Is Self-assessment a Credible Indicator of Translation Competence?

Falih S. M. Al-Emara
University of Leeds
Leeds, United Kingdom

Abstract
In his inverse translation competence model Campbell (1998) investigates self-assessment as an element in the monitoring of translation into the second language. Relatedly, the current study attempts to test the applicability of that model on translating into the first language through a conceptual replication experiment. Thus, it presents a quantitative methodological analysis of multiple translations produced by student-translators. Hence, this paper, as a part of a larger study, reports on the replication of only one of the two dimensions of monitoring (self-assessment and real-time editing). Specifically, it investigates whether the student-translators' general assessments of their own ability to translate relate to their translation competence and, if so, to what extent do they specify it. The results of this study lead to the conclusion that self-assessments, based on the participants' awareness of their output, are less credible in estimating translation competence than the general assessment of the tutor, and of quality assessment by multiple raters.
Keywords: quality assessment, self-assessment, translation competence, tutor's assessment

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

1. Introduction

This paper describes an empirical investigation of the notion of self-assessment as an element in translation monitoring, which is a component in Campbell's (1998) translation competence (TC) model. Translation monitoring is one of the three components of Campbell's model which includes, in addition to monitoring, textual competence and disposition. This study is based on a conceptual replication of this model of translating into the second language (L2) to test its applicability on translating into the first language (L1)[1]. However, admits that textual competence and disposition were studied because of the need for a "general model to underpin the teaching and learning of translation" (p. 126), whereas translation monitoring was studied for a purely practical need. It was motivated by the difficulty of convincing students who fail assignments about their real abilities since "(…) they often expressed inordinate surprise; some students seemed to think they were much better translators than they really were" (Campbell, 1998, p. 126). Studying this matter, led him to the inference that it was a facet of TC.

In addition, Campbell (1998) admits that this component was not 'theoretically underpinned', which was a major problem in investigating it. He was at pains to mention it, and stressed that his study owes nothing to proposals of monitoring in the context of language acquisition and cognitive psychology, such as Krashen (1977) and O'Malley & Chamot (1990). Thus, Campbell's (1998) study is based on purely empirical evidence derived from empirically investigating this problem in particular (p. 153).

The current paper reports on the conceptual replication of only one dimension from the two dimensions which Campbell tackles in his investigation of monitoring competence. To clarify, his original study of monitoring includes self-assessment and real-time editing. The first dimension, which is the subject of this paper, refers to the students' general assessment of their own ability to translate and how it relates to the other components of TC. So, Campbell assumes that their awareness of the quality of their output (self-assessment) can be proposed as a relevant factor in the characterization of TC and, consequently, as one of its indicators. However, the other dimension, for which an independent paper will be devoted, deals with the translator's opportunity to intervene to improve the output through real-time revision.

2. Definition of Self-assessment

This term is defined by the Oxford English Dictionary as the "assessment or evaluation of oneself or one's actions, attitudes, or performance". Similarly, it is defined by the Oxford US English Dictionary as the "(…) assessment or evaluation of oneself or one's actions and attitudes, in particular, of one's performance at a job or learning task considered in relation to an objective standard". Yet, it is a term that is widely used in English in almost all fields of life when judgments are required or made by a person about issues concerning him, his actions and his performance of tasks. The sense in which this term is used in the present study is limited to the one used in learning and teaching, especially of language and translation. However, there are a few studies about self-assessment in translation studies (TS) which will be reviewed after defining and outlining self-assessment in teaching and learning in general.

Generally speaking, self-assessment in teaching and learning is a relatively new concept that is applied and practiced in the processes of learning and teaching at large. It is defined by Boud (1991, cited by Mills & Glover, 2007, p. 2) as "(…) the involvement of students in identifying standards and/or criteria to apply to their work, and making judgments about the

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

extent to which they have met these criteria and standards". As Mills and Glover agree to this definition, they suggest that a student's involvement in the activity of self-assessment develops his reflection and analysis abilities of both his work and the learning outcomes.

Similarly, Blanche and Merino (1989, p. 313) define self-assessment as the information about the learners provided by the learners themselves, about their abilities, the progress they think they are making and what they think they can or cannot do yet with what they have learned in a course. While Coranado-Aliegro (2008, 1-3) relates self-assessment to what he calls 'self-efficacy', and contemplates that self-assessment is basically the feeling of mastery which the learner develops over a given task that he performs.

Likewise, Harris and McCann (1994:36) stress self-assessment as a monitoring of the learning process in formative assessment and feedback helps to establish self-regulation in the learners which positively affects their learning and their progress towards their goals. In like manner, Lenski, Ehlers-Zavala et al (2006:32) consider self-assessment as one of the basic instruments to assess language learners' progress, especially when they 'self-monitor' their performance tasks. Similarly, Gardener (2000:49) assumes that self-assessment is an effective tool in autonomous learning as a testing device for both accreditation and self-monitoring. While Dickinson (1987), earlier than Gardener, stresses the importance of self-assessment or 'self-evaluation' for language learners in general and for autonomous language learners in particular. By the same token, Rust (2002) views self-assessment as a device to help learners monitor their level of success in specific learning tasks.

3. Reliability and validity

Since self-assessment could be taken as a measure of the learners' awareness of the quality of their performance, a look at the questions of its reliability and validity is essential. Ross (2006, p. 3), for instance, concludes that the "psychometric properties of self-assessment" indicate that it is a reliable technique to assess and to yield consistent and dependable results. However, when considering validity, he arrives at the general conviction that, students commonly give higher estimations of their performance and abilities than what their tutors give them. Formerly, Boud and Fachikov (1989) suggest that overestimations are more likely to be found where the self-assessment contributes to the student's grade on a course. Whatever the discrepancy between student and teacher assessment, it cannot be attributed but to what each party assesses. After reviewing a number of studies, Ross (2006, p. 4) submits that, though self-assessment studies give information about student achievement, such information corresponds only partially to the information given by teacher assessments. The variation is attributed to 'interest bias' or to the assessment criteria and information.

Nevertheless, MacIntyre, Noels et al (1997, p. 265-28) focus on the role of language anxiety in instigating biases in self-ratings of second language proficiency. This goes in line with the review made by Blanche and Merino (1989, p. 315), who determined that when the skills to be assessed are clear and detailed self-assessments will consistently agree with the ratings given by external measures. However, agreement of student self-assessment with external measures cannot be taken for granted because students do not necessarily assess accurately. The authors confirm that language learners mostly overestimate or underestimate their proficiency in language. This, of course, leads to the failure of the assessment to correspond to objective external measures such as tutor's assessment.

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

Arguments continue as to whether self-assessment is reliable. Dickinson (1987) and Blue (1988), for example, question its reliability and favour assessments made by teachers and specialists. Other studies arrived at similar findings such as Janssen-van Dieten (1989) and Thomson (1996) also discredit the reliability of the learner's self-assessment and favor the teacher's evaluation. Conversely, there are studies which accredit the reliability of self-assessment such as Bachman & Palmer (1989) and Blanche (1990).

In fact, it is difficult to account for the inconsistency in the findings of the various studies regarding the issue of reliability due to the differences in the variables which decide, to a large degree, the reliability of the findings. In other words, factors such as the size of the sample, the suitability of the setting, the clarity of the directions and the efficiency of administration definitely affect the reliability of a test. In addition, the characteristics of the participants in the studies, including their age, sex, education, social and cultural background, and the skill and experience they have in self-assessment procedures, all contribute to that variation. Other variables like the test format and the skills being compared can act as additional sources of reliability variation (Bachman, 1990, p. 160-223).

The question of whether to use self-assessment as a measuring device on its own or to compare it with some other well-established external criteria, in terms of validity and reliability, is challenging. Nonetheless, seeking measures with absolute validity and reliability in measuring skills related to language learning could be futile because of the improbability of fully controlling all the variables involved in the process. Consequently, it seems acceptable to use self-assessment and tolerate its margin of error in the same way other measures are accepted and adopted. This conclusion agrees with Gardner's sum up of the conflicting notions and arguments on using self-assessment (Gardner, 2000, p. 53). Still, pedagogically speaking, self-assessment is one of the tools that are stimulated in the more modern learner-centered approach to language teaching and fit under the social constructivist [Swan, 2005] paradigm of learning (Saltourides (2006, p. 55).

4. Self-assessment in TS

The studies that have investigated the use of self-assessment in TS, both in translation and interpretation, are scarce. Below is a brief survey of the most focal ones which highlight the function and vitality of this measure.

4.1. Self-assessment in interpretation research

Self-assessment is recommended and employed in the case of training interpreters to improve the quality of performance. Chiaro and Nocella (2004, p. 291), for example, suggest three main areas of operation including training, where the procedures of self-assessment are incorporated. Relatedly, Fowler (2007) empirically investigates the role of self-assessment, along with peer assessment and evaluation, in the training of professional interpreters. It is an attempt to validate the use of those forms of assessment and to inform trainee interpreters to use the feedback in their professional performance. The study concludes that self-assessment, along with peer assessment, are necessary in the training of interpreters because they foster self-awareness of the flaws and errors which accompany performance.

Similarly, Bartłomiejczyk (2007), in a seminal study, recruits eighteen subjects at the same stage of training and asked them to self-assess their output after they interpreted a text from

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

English into Polish in the light of their strategic processing. The results indicate that there was a noteworthy tendency to negatively assess their performance in terms of faithfulness to the original text and to its completeness, with almost complete negligence of matters related to presentation such as voice quality, intonation, pauses and hesitancy. In addition, she concludes that the results of the study generally cast some doubts on the appropriateness of the procedure of self-evaluation when conducted in the same unstructured way applied in her study. As a result, she suggests the use of assessment sheets, similar to those recommended by Schjoldager (1996) or Hartley, Mason et al. (2003), to attain better results in diagnosing problems of interpreter output (Bartłomiejczyk, 2007, p. 263-4).

Lee (2005) also investigates the usefulness of self-assessment in the teaching of interpretation. Graduate students of translation and interpretation are asked to self-assess their performance to their tutor. The results of the survey disclose that there are, from the point of view of the trainee interpreters, positive aspects in the identification and diagnosis of weaknesses and strengths, enabling them to orient their practice and to allow them monitor and appraise their progress.

Likewise, Arumi and Esteve (2006, p. 159) believe that assessment and self-assessment procedures form a component in the training of interpreters which plays an important role by encouraging self-regulation processes in consecutive interpreting. Postigo Pinazo (2008, p. 208) agrees on the importance of training interpreters to self-assess their knowledge and ability stressing the necessity of developing the habit of identifying weaknesses and strengths, gaps in their knowledge and utilization of skills. She concludes with the suggestion to integrate self-assessment into the different aspects of teaching.

4.2. Self-assessment in translation research
A pioneer study, by Fanghanel and Voela (2001), accomplishes through encouraging nine postgraduate students doing their masters in translation to perform formative self-assessments, contend that it is problematic for two reasons. The first is the way of dealing with the notion of "correctness" in translation which, unlike most other disciplines, does not yield itself well to this notion. The second is associated with the nature of translation as an interdisciplinary activity which comprises various "cognitive, social, textual and pragmatic skills and knowledge" (p. 47).

Correspondingly, Martinez and Hurtado (2001, p. 285) consider student self-assessment records as one of the basic evaluation tools in translator training, along with other tools such as teacher's observation records, translation diaries, documentation sources, error inventories and so on. Similarly, Kose (2011, p. 484-85) uses self-assessment scales to identify the levels of his subjects' language skills in his study of the effect of form and meaning in translation focused instruction. His self-assessment scale (p. 488) includes six skills: reading, writing, listening, speaking, grammar, and vocabulary and idioms.

In the same way, in a vital study, Waddington (2001, p. 311-325) employs self-assessment, along with teachers' assessment and a number of other factors, as a tool in the identification of TC. In the study, he uses students' self-assessment of their ability to translate from Spanish into English. This study is extremely important in that it statistically discloses the relationship between self-assessment and TC. The correlations were significant between TC and native language competence and self-assessment. In the conclusion, Waddington contends that

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

the main underlying factor is TC, which is closely related to student self-assessment of that aspect and to student native language competence.

A more recent study, conducted by Robinson, Lopez Rodriguez et al. (2006), investigates the introduction of e-learning in the Spanish university system and the opportunities it has provided to 'reorient translator training'. The study concludes with highlighting the importance of self-assessment describing it as a 'logical component' in translator professional training (p. 136).

Finally, a recent study by Fernandez and Zabalbeascoa (2012), which is very close to the current one in its aims and procedures, has investigated the relationship between self-assessment and the performance of trainee translators by correlating their self-evaluation results, based on their answers to post translation metacognitive questionnaires, with their teacher's assessment. It has focused on the trainees' identification of translation problems and the justification they give for their own solutions to those problems. It was revealed that the "best-performing students were more strategically and translationally aware in self-evaluating their own translating" (p. 463). The study concludes with the affirmation that there is a significant correlation between the students' self-evaluation and their level of performance in terms of identifying and solving translation problems. Thus, it confirms (p. 476) that the pedagogy and training must aim at raising this awareness to improve the translation performance of trainee translators.

In conclusion, the use of self-assessment in translation studies research has revealed that it is a relatively appropriate mode to be used in translator and interpreter training because it ensures the trainee's involvement and amplifies the sense of responsibility towards learning and future work. It is also typically associated with the assessment of the tutor to a degree that can be described as an established relationship. Most of the studies which were reviewed above show the dependency of self-assessment study on the tutor's evaluation as an external factor to establish its relevance and dependability.

5. The Current Study
5.1. Aim
It aims to examine the relationship between the participants' self-assessment and their tutor's assessment as an external measure. In doing so, it intends to confirm or otherwise falsify the assumptions made by Campbell (1998, p. 135-6) that students have good awareness of their ability to translate into their L1. It also aims to explore the extent to which students may consistently overestimate or underestimate that ability. Particularly, Campbell's (1998, p. 136) conclusion that "Arabic students greatly overestimate their ability [to translate] into their first language" is of great interest in this experiment. It is important here to empirically test the idea that students' general assessment of their own ability to translate, validated by its correlation with their tutor's assessment, relate to the other components of TC and can be proposed as a relevant factor which assists in its characterization.

5.2. Participants
The participants of the study were twenty-five MA student-translators for whom Arabic is L1 and English is L2, taking their courses at university in the UK. At the time of the study, they were enrolled in a module of English into Arabic translation. They were doing their Masters in translation at the University of Durham (18 participants) and the University of Salford (7

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

participants). They were 14 females and 11 males, with an age range between 22 and 41 years, and an age average of 28.28 years.

However, only eighteen participants of them were enrolled in the self-assessment study because one of the tutors declined to give his assessment for private reasons. Thus, the study was disadvantaged by the lack of tutor assessment of seven participants, which reduced the sample to eighteen participants only. This matter weakened but did not eliminate the ability of the study to investigate tutor rating reliability, as opposed to participant self-assessment albeit on a smaller sample than it was originally desired. Consequently, the results of quality assessments (section 6) were used as an additional external measure to secure the reliability and objectivity of the initial findings.

5.3. Data

It comprises qualitative data derived from the participants' self-assessments of their output on two texts they translated into their L1 (Arabic) in an experiment which was originally conducted for collecting data for a PhD research. The texts were two prose written texts in English, 220 words each, taken from press editorials. Some parts of the original texts were taken out for practical considerations of brevity without affecting the overall meaning and build-up of the texts. A question was addressed to the subjects to self-assess their output at the end of translating each text on a continuum of ten points. In addition, the tutor was asked to give a general cumulative assessment of the level of TC of those particular participants, based on his sustained observation during their studies.

The choice of press editorials over other genre types for this study is well justified and is accepted as a proper selection criterion. Campbell (1998) believes that "while many of the other genres are represented in the materials of translator training courses, this type seems to predominate and is very typical of accreditation examination scripts" (p. 76).

Fortunately, all the participants made self-assessments of their performance on both texts, which made correlation with the tutor's assessment possible.

5.4. The results

The eighteen participants completed assessments on a continuum of ten values after they finished translating a text as follows:

On the scale of ten below, please, estimate your translation quality of the above text by ticking the box below the score of your choice (10 being the highest):

1	2	3	4	5	6	7	8	9	10

It is the same scale used by the tutor in his general assessment of the participants. In fact, the scale and the criteria that were used by both were easy and explicit. The two texts are referred to as Text One [T1] and Text Two [T2] from now onward. The results, which were ranked and displayed in Table 1 below, illustrate the following:

a) Tutor assessment starts at score 4 on the assessment scale and extends to the highest score (10).

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

b) The frequency of tutor assessment bunches at scores 5, 6, 7, and 8, comprising the majority of the participants (14).

c) Participants' assessments of their performance on T1 slightly differ from that of T2 in both the range of assessment and constellation.

d) In T1 the assessment, analogous to that of the tutor, starts at score 4 but, dissimilarly, ceases at score 8. The frequency of the results bunches at scores 5, 6, 7, and 8, (17 participants) with score 7 being the most frequent.

e) By contrast, participant assessment of T2 starts at a lower score (2) than that of T1 and of the tutor. However, it extends to the same range of T1 at score 8 only. Frequency bunching is a bit different; starting at score 4 to 5 then to 7 and 8, excluding 6 where only one participant opted there. Score 5 was the most frequent.

Table 1 *Tutor and Participant Assessment Ranking*

Scale Values	1	2	3	4	5	6	7	8	9	10
Tutor's Assessment	0	0	0	1	3	4	4	3	1	2
T1 Self-assessment	0	0	0	1	4	3	7	3	0	0
T2 Self-assessment	0	1	1	4	5	1	3	3	0	0

The differences in assessment are reflected in Table 2 which shows the match and mismatch between the tutor and participant assessment on each text for each participant. When participant assessment is higher than that of the tutor the deviation is positively marked with a plus (+), and when it is lower it is negatively marked with a minus (-), whereas matching assessment between the participant and the tutor is marked with (0) disparity. The positive marking indicates overestimation, whereas the negative one indicates underestimation.

Table 2 *Participant's Over/Under-estimation*

Participant	Tutor scores	Participant- T1		Participant- T2		Disparity Mean
		score	disparity	Score	disparity	
1	8	7	-1	5	-3	-2
2	6	8	+2	8	+2	2
3	5	7	+2	4	-1	0.5
4	7	8	+1	8	+1	1
5	7	7	0	7	0	0
6	5	5	0	2	-3	-1.5
7	4	5	+1	5	+1	1
8	6	7	+1	7	+1	1
9	6	7	+1	5	-1	0
10	8	7	-1	4	-4	-2.5

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

11	7	5	-2	4	-3	-2.5
12	7	6	-1	7	0	-0.5
13	10	8	-2	8	-2	-2
14	10	6	-4	5	-5	-4.5
15	8	7	-1	6	-2	-1.5
16	9	5	-4	5	-4	-4
17	6	6	0	3	-3	-1.5
18	5	4	-1	4	-1	-1
Mean	6.89	6.39	-0.5	5.39	-1.5	-1

The participants' estimation results in Table 2, are summarized in Table 3:

Table 3 *Estimation Summary*

Level of Self-Estimation	T1		T2		Mean	
	No.	%	No.	%	No.	%
Over-estimators	6	33.33	4	22.22	5	27.78
Matching-estimators	3	16.67	2	11.11	2.5	13.89
Under-estimators	9	50.00	12	66.67	10.5	58.33

As it is hoped to disclose whether the overestimation or underestimation of one's performance notably relates to high or low levels of participant TC, The results reveal that over-estimators for T1 represent only one third of the sample whereas under-estimators for the same text represent one half of the population. On the other hand, underestimation in T2 is stronger than that in T1 as just a little more than one quarter of the participants overestimated their performance, whereas the other two thirds underestimated their output.

5.5. The statistical analysis

The results of the correlations between the participant and the tutor assessments are displayed in Tables 4 and 5 below, which show significant correlation between the participant assessments on both texts. This indicates the reliability of participant assessment despite the difference between the texts in the level of difficulty and structure. By contrast, the absence of a significant correlation between tutor and participant assessment reflects the lack of validity in the assessment.

Table 4 *Tutor and Self-assessment Correlations*

Correlations		T1	T2	Tutor
T1	Pearson Correlation	1.000	0.669**	0.310
	Sig.(2-tailed	-	0.002	0.211
	N	18	18	18
T2	Pearson Correlation	0.669**	1.000	0.323
	Sig. (2-tailed)	0.002	-	0.191
	N	18	18	18
Tutor	Pearson Correlation	0.310	0.323	1.000

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

	Sig. (2-tailed	0.211	0.323	-
	N	18	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

5.6 Interpretation of the results
These results can be interpreted in the following way:

a) The significant correlation of participant assessment between the two texts indicates the reliability of participant self-assessment across the two texts (whether they are overestimating or underestimating).

b) If we take into our account the fact that the participants are native speakers with considerable control on their language, the results do not agree with Campbell's (1998, p. 136) statement that "Arabic students greatly overestimate their ability [to translate] into their first language". They rather agree with another statement by him that the ability to self-assess one's translation ability is related to language competence and differs "(…) more fundamentally between types of bilingualism and that poor language competence is linked to overestimation and good language competence to under-estimation (Campbell, 1998, p. 137)". This entails that the good language competence of the participants could be responsible for their general tendency to underestimate.

c) Retrospectively, a more subjective interpretation based on the author's personal experience and discussions with specialists suggest another possible reason to this general tendency to underestimate in self-assessment. It is supposed here that the current situation can be partly attributed to the kind of teacher assessment those participants became used to in their past exams and assignments in their schools. As a subject of study, Arabic language is largely treated with dignity and reverence in the Arab World for reasons of nationality, education and most importantly of religion, especially by teachers of Arabic who are schooled in this prescriptive tradition to consider themselves guardians of the classical language. Consequently, under their effect, the students establish the conviction that only superior performance is expected and positively assessed. Congruently, the current study participants are likely to have transferred this experience to their personal assessment of their own output, which results in an underestimation of their translation performance in their L1. Unfortunately, it is not possible to scrutinize this issue in this study, and it deserves some future investigation.

6. Quality Assessment, Self-assessment and Tutor assessment
Due to the unexpected and untimely decline of one of the tutors, as mentioned earlier, which did not allow for adequate comparisons to ensure the objectivity of the results, it is sought to validate them against another external measure. This measure consists of the results of quality assessment of the translation of the two texts. Thus, copies of the translations of the participants (two texts each) were submitted to three expert raters to individually assess them according to an assessment chart. The chart was explained by an assessment sheet made up of a number of behavioral statements which describe the levels of output expected from translators on each aspect of the chart. The raters all were Arabic native speakers with experience in translation teaching and assessment. At the time of performing the assessment, two of them, a female and a male, were PhD holders in TS whereas the third (a male) has two bachelors; one in Arabic and one in English, an MA in Arabic and was doing a PhD in English Literature in a UK university. The assessment sheet was derived from the code of practice in the School of Languages, Cultures

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

and Societies at the University of Leeds, and vividly stated and simplified by the researcher to be used easily and reliably as follows:

Translation Assessment Sheet

Source Language [SL] Comprehension:

5--- Perfect comprehension with no traces of miscomprehension at all.

4--- Few comprehension problems slightly affect the translation.

3--- Minor comprehension problems partly affect the translation.

2---Predominant comprehension problems entirely affect the translation.

1--- Comprehension problems so severe that they distort the translation.

Command on Subject-matter:

5--- Full command of the subject-matter to carry out the translation.

4--- Few subject-matter problems which slightly affect the translation.

3--- Minor subject-matter problems which partly affect the translation.

2--- Predominant subject-matter problems which entirely affect the translation.

1--- Severe subject-matter problems which distort the translation.

Target Language [TL] Appropriateness:

5--- Completely appropriate TL.

4--- Few traces of TL inappropriateness slightly affect the translation.

3--- Minor TL inappropriateness problems partly affect the translation.

2--- Predominant TL inappropriateness problems entirely affect the translation.

1--- Completely inappropriate TL distorts the translation

Target Language Accuracy:

5--- Completely accurate TL.

4--- Few traces of TL inaccuracy very slightly affect the translation.

3--- Minor TL inaccuracy problems partly affect the translation.

2--- Predominant TL inaccuracy problems entirely affect the translation.

1--- Severe TL inaccuracy problems distort the translation.

The assessment was recorded on a chart designed by the researcher, comprising the four aspects that were assessed according to the guidance provided in the assessment sheet above. Each aspect was evaluated on a scale of five points. They start with number 5 as the score for the highest desirable output and end up with number 1 as the lowest possible score for the output. The total mark represents the sum of the scores a participant gets on the different components which, theoretically, does not exceed 20. Below is a sample of the assessment chart:

Translation Assessment Chart

Participant No. (   )   Rater (   )

Evaluation	5	4	3	2	1
Comprehension of the SL					

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

Command of subject matter					
Appropriateness of the TL					
Accuracy of the TL					

Total Mark: (   )
Rater's Signature

The assessments of each rater [R1, R2, and R3] on both texts were then recorded and displayed in separate tables to find out the relationship between the two texts in general and between each component across the two texts in particular. R1 results show that the mean of T1 scores (11.08) is higher than that of T2 (9.92). The variation could be partly attributed to the level of difficulty of each text. Yet, it is perceived that the total average score of both texts (21/40) is relatively low and could be attributed to the possible rigorousness of this rater. However, the total scores of the participants reflected considerable distribution ranging from a least score of 8 marks to the most score of 36 marks. This, in some way, indicates that the rater highly discriminates among the levels of performance. Generally, R1's evaluation of the different evaluated aspects and of the two texts has yielded the correlations summarized in Table 5:

Table 5 *Rater1- T1 vs. T2 Correlations*

Aspect	Correlation
Source Language Comprehension	0.688**
Command of subject matter	0.674**
Target Language Appropriateness	0.494*
Target Language Accuracy	0.606**
Sum of the Two Texts	0.690**

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

The results show that there is a statistically strong relationship among the various components and also between the results of the two texts. This also suggests the consistency of the rater's assessment.

Although R2 yielded slightly higher results than those of R2, almost the same pattern is retained in that higher scores were recorded on T1 (mean=14.08) as compared to T2 (mean=11). However, the total average score was nearly five marks higher than that of R1. The lowest score was 12 marks and the highest was 35 suggesting lower distribution of scores, and subsequently less discrimination ability than R1. Similarly, the correlations show that there is a statistically moderate relationship between ratings of the two texts and also among the four assessed aspects indicating the consistency of the rater, though relatively weaker than that seen in R1's results displayed in Table 6:

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

Table 6 *Rater2- T1 vs. T2 Correlations*

Aspect	Correlation
Source Language Comprehension	0.475*
Command of subject matter	0.465
Target Language Appropriateness	0.408
Target Language Accuracy	0.412
Sum	0.467

*. Correlation is significant at the 0.05 level (2-tailed).

Similarly, the results of R3 are not overtly different from the other two. The total scores and the average, of this rater, fall somewhere between their counterparts of the two other raters. Scores on T1 (mean=12.4) are similarly higher than those of T2 (mean10.88). The lowest score obtained was 9 and the highest was 32. However, unlike R2, this rater's assessments show significant correlations among the evaluated aspects. Table 7 below suggests that there is a very strong relationship among the four aspects suggesting very high consistency of the rater.

Table 7 *Rater3 T1 vs. T2 Correlations*

Aspect	Correlation
Source Language Comprehension	0.757***
Command of subject matter	0.683**
Target Language Appropriateness	0.796***
Target Language Accuracy	0.650**
Sum	0.790***

**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.001 level (2-tailed).

However, when the total results of the three raters are pooled together, it is found that the average rater correlations (0.752) show very strong relationships, and suggest the high reliability of the raters and the validity of the assessment procedure as shown in Table 8 below:

Table 8 *Correlations among the Three Raters' Assessments*

Rater	Aspect	T1	T2
R1 vs. R2	Source Language Comprehension	0.322	0.593**
	Command of subject matter	0.518*	0.423
	Target Language Appropriateness	0.551*	0.296
	Target Language Accuracy	0.575*	0.326
	Sum	0.520*	0.423
	Total	0.545*	

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

R1 Vs R3	Source Language Comprehension	0.808***	0.806***
	Command of subject matter	0.860***	0.628**
	Target Language Appropriateness	0.737***	0.767***
	Target Language Accuracy	0.866***	0.724***
	Sum	0.917***	0.783***
	Total	0.894***	
R2 vs R3	Source Language Comprehension	0.634**	0.834***
	Command of subject matter	0.700**	0.782***
	Target Language Appropriateness	0.578*	0.643**
	Target Language Accuracy	0.570*	0.803***
	Sum	0.679**	0.809***
	Total	0.817***	

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.001 level (2-tailed).

It is clear from Table 8 that the total correlations reflect a range between strong and very strong relations. Although the correlations between single aspects in the comparison of R1 and R2 show some moderate to strong correlations, the total correlation is strong. In addition, the correlation between R1 and R3 and between R2 and R3 are very strong, to the degree that they can be treated as identical.

However, the correlations between the participant self-assessment results and the raters' quality assessment show mostly low relationships. On the other hand, the correlations between raters' assessments and tutor's assessment show only low to moderate relationships. Table 9 below displays those results:

Table 9 *Rater's Assessment vs. Self and Tutor's Assessment*

Aspect	R1	R2	R3
Self-assessment of T1 vs. Tutor's	0.214	-0.091	0.267
Self-assessment of T2 vs. Tutor's	0.381	0.209	0.289
Rater's assessment of T1 vs Tutor's	0.341	0.332	0.464
Rater's assessment of T2 vs Tutor's	0.580*	0.227	0.434
Rater's assessment of both texts vs. Tutor's	0.515*	0.328	0.473*

*. Correlation is significant at the 0.05 level (2-tailed).

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

7. Conclusion

The results of this study lead to the conclusion that the participants' self-assessments (or their awareness of own output) are less credible than the general assessment of their tutor. Specifically, it does not correlate well with the results of the external measures of the tutor as well as of the quality assessment raters. In other words, the study has shown that the self-assessment technique has unapproved reliability and validity to be confidently used on its own as an element in assessing TC. However, it can be an effective motivational device, as suggested by some studies, reviewed in this paper, to help trainee translators develop an awareness of their abilities or level of professionalism. In this case, it must be used in a guided and moderate way, urging students to take it seriously. To conclude, its reliability is especially questionable in translation if we take into account the unique nature of the translation process as far as the notion of correctness is concerned, and also the possibility of having multiple correct translations. So, the students may not be able to properly estimate what scores they deserve because of the inconsistent evaluation criteria they have in mind, which commonly allow for subjectivity. Subsequently, self-assessment is not perceived as a dependable measure or even as a credible indicator in measuring TC[2].

Notes

[1.] This paper reports on the results of a PhD research completed at the University of Leeds (September 2014), by the present author.

[2.] In accordance with this conclusion, self-assessment as an element in profiling translation competence was discarded from the original study.

References

Arumí, M. & Esteve, O. (2006). Using Instruments Aimed at Self-regulation in the Consecutive Interpreting Classroom: Two Case Studies. *Electronic Journal of Foreign Language Teaching*. *3* (2), p.158-189. Retrieved 08.03.2013, from http://e-flt.nus.edu.sg/v3n22006/estevearumi.pdf.

Bachman, L. & Palmer, A. (1989). The Construct Validation of Self-ratings of Communicative Language Ability. *Language Testing, 6* (1), 14-29.

Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Bartlomiejczyk, M. (2007). Interpreting Quality as Perceived by Trainee Interpreters. *The Interpreter and Translator Trainer, 1* (2), 247-267.

Blanche, P. & Merino, B. (1989). Self-assessment of Foreign-language Skills: Implications for Teachers and Researchers. *Language Learning, 39* (3), 313-340.

Blanche, P. (1990). Using Standardized Achievement and Oral Proficiency Tests for Self-assessment Purposes: the DLIFC study. *Language Testing, 7* (2), 202-29.

Blue, G. (1988). Self-assessment: The Limit of Learner Independence. In: Arthur Brookes & Peter Grundy (eds), *Individualization and Autonomy in Language Learning*. ELT documents. London: Macmillan, 100-118.

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

Boud, D. (1991). *Implementing Student Self-assessment*. Second edition. Sydney: HERDSA.

Boud, D. & Fachikov, N. (1989). Self-assessment in Higher Education: A Meta-Analysis. *Review of Educational Research, 59* (4), 395-430.

Campbell, S. (1998). *Translation into the Second Language*. London and NY: Longman.

Campbell, S. (2000). Critical Structures in the Evaluation of Translations from Arabic into English as a Second Language, *The Translator, 6* (2), 211-229.

Chiaro, D. & Nocella G. (2004). Interpreters' Perception of Linguistic and Non-Linguistic Factors Affecting Quality: A Survey through the World Wide Web. *Meta,* 49 (2), 278-293.

Coronado-Aliegro, J. (2008). The Relationship between Self-Efficacy and Self-Assessment in Foreign Language Education: A Pilot Study. *Journal of Literature, Language and Linguistics,* 2 (1), 1-4. Retrieved 3-8-2013 from http://www.scientificjournals.org/journals2008/articles.

Dickinson, L. (1987). *Self-Instruction in Language Learning*. Cambridge: Cambridge University Press.

Fanghanel, J. & Voela A. (2001). Using Self-assessment to Bridge the Gap between Academic and Professional Translation. In: Martha Tennent (ed), *Training translators and interpreters in the new millennium. Portsmouth*: University of Portsmouth, 45-65.

Fernández, F. & Zabalbeascoa P. (2012). Correlating Trainees' Translating Performance with the Quality of Their Metacognitive Self-evaluation. *Perspectives: Studies in Translatology, 20* (4), 463-478.

Fowler, Y. (2007). Formative Assessment Using Peer and Self-assessment in Interpreter Training. In: Cecillia Wadensjö and Birgitta Englund Dimitrova (eds), *Critical Link 4 Professionalisation of Interpreting in the Community*. Amsterdam: John Benjamins, 253–262.

Gardner, D. (2000). Self-assessment for Autonomous Language Learners. *Links & Letters*. 7, 49-60.

Harris, M. & McCann, P. (1994). *Assessment*. London: Macmillan.

Hartley, A., Mason, I. Gracie, P. et al. (2003). Peer and Self-assessment in Conference Interpreter Training: 1-24. Retrieved 12.06.2013 fromwww.lang.ltsn.ac.uk/prf.aspx#lang1

Janssen-Van, D. (1989). The Development of a Test of Dutch as a Foreign Language: The Validity of Self-assessment by Inexperienced Subjects. *Language Testing,* 6 (1), 30-46.

Krashen, S. (1977). The Monitor Model for Adult Second Language Performance. In: Marina Burt, Heidi Dulay & Mary Finocchiaro (eds), *Viewpoints on English as a Second Language.* NY: Regents, 152-61.

Lee, Y. (2005). Self-assessment as an Autonomous Learning Tool in an Interpretation Classroom, in *Meta,* 50 (4) [no pagination]. Retrieved 15.07.2013 from http://id.erudit.org/iderudit/019869ar.

Lenski, S., Ehlers-Zavala, F., Daniel M. et al. (2006). Assessing English-language Learners in Mainstream Classrooms. *International Reading Association*, 24–34. Retrieved 12.05.2013 from onlinelibrary.wiley.com.

MacIntyre, P., Noels K. & Clement R. (1997). Biases in Self-Ratings of Second Language Proficiency: The Role of Language Anxiety. *Language Learning,* 47 (2), 265 287.

Martínez, M. N. & Albir A. H. (2001). Assessment in Translation Studies: Research Needs. *Meta*. *46* (2), 272-287.

Mills, J. & Glover C. (2007). Who Provides the Feedback- Self and Peer Assessment? *FAST,* 1-4. Retrieved 15.07.2013 from http://www.open.ac.uk/fast/pdfs/Mills%20and%20Glover.

Arab World English Journal (AWEJ)Special Issue on Translation No.5 May, 2016

Is Self-assessment a Credible Indicator of Translation Competence?   Al-Emara

O'Malley, M. & Chamot, A. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.

Orozco, M. & Albir A. H. (2002). Measuring Translation Competence Acquisition. *Meta, 47* (3), 375- 402.

Oxford American English Dictionary. (2013). S.v. Self-assessment. Retrieved 12.06.2013 from http://oxforddictionaries.com/definition/american_english/self--assessment

Oxford English Dictionary. (2013). S.v. self-assessment. Retrieved 12.06.2013 from http://oxforddictionaries.com/definition/english/self--assessment.

Postigo Pinazo, E. (2008). Self-Assessment in Teaching Interpreting. *TTR: traduction, terminologie, redaction*. 2 (1), 173-209. Retrieved 12.5.2013 from http://www.erudit.org/revue/ttr/2008/v21/n1/029690ar.html?

Robinson, B., Rodrígueza C. L., & Sáncheza M. T. (2006). Self-assessment in Translator Training. *Perspectives: Studies in Translatology,* 14 (2), 115-138.

Ross, J. (2006). The Reliability, Validity, and Utility of Self-Assessment. *Practical Assessment, Research & Evaluation,* 11 (10), 1-13. Retrieved 15.07.2013 from http://pareonline.net/pdf/v11n10.pdf.

Rust, C. (2002). *Purposes and Principles of Assessment*. Oxford: Brookes University, 1-5. Retrieved 15.07.2013 from www.brookes.ac.uk/services/ocsld/resources/briefingpapers/ppassessm.

Saltourides, E. (2006). Discovering Language Strategies through Self-assessment: Students Becoming Independent Learners. *Arizona Working Papers in SLAT*. 13, 55-76. Retrieved 18.04.2013 from http://slat.arizona.edu/sites/slat/files/page/awp13saltourides.pdf.

Schjoldager, A. (1996). Assessment of Simultaneous Interpreting. In: Cay Dollerup & Vibeke Appel (eds), *Teaching Translation and Interpreting 3: New Horizons*. Amsterdam & Philadelphia: John Benjamins, 187-96.

Swan, K. (2005). A Constructivist Model for Thinking about Learning Online. In: John Bourne & Janet Moore (eds), *Elements of Quality Online Education: Engaging Communities*. Needham: MA: Sloan-C.

Thomson, C. (1996). Self-assessment in self-directed learning: issues of learner diversity. In Richard Pemberton, Edward Li & Winnie Or (eds.), *Taking Control: Autonomy in Language Learning*. Hong Kong: Hong Kong University Press.

Waddington, C. (2001). Different Methods of Evaluating Student Translations: The Question of Validity, *Meta,* 46 (2), 311-325.