing 12 translation related subjects and submitting a dissertation. Its content is largely based on the translation oriented Williams and Chesterman's (2002) book: The Map and extra supplementary materials. Students enrolled in this course are supposed to be familiar with the basics of general research from their previous underground studies and experiences. Types of assessment in this course include both assignments and exams. In this study the focus is only on the final examination results in one type of exam questions, namely the True or False (T or F) type.

In this paper, a study investigating the way students handle T or F type of questions with regard to managing tricky segments, checking cues, retaining information and scoring marks is presented. The study aims to assess the value of this type of test in translation education and orientation in conducting research in translation to bring to light the benefits and limitations of such a test in translation education. The study is conducted at the end of a sixteen weeks course at the Department of Translation/ School of Languages/ The Libyan Academy for graduate studies, Spring 2017. The population of the study is twenty three students sitting for a final examination in the subject of TORM. Data is collected from students' answers to the T or F question (10 statements; five are true and five are false). Data is mainly analyzed quantitatively, but also qualitatively when relevant. With regard to performance in the T or F question in comparison with the other types of questions in the whole exam, the highest/ lowest scores in the five different questions in the exam sheet (True or False, Fill in the Blank, Explain, Discuss and a Short Essay) are compared. Because of the nature of the data

46

collection process being part of a formal exam, the study is only a product based one that employs no triangulation of methods.

2.Literature review

With advances in all aspects of the educational disciplines, emphasis on the quality of designing questions for testing students' achievements is becoming increasingly pronounced. Based on early thinking in learning and achieving that goes back as far as the Socratic questioning methods more than 2000 years ago, taxonomies have been created and others have been revisited and improved to promote higher forms of thinking in education. For example, Bloom's 1956 taxonomy of learning domains (creating, evaluating, analyzing, applying, understanding and remembering), has been revised, changed and rearranged, Anderson, et. al. (2001). Most studies in learning and testing, however, focus on the cognitive and psychological aspect with varying numbers of categorizations. Bloom's original cognitive taxonomy (cited in Huitt. 2009) matching the six learning domains above refers to three levels of knowledge: factual, conceptual and procedural. After being revised by Anderson and team (Ibid.) metacognition was added. Some refer to the three categories under different labels: knowledge, Skill and Attitude KSA others refer to five under the model of Structure of Observed Learning Outcome (SOLO) taxonomy: pre-structural, uni-structural, multi-structural, relational, extended abstract (Biggs and Collis, 1982). (Clark, Chopeta, 2004; Clark, Mayer, 2008) have also identified five categories: facts, concepts, process, procedures and principles.

Any learning process whether in school, university or beyond would end up with a form of evaluation of the outcome to assess the degree of achievement. Through history and modern times, various methods have been invented and developed to test students achievements during and after any educational process that is subject to instruction. In higher education, systems such as Voluntary System of Accountability VSA (http://www.voluntarysystem.org/) are set to measure learning outcomes even of the whole period students spent in college either longitudinal (the same group) or cross-sectional (different groups of freshmen and senior), liu (2009). In higher education, formal teaching in the area of research methodology in translation studies "opinions differ as to the value of such formal training", (Schaeffner: 2009: 11). However, in a system such as that (Pym, 2016: 1) refers to: "(…) master-apprentice models place authority in the instructor's experience rather than in the learners' discoveries", formal teaching is indispensible. With regard to training in research, it is not easy to attain the maximum goals set for efficiency in training to ensure quality in research. What might be ideal, of course, is if a 'research environment' and an 'intellectual climate' are guaranteed. In small institutions when students are expected to be taught more than develop during the course of conducting research, however, both requirements are less than ideal to demonstrate research skills and techniques students acquired. According to the Postgraduate Research Experience Survey (PRES, 2007) both requirements often fail to meet students expectations. Results show that non-EU students are the least positive about 'intellectual climate', (PRES, 2007: 3).

Unlike assessing competence and performance in pure science, in the area of humanities studies assessment is hard to maintain a 100% accuracy. One method, however, stands out: it is the T or F type of questions. Having

said that, T or F questions type can be considered as falling under the cognitive domain (Bloom, 1956), even though lower level cognitive skill in contrast to higher level cognitive skill (Lamb, 1976), that involves knowledge and the development of intellectual skills to recall and recognize specific facts (acquisition of knowledge comprehension of material). Interestingly enough, despite the fact that the field of evaluation has broadened to include so many types and tools of assessment, there is hardly anything about using the T or F test in translation studies and translation education. With regard to translation studies, the term is hardly mentioned, and when it is used, it is used in the context of equivalence. Baker (2009) identifies four meanings: prepositional, expressive, presupposed and evoked. According to her, "expressive cannot be judged as true or false". In translation education, T or F type of questions seem to be used in applied translation, i.e. classroom activities. Melis and Albir (2001: 285) suggest the use of multiple choices type of questions among four competence assessment instruments that can complement or even replace translation tasks. Since T or F is a kind of multiple choice type, by implication, it would be included.

2.1. True or False Type of questions

T or F questions are a kind of Multiple Choice question type: Yes or No, Agree or Disagree, etc. With T or F questions, students choose T or F in response to a statement question. They "are well suited for the assessment of both the learner's retention of specific information from a course and their general understanding of the material", (Treser, 2015).

The purpose and characteristics of True or False questions can be summed in the following three points put forward by the University of Waterloo, Centre for Teaching Excellence, they:

- Are most often used to assess familiarity with course content and to check for popular misconceptions.
- Allow students to respond quickly so exams can use a large number of them to test knowledge of a broad range of content.
- Are easy and quick to grade but time consuming to create.

(Univerisity of Waterloo, n.d.)

Like any type of measurement, T or F type has its advocates who see virtues in using it and those who hold it in low esteem. In an experiment to investigate the efficiency of T or F test items in comparison with multiple choice type of questions, Ebel (1971) found that true or False test can be as reliable as multiple choice test and, in fact, discriminating. This, according to Ebel, strengthens faith in the usefulness and value of T or F test. In defense of such type of questions against criticism of T or F questions of being often trivial or ambiguous and always susceptible to guessing, Ebele (1970) quoted in Ebel (1975: 31) argues that "these faults are not inherent in the form, and need not seriously limit its usefulness". T or F questions, whether solo or among other types of questions, are generally favoured by students in exams for being light, quick to process and offer a way to escape assessing their language proficiency with regard to grammar and spelling. "When asked, students often say that they prefer true-false tests" (Pennington publishing, n.d.). True or False questions are unlike Essay Questions or even Fill in

the blank questions; there is nothing to write but use the letters (T or F) or the symbols (√ or ×) when appropriate. Such a flexibility is expressed by Treser, (2015) below:

> **A learner can answer a large number of true/false questions within a short period of time. This enables you to cover a wide area of knowledge and test for understanding of every significant point even if you are on a tight schedule. Note that despite the name, you are not forced to use the "true/false" pair of answers specifically - "yes/no" will serve just as well. Use whatever option works best for you.**

(Treser, 2014, unpaginated)

49

Instead of the letters T or F, another option is also available: C or I for Correct and Incorrect answers respectively. The disadvantage with T and F questions, however, is that the matter is 'either/ or'; no room for taking into consideration any interpretations of the answers in favour of the student rather than whether the answer is "overtly" right. Some exams, IELTS, for example, add the "not given" option besides T or F choices.

In general, there are three types of T or F questions: Simple, complex and compound. With simple T or F questions, the choice is between True or False only. With the complex type, the choice is between True or False plus giving opinion. With the compound answer, if the answer is False, an explanation of why the answer is False has to be given.

2.2. Strengths and limitations of T or F questions

On the scale of advantages and disadvantages, T or F tests have strength and limitation points. Among the strengths and limitations as put forward by the University of Texas (n.d.) are the following points:

Strengths:

- Can cover a lot of content in a short time (about two questions per minute of testing time)
- The question is useful when there are only two possible alternatives.
- Less demand is placed on reading ability than in multiple-choice questions.
- Can measure complex outcomes when used with interpretive exercises.
- Scoring is easy and reliable.

Limitations:

- Difficult to write questions beyond the knowledge level that are free from ambiguity.
- False statements provide no evidence that the student knows the correct answer.
- Scores are more influenced by guessing than with any other question type.
- Cannot discriminate between students of varying ability as well as other questions.
- Requires that the answer to the question is absolutely true or false.

Luck sometimes works even when the choice of T or F is made on a mere hit-or-miss basis but questions can sometimes be too tricky to guarantee a *hit* answer. Tips and advice on how to handle tests and exams are often given in different ways to students to succeed particularly by university advice and counseling centers. With regard to T or F questions, varieties of tips are given, not only on how to answer questions but also on how to study (e.g. http://www.how-to-study.com/study-skills-articles/true-false-tests.asp). Tips of how to prepare for taking the test in different forms including videos are also available (cf. Monkey See, n.d.).

Test takers are also alerted to which statements tend to be True and which ones tend to be False and advice on working strategies to score marks in this type of tests is also given (cf. http://blog.penningtonpublishing.com/study_skills/the-top-nine-tips-to-taking-true-false-tests/. For example, one of the tips given to T or F test takers suggest that there is a greater chance that longer statements tend to be false "because it only takes one part of the statement to be false to make the whole statement false", (Ibid.). A statement like "there typically tend to be more true-false questions that are true than false" (College Atlas, n.d.) is often repeated. Sometimes a reason is given, "because teachers want you to know what is true, most true-false questions are true" (Life Solutions, n.d.).

Similarly, tips and advice are given to examiners on how to compose T or F questions. For example, University of Waterloo, Centre for Teaching Excellence gives the following advices in what to avoid and what to use when composing T or F questions under the title: Tips for Writing Good True/False Items:

Avoid	**Do use**
• Negative and double-negatives	• Your own words
• Long / complex sentences	• The same number of true or false statements (50 / 50) or slightly more false statements than true (60 / 40)
• Trivial material	
• Broad generalizations	
• Ambiguous or indefinite terms	• One central idea in each item

(Univerisity of Waterloo, n.d.)

In more details, Colorado State University, gives the following advice when composing True/False questions,

1. Avoid using more than one idea in a True or False question. Make your main point prominent.
2. Keep the statement short and simple. The question should be based on the learner's knowledge and not their ability to interpret the question.
3. True statements should be true under all circumstances. Avoid using may, seldom, possible, often, and other qualifiers
4. Use negative statements sparingly and do not use double negatives. Negative words are often overlooked and should be underlined or in capital letters.
5. Opinion statements should be attributed to some source. Instead of agreeing with the stated opinion, the students should be aware of the opinions of the organization or individuals.
6. When cause and effect relationships are being measured; use only true propositions.
7. Avoid extraneous clues to the answer. For example, always, never, none, all, only, etc.

(Colorado State University, n.d.)

In comparison between T or F type of questions and essay type of questions, the former is described to be "objective" and the latter is described as "subjective":

> **There are two general categories of test items: (1) objective items which require students to select the correct response from several alternatives or to supply a word or short phrase to answer a question or complete a statement; and (2) subjective or essay items which permit the student to organize and present an original answer. Objective items include multiple-choice, true-false, matching and completion, while subjective items include short-answer essay, extended-response essay, problem solving and performance test items. For some instructional purposes one or the other item types may prove more efficient and appropriate.**

(CITL, n.d., unpaginated)

Despite the charges of superficiality, ambiguity and guesswork against objective type of tests, including T or F type, they are, no doubt, fair and impartial. They are described as objective for a good reason. According to Ebel (1977: 11) the scorer of an objective test indicates unequivocally which answer he considered correct to each question. Each student's performance is judged against the same standard, thus an objective test score is likely to be fairer to the students than essay test score.

The present study is specifically intended to find out to what extent students manage to handle T or F questions in exams. It examines when do students miss the point in answering this type of questions. Tricky parts of the T or F questions in the present study could be challenging in a way that they genuinely test how well students understood the material in the subject under consideration but by no means intended to be misleading. In other words, the prime aim here is not to test students ability to decipher clues in answering confusing questions as such but to test their competence to truly distinguish between what is a TRUE answer and what is a FALSE one. This kind of test falls under what is termed as *summative assessment* (midpoint and/or end of instruction) in contrast with *formative assessment*, i.e. assessment during the instruction period (Goff-Kfouri, 2004). The findings are hoped to alert teachers to the misconception among students that T or F questions are meant to aid them get marks *free of charge* as a compensation of otherwise more *difficult* questions with which they, comparatively, often score lower marks.

3. Methodology

A test of 10 T or F questions is given to 23 MA students at the Translation Department, School of Languages, the Libyan Academy to answer among other questions (four more) in a three-hours final examination at the end of a 16 weeks course in the subject of Translation Oriented Research Methodology. The T or F questions test is placed at the top of the examination sheets of four pages but, obviously, with no obligation to answer it first.

Data is analyzed quantitatively on the basis of the number of correct and incorrect answers. Qualitative analysis is also carried out when necessary; when students voluntarily add information they are not asked for, for example.

Measurement is carried out according to averages of marks per individual student (1student per 10 questions) and per group (1 student per 23 students).

This study is a product based one. Data is collected from what is shown on the answering sheets. Because of the nature of the test being part of a formal final examination, no post-task interview or questionnaire were conducted. Understandably, however, this would halt any chance to find out whether students resorted to the guess strategy based on the 50-50 chance of the answer being right.

4. Hypotheses

The present study hypothesizes that:

1. The common belief that T or F questions are easy to handle by students is not necessarily true.
2. Familiarity with True or False questions does not guarantee that students fully manage to use cues and clues and master the skills to answer such questions.
3. Contrary to the common belief, students not necessarily hit the highest marks with T or F questions when they are used together with other types of questions.

5. Population

The population of this study consists of 23 MA students (17 female and 6 male) from the Translation Department/ School of Languages/ The Libyan Academy. They are all freshmen (first semester), their ages are between 23 and 30. Neither gender nor age is a variable in the present study. Students are familiar with this type of questions in different exams during their undergraduate studies. With regard to the present T or F type of questions, they vociferously asked for it to be included in the final examination after they wondered why it had not been included in the mid-term exam claiming that their marks would have been better if this type was included. In fact, it is a usual post-test complaint when T or F question is not included.

6. Material

The T or F test is part of a final examination consisting of five questions for a total score of 50 marks equally divided (10 marks each). The T or F question consists of ten statements that only require students to indicate whether the statement is T or F in the space provided (simple type; see above); no 'not given' option is given neither a slot for correction or opinion is provided. The 'to explain' option is saved for a separate question on the exam sheets. Students' answers should be either True or False depending on the content of the statement. The selection of the statements in the test is based on facts presented in the TORM course. The textbook used in the course is The Map: A Beginner's Guide to Doing Research in Translation Studies by Williams, J and Chesterman, A. (2002) besides other supplementary material discussed in class in a form of lectures, class discussions and presentations. Statements are designed, however, to be somehow challenging. The choice of the answer totally depends on the student's understanding of the statement. Questions are randomized with regard to the order and equally distributed: five are True and five are False.

One limitation in the test that has to be noted is the fact that it is not designed to test uncertainty of students choices. That is, if any student used the "hit-or-miss" method to merely score marks, there will be no way to find out. This can be an interfering factor though taking the level of the students into consideration, they are unlikely to resort to such a method. To get the most of the exam, particularly with

52

this part of questions, they understand to be extra cautious not to take risk. Moreover, the test is not a 'rapid-fire questions' type that any instant reaction answer could be considered. On the contrary, students need to think twice before making up their mind about the answers. With regard to speed, it is up to the student to divide the 'airtime' available from the three- hours-time allocated to the exam as a whole proportionally between questions, or otherwise.

T or F question in the exam, as in the case of the rest of the questions on the exam sheets, is primarily aimed to test students command of the subject. In terms of technicality, questions are designed to meet the maximum possible criteria of True or False questions.

53

The ten T or F statements appear on the exam sheet as below:

I.   State whether the following statements are False (F) or True (T)   10 marks

1- Obtaining a consent of all those to be involved in your research is essential, but not necessarily written. (…)
2-The advantage of using the think aloud technique is that data recorded would not need to be transcribed. (…)
3-Target language oriented translation quality assessment approach is not concerned with equivalence. (…)
4-Unlike note taking, using video or audio as a method of recording talk during interviews is distracting. (….)
5- In observation findings are recorded simultaneously not afterwards. (…)
6-There is no way to set up international standards to control or assure quality in translation. (….)
7-With regard to ethics, clash between loyalty to author and loyalty to readers is quite possible. (….)
8- Because of the nature of interpreting, only cognitive studies are possible. (…)
9-The phases of research planning take place more or less simultaneously. (…)
10-   In research, the two terms *References* and *Bibliography* are different. (…)

With regard to length of the questions, it was not feasible to maintain consistency. In word count, the longest statement is 19 words (Q2) and the shortest is 8 words (Q5). However, five statements are ranging between 8 to 13 words, two of which are exactly 11 words (Qs 8 and 10) while the other five are between 17 and 19 words, two of which are exactly 18 words (Qs 1 and 4). Correlation between length and tendency of being T or F is not one of the criteria upon which the questions are designed. To illustrate the none-correlation factor in the above test, the two second long sentences (both 18 words), one is T and the other is F and the second shortest (both 11 words), also, one is T and one is F.

While the T or F question is intended to test students' knowledge of specific factual information, the other questions on the exam sheets, apart from the Fill in the Blank one, are the sort of "open-ended questions focusing on broad general ideas. Questions on the exam sheet appear as below:

   I.   State whether the following statements are True (T) or False (F)
  II.   Fill in the Blank
 III.   Explain
 IV.   Discuss
  V.   The following extract is from the conclusion of a research paper in translation. Parts of it are altered and replaced by items or statements which are considered, from a research point of view, inappropriate. Underline those parts and, below, explain what is incorrect with them.

7. Data Analysis

Data collected from the 23 exam sheets is simply logged in a table where students, labeled in numbers, are vertically listed in an alphabetical order, each number is preceded by the letter S in reference to 'student' while questions are listed horizontally from 1 to 10 preceded by the letter Q in reference to the word 'question'. The symbols √ or × are used with reference to the correct and incorrect answers respectively in the corresponding cells. The total of correct answers per student is given towards the end of the relevant grid while the total of the incorrect answers per question is given at the end of each question's column as shown in Table 1 below:

54

Table 1: Numbers of correct answers (√) per student and incorrect answers (×) by question.

Students	Questions										Total
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**√s**
S1	√	√	√	√	√	×	√	√	√	√	9
S2	√	√	×	√	×	×	√	√	√	√	7
S3	×	×	√	√	×	√	√	√	√	√	7
S4	√	×	×	√	×	√	√	√	√	√	7
S5	√	√	×	√	×	×	√	√	√	√	7
S6	×	×	√	√	×	×	×	√	√	√	5
S7	√	×	√	√	√	√	√	√	×	√	8
S8	×	×	×	√	×	√	√	√	×	×	4
S9	√	√	√	×	√	×	×	√	√	√	7
S10	×	√	×	√	×	×	√	√	×	√	5
S11	√	√	×	√	√	×	√	√	√	√	8
S12	×	√	×	√	×	√	√	√	×	√	6
S13	×	√	×	√	×	√	√	×	√	√	7
S14	×	√	√	√	×	×	√	√	√	√	7
S15	×	×	×	×	√	×	√	√	√	√	5
S16	√	×	√	√	√	×	×	×	√	√	6
S17	×	×	√	√		×	√	√	√	√	7
S18	√	√	×	√	√	×	√	×	×	×	5
S19	×	√	√	√	√	√	×	×	√	×	6
S20	×	√	√	√	×	×	√	√	√	√	7
S21	√	√	√	×	√	√	×	√	√	√	8
S22	×	√	√	×	×	×	√	√	√	√	6
S23	×	√	√	√	×	√	√	√	√	√	8
Total ×s	13	8	10	4	13	14	5	4	5	3	

A close look at Table 1 above shows that Q6 got the highest number of ×s (incorrect answers) because of what seems to be an influence of the usually repeated statement that agreement on translation quality assessment standards is hard to attain. The statement "there is no way to set up international standards to control or assure quality in translation" is too absolute using the form 'no plus noun' to be considered T, but, nevertheless, does not sound unusual to students to be marked as F. However, the fact that some general principles in the area do exist, apparently, led student 14 (who got the overall highest mark in the exam but still

got this one incorrect) to voluntarily note in the margin that "however, certain approaches can be followed … ". Moreover, it is the only question whose first answer F to be crossed and incorrectly mended outside the brackets as T by the same student (S14). The fact that Q6 is the one that prevented S1 (see table above) from getting a full mark (9 out of 10) is also indicative of how far this question is subtle.

Next highest in getting ×s are Qs 1 and 5 (total each 13×s). The delicate part in Q1 is in the statement "not necessarily written". In research ethics, voluntary informed consent is a prerequisite for a subject's participation but in some cases it can be verbal. Hence, the statement is T. With regard to Q5, it is false because the option to record findings during or afterwards in observation is available. The 13 students who got it wrong by marking it as T seem to have taken the statement "findings are recorded afterwards" as if it is based on common sense.

What seems to be puzzling about Q3 that made 10 students get it incorrect is the strong claim involving "equivalence", around which translation traditionally revolves. To claim it is not an issue here, even if the focus is TT oriented, is hard to believe unless the student is so alert to details of the course material and class discussion.

Next highest with ×s is Q2, 8 students got it incorrect. What seems to make Q2 sound true to those who got it incorrect, while it is in fact false, is the strong start of the statement: "the advantage of …". This declarative statement referring to the less commonly used technique, namely, the think aloud protocol compared to other historically common data collection methods such as interviews gives the impression that the statement makes sense.

The rest five questions are as follows: two 5×s, two 4×s and one 3×s (Q7 and Q9, Q4 and Q8 and Q10 respectively). With regard to Q7, the problem for those who got it incorrect by marking it as F could be in the qualifier "quite" preceding the word "possible", otherwise, the statement is a straight True. There is little explanation why would students get it incorrect particularly they are familiar with such couplets (author and reader, theory and practice, product and process … etc) associated with words such as clash, tension, contradiction, … etc, in translation studies. As for Q9, the statement is copied straightforwardly from the taught material in the course. It is a true statement, answering it correct almost entirely depends on the student's understanding of the course material and the degree of retention of information in the exam. Moreover, the phrase "more or less" detaches it from being absolute and can be used as a cue to hit a correct answer. In Q4, the strong starting phrase beginning with "unlike" attached to "note taking" makes one pauses before taking any unduly decision. The 100% true afterwards sentence and the apparently true preceding phrase from the point of view of the student credit the whole statement to be true. Contrary to what sounds to be true, Q4, in fact, is false bearing in mind that "note taking" can also be "distracting". With regard to Q8 the starting phrase "because of the nature of interpreting" is widely used to the extent it is almost becoming a clichés for the fact that interpreting is a unique specialty. It gives the impression that the statement is true particularly when associated with what comes later: an unarguable association between interpreting and cognition. In fact the statement is false. The fact that 3 students got Q10 incorrect, the least ×s it got among the other questions, is not a big deal though it does not deserve the effort to think whether it is T or F. It is a straight true statement based on the often flagged out distinction between "References" and "Bibliography" in research methodology courses. Getting it incorrect can be explained by an ongoing misunderstanding on the part of the students of the no-difference between the two terms. The fact it is stated otherwise in the course can easily pass unnoticed as students would pay little attention to what might be considered by them as only trivial information. What explains it even further is that the three students who got

55

it incorrect (S8, S18 and S19), they also got the least √s (4√s, 5√s and 6√s respectively) compared to the rest of the students.

With regard to the second hypothesis, marks scored in each of the other four questions (II, III, IV and V) on the exam sheets do not show that the T or F question (QI on the exam sheets) scored the highest mark by most of the individual students. In fact, it is the second lowest among the five exam questions; only four out of the twenty three students had it the highest mark compared to scores gained in the other questions. See Table 2 below:

Table 2: **Total of highest and lowest scores per question per students.**

Students	Questions					Highest score
	QI	QII	QIII	QIV	QV	
s1	9	7	8	5	8	QI
s2	7	3	4	7	9	QV
s3	7	8	4	4	3	QI
s4	7	6	6	7	7	QI, QIV and QV
s5	7	8	7	5	6	QII
s6	5	9	3	7	4	QII
s7	8	5	3	10	8	QIV
s8	4	3	5	5	6	QV
s9	7	1	4	7	5	QI and QIV
s10	5	9	8	10	5	QIV
s11	8	6	8	9	6	QIV
s12	6	10	5	7	4	QII
s13	7	8	7	4	9	QV
s14	7	10	10	9	9	QII and III
s15	5	3	5	4	2	QI and QII
s16	6	6	8	10	6	Q IV
s17	7	9	6	5	4	QII
s18	5	2	4	2	7	Q V
s19	6	4	3	8	4	QIV
s20	7	4	7	10	9	QIV
s21	8	2	8	9	5	QIV
s22	6	9	7	6	5	QII
s23	7	5	8	10	9	QIV
Total highest score	**4**	**7**	**1**	**10**	**5**	

From Table 2 above, we notice that, interestingly, QIV (Discuss) scored the top mark (10) as ten students out of the 23 had got it as the highest mark out of the five questions; five students got a full mark (10 out of 10). On the other hand, Q1 (T or F) scored the second lowest as only four students had it the highest mark on their exam sheets; none of the students got a full mark in QI (T or F).

8. Discussion

Findings from the above (Table 1) show inconsistency between the vertical and horizontal columns with regard to the lowest ×s (per question) and the highest √s (per student) respectively. Findings also prove that the belief among students that T or F questions are easy to handle is only a misconception. In fact, they can be as perplexing as any other type of questions, if not more.

It would be an extra advantage if students were observed whether they answered the T or F question first being on the top of the answering sheets.

Whether answered first, last or in between could be an indication of students' attitude towards this type of question. Also, apart from the one student (S21) who submitted a draft with the answering sheet, there is little that can be used as a proof of reviewing the answers. Temptation to change answers was, however, noticed on a number of occasions. What cannot be determined is whether such reconsideration of answers that had already been made were based on detecting inter-item cues (clues in one item which tell the examinee about another) or not. For example, S13, who is on average one of the top of the list as previous performance record tells and the overall grade of the present exam shows, corrected three answers (Q1, Q3 and Q8) but, in fact, got them all incorrect. The amendment S13 made outside the space provided (2 Ts and 1 F respectively), vigorously crossing the previously chosen ones inside the brackets, led to the deduction of 3 marks that could have been secured by maintaining the first choice. If she did not amend the answers, she would have scored a full mark: 10 out of 10. What cannot be detected, however, whether this is done for a good reason from the point of view of the student or otherwise. In other words, only a direct question to the examinee to obtain an answer would solve the puzzle, otherwise it can only be speculated that this is due to over-analysis of the statement. On the other hand, the single draft submitted shows that three questions (Q7, Q9 and Q10) were left unanswered on the draft in pencil and answered correctly on the final draft in pen, apparently just prior to submission after all the other exam questions were answered on the draft. Also, Q5 is incorrectly marked (T) on the draft but correctly (F) on the final version; an indication of a second thought on part of the student after revision.

A study based on product only data like the present one would not show what strategies students incorporated to answer the questions; strategies such as those suggested as tips in test taking manuals or even those based on common sense. Only three students circled and/or underlined parts of the statements to highlight key phrases or clue words that they probably think they provide insight into the questions (usually qualifiers). This, however, does not always correlate with right answers.

When not absolute qualifiers are used in the T or F statements, there is no indication that students used the strategy of substituting other qualifiers to check which one makes the answer just right. Students are usually advised to use such strategy among others in cues lists to guide them to be test-wise with this type of questions to score marks. Despite the fact that there is no way to find out whether students used the guessing strategies to answer the questions, resorting to such a strategy is not ruled out since no penalties are incurred for guessing such as marks deduction from the right answers when the incorrect choice is made.

By means of scrutinizing certain answers, signs of quibbling were noticed in five instances. Writing T in a form of either T or F and vice versa in a pen and paper exam form is seen here as a sign of uncertainty that the student is not sure of the right answer. Moreover, it can be considered as a tactic to leave the choice to the teacher to consider the answer right either way in favour of the student. Interestingly enough, on the other hand, only the top two students, who usually rigorously follow instructions and carefully avoid any self-imposed interpretation, used the two symbols √ and x instead of the two letters T and F instructed in the question. This is taken here as an extra precaution on the part of the students not to leave room for speculation on the part of the teacher to read F as T or vice versa.

From the results presented above, one can assume that getting overtly clear statements to be T or F incorrect could be a result of unnecessary over analysis of the statement under consideration on the part of the student. However, rushing to move to the other questions on the exam sheets (four more, some of which are essay type) could also be a possible reason.

57

Finally, acquaintance with the individual students involved in the test could help to generally categorize them as those who "work smarter, not harder" and those who do not. For example, those who unnecessarily wrote comments to explain their choices or voluntarily provided corrections to the false statements, forgetting that this type of questions, as instructed, requires no explanation, they only wasted time and effort that could have been invested otherwise and elsewhere. Those who used a pencil first and then a pen afterwards to give themselves a chance for a second thought only missed the opportunity of the intuitive snap decision opportunity based on gut feeling, not necessarily rational, that usually works with this type of questions.

58

9. Conclusion

With reference to the aim stated early in this paper, the overall marks obtained by students show that, in general, they managed to handle the True or False questions relatively well. When it comes to details, however, students did make some unjustified incorrect choices. Some of such incorrect choices can be attributed to certain reasons while others, as discussed above, are hard to explain. One obvious reason for choosing incorrect answers could be that simply students did not study well enough to be able to retrieve information. Another reason could be they just read the test in its totality overlooking important details that offer clues to the correct answers although, in general, with regard to the students' ability to manage using cues and clues, it has been proved to be relatively true. Interestingly, however, such ability does not often correlate with the individual student's overall performance. In other words, a very good student with an overall high mark can get a true or false question that includes obvious clues incorrect while an average or weak student gets a correct answer.

With regard to the main hypothesis here, findings from the study show that answering T or F questions in exams is not, as some might think, a straightforward process that secures better scores in comparison with other types of questions or, in other words, guarantees obtaining the highest marks for individual exam takers. In fact, they can be as difficult to handle as other types of questions. Moreover, familiarity with such type of questions does not necessarily guarantee that students master the skills to answer them.

One observation emerged from the study is that to prevent students quibbling while answering T or F questions, it would be better if T and F symbols are provided on the test sheet and the choice between the two is made by circling the selected letter rather than writing it down. This advice, among other tips, is also given by the University of Texas, Faculty Innovation Center https://facultyinnovate.utexas.edu/teaching/check-learning/question-types/true-false, and here it is proven that it would work better than asking students to write the symbols down.

Finally, T or F type of questions are meant to save time and trigger fast response. Accordingly, intuition plays a major role in answering them. Hence, training how to handle such type of questions is recommended at the early stages of the learning process to unable students handle such type of questions in exams with no misconceptions.

To sum up, in the wider perspective of translation education and in the context of the "Changing Landscape in Translation", and the "Rapid and Radical Changes in Translation and Translation Studies", (Gabier, 2014 and 2016 respectively) nothing is trivial. Despite its limitations such as the fact that it can hardly be of use to evaluate students' performance in advanced translation proper tasks, complex and compound types of such test can be beneficial in early translation training when the choice between True or False translation is

accompanied by giving opinion and explanation. This type of test can also be beneficial in testing knowledge of a broad range of content in the area of translation studies and translator education. It can verify that knowledge has been acquired. It can also benefit training students in research methodology specifically that employed in translation studies research. It opens students' eyes on what to do and not to do when conducting research in translation studies. What merits this type of test to be integrated in the pedagogical practice in translation and incorporated into the learning process is the fact that it is impartial while all translation assessment methods are often described to be neither just nor objective. Another potential area that can benefit from this type of test is assessing the development of the individual's aptitude or the applicants' aptitude to study translation and interpreting since most other aptitude tests are criticized. According to Campbell and Hale (2003: 212). "common to all aptitude tests described in the literature are the competencies, the test aims to assess, the subjective marking criteria, and the high failure rate".

Having stated the above, with the limitations of the present study for the reasons discussed earlier, findings presented here are not intended to be unreservedly generalized. They can only be considered as indications of certain tendencies that can be further confirmed under different research conditions. Moreover, the findings in the present study are based on exam grades, it would be beneficial if similar studies are conducted to assess other course outcomes such as satisfaction and metacognition.

References

Anderson, L. et al. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York: Pearson, Allyn & Bacon. Retrieved 28 Feb 2017.

ARTIS. (n.d.) https://initiative.org/.

Baker, M. (2009). In other words: a textbook on translation. 2nd edition. London and New York: Routledge

Berber, D, and Birgit Kretschmann. (2016). "The multilingual approach to interpreter training as a solution". *Journal of Translator Education and Translation Studies* 1(1): 94-104. Retrieved from ??? http://tinyurl.com/j5jtv5f. Accessed October 2016.

Biggs, J. and Collis, K. (1982). *Evaluating the quality of learning: the SOLO taxonomy*. New York, Academic Press

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). *Taxonomy of educational objectives, handbook I: the cognitive domain*. New York: David McKay Co Inc.

Bologna process. (n.d.). http://www.wg.aegee.org/ewg/bologna.htm. Retrieved, 28 Feb 2017.

Campbell, S. and Hale, Sandra. (2003). "Translation and Interpreting Assessment in the Context of Educational Measurement.". In Gunilla. M. Anderman and Margaret Rogers (eds), *Translation Today: Trends and Perspectives.* 205–220. Clevedon: Multilingual Matters.

Center for Innovation in Teachign and Learning (CILT). (n.d.). https://citl.illinois.edu/citl-101/measurement-evaluation/exam-scoring/improving-your-test-questions. Retrieved 28 Feb 2017.

Chandler, D. (1993). Writing strategies and writers' tools. *English Today*, v9 n2 pp32-38.

Clark, R., Chopeta, L. (2004). Graphics for learning : proven guidelines for planning, designing, and evaluating visuals in training materials. San Francisco: Jossey-Bass/Pfeiffer.

Clark, R., & Mayer, R. (2008). Learning by viewing versus learning by doing: evidence-based guidelines for principled learning environments. *Performance Improvement*, 47, 5–13. Dublin descriptors (2004). "Shared 'Dublin' descriptors for Short Cycle, First Cycle, Second Cycle and Third Cycle Awards." A report from a Joint Quality Initiative informal group. Retrieved 28 August 2017 from http://www.jointquality.nl/.

College Atlas. (n.d.). https://www.collegeatlas.org/true-false-exams.html. Retrieved 28 Feb 2017.

Colorado State University. (n.d.). https://tilt.colostate.edu/teachingResources. Retrieved 28 August 2017.

Ebel, R. (1971). *The comparative effectiveness of true-false and multiple choice achievement test Items*. Michigan State University.

Ebel, R. (1975). Can teachers write good true-false test items? *Journal of educational measurement*, Vol. 12, No 1, pp 31-35

Ebel, R. (1977). *The use of standardized testing*. The Phi Delta Kappa Educational Foundation Bloomington, Indiana.

Gambier, Y. (2009). *Challenges in research on audiovisual translation.* In Anthony Pym and Alexander Perekrestenko (eds.), Translation Research Projects 2, Tarragona: Intercultural Studies Group, pp. 17-26.

Gambier, Y. (2014). Changing landscape in translation. *International Journal of Society, Culture & Language*, 2(2),

Gambier, Y. (2016). Rapid and Radical Changes in Translation and Translation Studies. *International Journal of Communication*, Vol. 10, pp887-906

García, Ignacio (2009). *Research on translation tools. .* In Anthony Pym and Alexander Perekrestenko (eds.), Translation Research Projects 2, Tarragona: Intercultural Studies Group, pp. 27-34.

Gile, D. (1990). Scientific Research Vs. Personal Theories In the Investigation Of Interpretation. In Gran, Laura & Christopher Taylor (eds). *Aspects of Applied and Experimental Research on Conference Interpretation*. Udine: Campanotto Editore. pp.28-41

Gile, D. (2009). Research for training, research for society Tarragona: Intercultural Studies Group, 2009. pp. 35-40. Retrieved 28 August 2017 from http://isg.urv.es/publicity/isg/publications/trp_2_2009/index.htm

Goff-Kfouri, C. (2004). Testing and evaluation in the translation classroom. *Translation Journal* 8 (3).

Huitt, W. (2009). Bloom et al.'s taxonomy of the cognitive domain. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University

Lamb, W. (1976). Ask a higher-level question, get a higher-level answer. *Science Teacher*, 43(4), 22-23.

League of European Research Universties. (2010). *Training talented researchers for society: Doctoral studies beyond 2010*. Retrieved 28 August 2017 from http://www.leru.org/files/publications/LERU_Doctoral_degrees_beyond_2010.pdf

Life Solutions. (n.d.). http://www.drmontedavenport.com/true-false-questions/. Retrieved 28 Feb 2017.

Liu, O. L. (2009a). Measuring learning outcomes in higher education (Report No. RDC-10). Princeton, NJ: ETS.

Martinez Melis, Nicole and Amparo Hurtado Albir. (2001). Assessment in translation studies: research needs. *Meta* 46:2. 272-287. Available as pdf at http://www.erudit.org/revue/meta/2001 /v46/n2/003624ar.pdf

Monkey See. (n.d.). http://monkeysee.com/what-are-some-tips-for-answering-true-or-false-questions/. Retrieved 28 Feb 2017.

Moser-Mercer, B. (1994). Aptitude testing for conference interpreting: why, when and how. In Sylvie Lambert and Barbara Moser-Mercer (eds). *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, 57−68. Amsterdam/Philadelphia: John Benjamins.

Reisz, Matthew. (2008). Doctor, doctor, quick, quick. *Times Higher Education*, 4 December 2008, pp. 31-35.

Öner, S. (2013). Translator as researcher: perspectives on training and life-long professional improvement of translators. *Procedia - Social and Behavioral Sciences* 106 (2013) 3159 – 3166.

Pennington Publishing. (n.d.). http://blog.penningtonpublishing.com/study _skills/the-top-nine-tips-to-taking-true-false-tests/. Retrieved 28 Feb 2017.

PRES (2007). "Postgraduate Research Experience Survey". Retrieved 28 August 2017 from https://www.heacademy.ac.uk/knowledge-hub/postgraduate-research-experience-survey-2007.

Pyme, A. (2016). Teaching translation in a multilingual practicum. Retrieved 29 August 2017 from http://usuaris.tinet.cat/apym/on-line/training/2016_ multilingual.pdf. Retrieved 28 Feb 2017.

Schäffner, Christina. (2009). Research training: how specific does it need to be? In Anthony Pym and Alexander Perekrestenko (eds). *Translation Research Projects 2. Tarragona: Intercultural Studies Group, 2009*. pp. 9-15.

Treser, M. (2015). True/false questions: getting them right. Retrieved 28 Feb. 2017 from https://elearningindustry.com/true-false-questions-getting-right,

Univerisity of Waterloo. (n.d.). https://uwaterloo.ca/centre-for-teaching-excellence/teaching-resources/teaching-tips/developing-assignments/exams/ questions-types-characteristics-suggestions. Retrieved 28 Feb 2017.

University of Texas. (n.d.). https://facultyinnovate.utexas.edu/teaching/check-learning/question-types/true-false. Retrieved 28 Feb 2017

Williams Jenny & Andrew Chesterman (2002). The Map – A. beginner's guide to doing re- search. Manchester, UK: St. Jerome.