Are Your Participants Multilingual? The Role of Self-assessment in SLA Research

Amy S. Thompson

University of South Florida, USA

Abstract

This study addresses the issue of using self-assessments to evaluate the language proficiency of participants in SLA research. When conducting research in the field of SLA, it is important to analyze the linguistic background of potential participants; however, finding comparable assessment tests in multiple languages to satisfy the linguistically diverse backgrounds of the language learners as well as the difficulty in finding time to administer the tests to those with experience in multiple languages impedes the successful collection of this language background data. The results of this study support an alternate way of determining language learners' linguistic background through the use of self-assessment, specifically with the use of the Common European Framework of Reference (CEFR) can-do statements. The results of this study provide empirical support for the use of self-assessment in SLA research.

1. Introduction

Finding a way to accurately assess the language background of participants is an issue for SLA researchers. When conducting research, a thorough analysis of the language learning experiences of the participants needs to be taken into consideration in order to be able to examine the data from multiple perspectives. However, many researchers monitor the learners' linguistic background only if the learners have a high proficiency in another language (e.g., Clyne, 1997; Clyne & Cassia, 1999; Nayak et al. 1990; Williams & Hammarberg, 1998). This idea has been challenged by several second language researchers, including De Angelis (1999, 2005a, 2005b, 2007), Selinker & Baumgartner-Cohen (1995), and Thompson (2013a), who found effects of previous language experience, even when that experience was minimal. Perhaps one reason for the limited consideration of participants' previous language learning experience is the lack of an efficient yet effective way to gather such information. Finding comparable assessment tests in multiple languages to satisfy the linguistically diverse needs of the language learners as well as the difficulty in finding time to administer the tests to those with experience in multiple

languages impedes the successful collection of this linguistic background data. The results of the present study support an alternate way of evaluating language learners' linguistic background through the use of self-assessment questionnaires, specifically using the Common European Framework of Reference (CEFR) can-do statements. The results of this study provide empirical support for the use of self-assessment in SLA research.

2. Literature Review

The use of self-assessment in SLA research is much debated: "It is with cyclical regularity that the issue of self-assessment...finds its way into journals dealing with educational measurement and applied linguistics" (Ross, 1998, p. 1). Can researchers rely on participants to give accurate portrayals of their linguistic abilities? The literature on the topic is rife with controversy, and a consensus on the usefulness of self-assessment scores has yet to be reached in the field of second language research. The purpose of this literature review is twofold. In the first part, the controversies surrounding self-assessment as an accurate measure of linguistic ability are presented. In the second part, the argument for the use of self-assessment in empirical research to accurately portray participants' language learning experiences and multilingual status is explored.

2.1 Controversies with self-assessment

How accurate are self-assessment tools? The results from the literature offer conflicting conclusions. There are those studies that find self-assessment an inaccurate measure (e.g. Peirce et al., 1993; Brantmeier, 2006) whereas there are those studies that show the accuracy of self-assessment scores (e.g. Birckbichler et al., 1993; Brantmeier, 2005; Hargan, 1994; Heilenman, 1991; von Elek, 1987).

Peirce et al. (1990) found that self-assessment only correlates weakly with actual language ability in Canadian French immersion programs. Similarly, Brantmeier (2006) found that the self-assessment of reading ability by advanced learners of Spanish entering into the university as freshmen did not correlate strongly with actual abilities in reading. In contrast, Brantmeier (2005) found advanced learners of Spanish to assess quite accurately their L2 reading ability. According to Blanche & Merino (1989), although self-assessment scores can vary depending on the tools used as well as other factors, the relationship between self-assessment scores and actual ability is "good to quite good" (p. 315). In Blanche & Merino's (1989) article, the authors provide a meta-analysis of 11 articles investigating the reliability of self-assessment in reading, listening, and speaking. The participants' countries of origin and L1s differed, as well as the languages that they were studying, although the most common of the target languages was English as a second language (ESL). Across the articles in the meta-analysis, the strongest correlations between the self-assessment scores and the actual abilities of the participants were found in the receptive skill of reading, followed closely by the receptive skill of listening. Although there were also correlations found in the area of speaking, the overall results showed that students are less accurate in assessing their own speaking abilities. Thus, the students proved to be more accurate in self-assessing the two receptive skills of reading and listening over the productive skill of speaking. In the article, there was no mention of the other productive skill, writing. Other authors have also found a positive correlation between self-assessment scores,

placement tests, and ability (e.g. Birckbichler et al., 1993; Brantmeier, 2005; Hargan, 1994; Heilenman, 1991; von Elek, 1987).

Blanche & Merino (1989) suggest two reasons that the results of self-assessments can vary. One reason is that the participants may interpret the questions differently. The second reason is that if a language learner has never experienced a situation mentioned in the self-assessment questionnaire, it might be difficult for the learner to self-estimate his or her ability to do it. One common tool used for self-assessment is the “can-do” statements from the CEFR; an example from this tool can help illustrate the issue of learners assessing their abilities in certain situations. Statement 17 for level B2 of the CEFR is “I can understand in detail what is said to me in standard spoken language even in a noisy environment.” If language learners have never been in a noisy environment trying to understand the language in question, how do those language learners know if they would be able to understand the foreign language in that situation?

Although it is not the focus of the current study, it is also important to briefly discuss the role of self-assessment for purposes other than research. Self-assessment scores are now included in language learning portfolios, such as the European Language Portfolio (ELP) (Ekbatani, 2000; Little, 2005); they can also be used as part of a needs analysis in a language classroom (Brown, 1995; Gravatt et. al, 1997; Richards, 2001). While using self-assessment as a partial indication of a student’s ability or as part of a needs analysis is encouraged, using self-assessment in place of high-stakes language exams is not common practice. Students have a tendency to rate their language skills differently if a high score means being exempt from a sort of requirement, such as being able to waive a language requirement at university, or if the result could impact their final course grade. In these cases, there is still a tendency to use language exams (Little, 2005). An in-depth discussion of the use of self-assessment in classroom settings is beyond the scope of the current study, whose focus is the use of self-assessment as a data collection technique in SLA research. It is important to note, however, that the relative accuracy of self-assessment differs based on context. While self-assessment is arguably an adequate measure in an SLA to be used to create an independent variable based on the language backgrounds of the participants, self-assessments might not be an accurate measure of proficiency in high-stakes situations.

2.2 Self-assessment in second language acquisition research

Self-assessment is invaluable in assessing the language ability of participants in SLA research. When conducting research involving multilingual participants, it is often difficult, if not impossible, to give a traditional test for an independent measure of language proficiency. Thus, when dealing with a relatively homogenous group of multilingual learners, self-assessment scores are a good way to assess the language skills for the languages that cannot easily be tested by an independent measure. According to Roever (2000), the current status of self-rating scores is that they are accepted for language testing purposes for experimental data collection, whereas in high-stakes testing situations, independent measures of language proficiency are favored.

However, it is the case that in-depth self-assessment measures are not used very frequently in language acquisition research; instead, the previous language experience of the participants is often not used in the data analysis. One of the issues in SLA research is the

operationalization of multilingualism itself (Wei, 2001), but once multilingualism has been operationalized, it can be used as an independent variable in the analysis of many different phenomena. There is not a consensus in the field of SLA as to what ability level is needed in a language to make monitoring the linguistics background of the participants relevant to the study. Many researchers monitor this linguistic background only when the ability is high (e.g., Clyne, 1997; Clyne & Cassia, 1999; Nayak et al. 1990; Williams & Hammarberg, 1998). However, other researchers do not agree with this assumption, with the most recent and persuasive arguments made by De Angelis (1999, 2005a, 2005b, 2007). De Angelis (2005a) illustrated the effect of minimal non-native language exposure with participant performance on overt subject insertion in Italian. In addition, Ringbom (1987) found influence at the lexical level even with a very low level of proficiency; similarly, Thompson (2013a) found that having studied more than one foreign language positively affected language aptitude, even at low levels of proficiency. Although there is not a consensus about the proficiency level of a non-native language needed to see an effect in subsequent language acquisition, a complete linguistic history of the participants in language acquisition research should be obtained for accurate research results considering the fact that "...as little as one or two years of formal instruction in a non-native language can affect the acquisition of another non-native language to a significant level" (De Angelis, 2007, p. 6). Thus, the concept of previous language exposure is crucial in determining successful subsequent language acquisition and could potentially affect the results of SLA data in empirical research.

Although self-rating proficiency scores have a variety of uses and advantages, "...there has been surprisingly little discussion of the value of self-assessment to more expensive and logistically viable approaches to proficiency and achievement assessment, particularly in the area of second and foreign language testing" (Ross, 1998, p.1), which could also be extended to the realm of second language research. Oftentimes administering several languages tests for an independent measure of proficiency is logistically impossible. Is there an accurate, alternate way to assess language ability? Several researchers have used self-report data in order to get a detailed view of the participants' language learning experiences (e.g. Thompson, 2013a; Thompson & Lee, 2014; Dewaele, 2010; Dewaele & Stavans, 2012; Dewaele & Wei, 2012). Very succinctly stated, Dewaele (2012) indicates why self-report data is an accurate measure of language proficiency in the SLA research context:

It [self-report data] gives you a rough indication, and is good enough for what we want to do. If we look at language profiles, it would be impossible to measure true, actual proficiency in 15 languages. It just wouldn't be possible. So we have to rely on what they say, and the only argument I would say is that a number of studies have shown a decent correlation between both. If I ask you to fill out that questionnaire, you would have no reason to lie. You won't win anything by telling me you are brilliant in 12 languages if you aren't (Dewaele, 2012).

Several recent studies have used participant self-assessment to operationalize multilingualism in order to be able to use it as an independent variable. For example, in the Israeli context, Dewaele & Stavans (2012) explored the relationship between immigration, multilingualism, acculturation and personality profiles (N=193). Dewaele & Wei (2012) conducted the research with monolinguals, bilinguals, and multilinguals worldwide (N=2158) to investigate the link between multilingualism and tolerance of ambiguity. Dewaele (2010) examined the effect of

multilingualism on self-perceived communicative competence and communicative anxiety those who had studied French as an additional language (N=953). In the Korean context, Thompson & Lee (2013) investigated the relationship between anxiety, proficiency, and multilingualism (N=123). In all of the aforementioned studies, the participants were asked to self-assess their language abilities as part of the background questionnaire, making it possible to operationalize multilingualism, thereby using it as an independent variable to investigate various constructs in SLA.

3. The Study

Using correlations between the CEFR can-do statements and the results of an independent measure of English language proficiency, the Michigan State University English Language Test (MSU-ELT), the results of the current study support the argument for the use of self-assessment scores as one way to assess the language ability of multilingual participants, thus helping to fill the gap in the literature regarding empirical validity of self-assessment in SLA research. This study examines the concept of self-assessment scores as a measure of language proficiency in order to collect detailed information about the language learning experiences of research participants. The research question is as follows: Are self-assessment tools, such as the CEFR can-do statements, a viable way to assess language proficiency in the context of SLA research?

Participants in this study were students enrolled in the *Casa de Cultura Britânica* (CCB) affiliated with the English Language Program of the *Universidade Federal do Ceará* (UFC - the Federal University of Ceará, a northern province of Brazil) in Fortaleza, Brazil. Other than the English program, the university also has language programs in several other languages: *Casa de Cultura Alemã* (German), *Casa de Cultura Francesa* (French), *Casa de Cultura Hispânica* (Spanish), *Casa de Cultura Italiana* (Italian), and *Casa de Cultura Portuguesa* (Portuguese). The English language program is the largest of these programs with 2,400 students out of the 5,600 total students in the programs¹.

The 79 participants in this study were all enrolled in the *curso basico*, the first set of English courses at the CCB consisting of seven levels. All the classes that were in session during the time of data collection were visited for recruiting purposes, resulting in 79 participants who completed all portions of the study. For consistency purposes, those students in programs other than the *curso basico* were not recruited. None of the participants were heritage language learners, none had study or travel abroad experience, and they had little to no contact with native speakers of the languages studied. The L2s of the multilingual participants varied slightly, although the most common L2 was English. Other L2s included Esperanto, French, German, Italian, and Spanish. Other languages studied by these participants included English, Esperanto, French, German, Greek, Italian, Japanese, Latin, Spanish, and unknown (in the case that the

¹ The CCB was founded in 1964 and was the third of the language programs or “houses” to be formed under the department called the Casas de Cultura Estrangeira (the houses of foreign culture). Although the Casas as a whole function like a department of UFC, the courses do not count for credit towards a degree. The department called Letras (similar to a world language department in an American university) offers multiple language degree programs. The students taking classes at the Casas are doing so for personal enrichment or to supplement their university studies; thus, not all of them are university students. A number of them are working adults who take a language course in addition to their other responsibilities. Getting into a language program at the Casas is quite competitive. Potential students must take a level-specific placement test, and those with the highest scores are granted enrollment. It is a government subsidized program, so the cost is minimal (the equivalent of about \$30.00 a semester for one language).

participant indicated that other languages had been studied but failed to mention the languages). The above descriptive information is provided solely for the multilingual participants. For the bilingual participants in this study, all of their L1s were Portuguese and their L2s English.

Variables in this study are additional language experience (as measured by the CEFR can-do statements) and English proficiency. The CEFR statements were translated into Portuguese for the ease of the participants, and the MSUFLT, the English level test, was given in English. Details of the materials used are described below.

CEFR is a framework that has been established by the Council of Europe as a way to standardize the language learning achievements of language learners across Europe. The self-assessment aspect of the CEFR can-do statements is one aspect of the complete assessment tool. The CEFR has six levels of assessment: A1 and A2 (basic user), B1 and B2 (intermediate user), and C1 and C2 (proficient user). For the current study, additional language experience, or multilingualism, was measured by the Reading and Listening sections of the CEFR. The CEFR has 29 questions to assess listening and 46 questions to assess reading. All of the questions are in a yes/no format of “Yes, I can do the skill described” or “No, I can’t do the skill described.” Both the listening and the reading questions test all CEFR levels (A1, A2, B1, B2, C1, and C2). The CEFR self-assessment questionnaires were distributed during the recruitment stage of the project, along with the motivation questionnaire (see Thompson, 2013b, for motivation results). The participants were asked to fill out a CEFR for every foreign language that they had studied/were studying, although for the purpose of this study, only the CEFR scores for English are analyzed. The researcher collected the questionnaires from the participants when they returned to take the language aptitude test (see Thompson, 2013a, for aptitude results). Only the reading and listening sections of the CEFR were used in data collection because those sections corresponded with the skills tested in the MSUFLT. There were no independent measures of the productive skills, speaking and writing, given to the participants; thus, these sections of the CEFR were also not given to the participants. In Tables 1 and 2 below, sample questions from each of the levels tested are listed.

Listening		Place a check mark here if you can do this.
A1	1 I can understand when someone speaks very slowly to me and articulates carefully, with long pauses for me to assimilate meaning.	
A2	5 I can understand what is said clearly, slowly, and directly to me in simple everyday conversation; it is possible to make me understand, if the speaker can take the trouble.	
B1	11 I can follow clearly articulated speech directed at me in everyday conversation, though I sometimes have to ask for repetition of particular words and phrases.	
B2	17 I can understand in detail what is said to me in standard spoken language even in a noisy environment.	
C1	23 I can follow extended speech even when it is not clearly	

		structured and when relationships are only implied and not signaled explicitly.
C2	29	I have no difficulty understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have time to get familiar with the accent.

Table 1. Sample questions from the listening CEFR

		Reading	Place a check mark here if you can do this.
A1	1	I can understand information about people (place of residence, age, etc.) in newspapers.	
A2	9	I can identify important information in news summaries or simple newspaper articles in which numbers and names play an important role and which are clearly structured and illustrated.	
B1	17	I can understand the main points in short newspaper articles about current and familiar topics.	
B2	25	I can rapidly grasp the content of the significance of news, articles, and reports on topics connected with my interests or my job, and decide if a closer reading is worthwhile.	
C1	33	I can understand fairly long demanding texts and summarize them orally.	
C2	41	I can recognize plays on words and appreciate texts whose real meaning is not explicit (for example irony, satire).	

Table 2. Sample questions from the reading CEFR

The independent measure of English language proficiency is the MSU ELP. This test consists of two parts: listening and reading. The listening section is first with 50 questions; it takes 25 minutes to complete. There are four main parts to this listening section. The first part is a short oral presentation on which the students must take notes as if they were in a lecture class. The second part consists of three short listening passages followed by a list of questions. The third part has recordings of conversations with two or more people followed by comprehension questions. The fourth part has questions about the short oral presentation presented in part one. The students can use their notes that they took in part one to complete part four. The reading section is done after the listening sections have been completed; there are 50 questions in the reading sections with a time limit of 45 minutes. Part one, questions 1-25, is reading comprehension, which consists of passages, followed by multiple choice questions based on the content of the reading passages. Part two, questions 26-50, is a vocabulary test. There is a sentence with a blank and the participant must choose the word out of four choices that best fits into the blank. This test has high measures of validity and reliability as discussed in Cook (2001). Because the students were currently taking English language classes, it was crucial that the

English level test be administered in a relatively short period of time so that the scores would be comparable; thus, the test was administered over a period of three weeks.²

4. Results

Although some researchers would argue that independent measures of language proficiency are preferred over self-assessment scores (see literature review for a discussion of this issue), there are times when administering such tests can be difficult because of time constraints of both the researcher and the participants, as well as because of the cost associated with such test administration. For some languages, these proficiency tests are not easily available, and if they are, the results they produce might not be comparable. For this reason, the reliability of self-assessment measures becomes an important question. To answer the research question: “Are self-assessment tools, such as the CEFR can-do statements, a viable way to assess language proficiency in the context of SLA research?” Pearson correlations between the participants’ MSUFLT scores and their English CEFR scores were carried out. Not only were the overall scores compared, but the sub-skill scores of reading and listening were also compared (see Table 3 for a summary of these results).

	MSUFLT: Total	CEFR: Total	MSUFLT: Listening	MSUFLT: Reading	CEFR: Listening	CEFR: Reading
MSUFLT: Total	1	.585**	.897**	.936**	.505**	.588**
CEFR: Total		1	.485**	.569**	.918**	.976**
MSUFLT: Listening			1	.693**	.449**	.472**
MSUFLT: Reading				1	.471**	.584**
CEFR: Listening					1	.810**
CEFR: Reading						1

**correlations significant at the $p < 0.01$ level

Table 3. Correlations between the MSUFLT and CEFR scores

As can be seen from the above table, all of the correlations were significant at the $p < 0.01$ level. Not only did the overall MSUFLT scores correlate significantly with the CEFR scores, but the sub-skills also significantly correlated with the sub-skills on the opposite test, indicating that this group of participants is relatively accurate at assessing their language abilities through the self-assessment tool.

² All of the sections of the MSUFLT require reading and answering questions in English. The “listening” versus the “reading” subsections refer to the types of input the participants receive to answer these questions.

The correlation coefficient of the total scores is 0.585, for the listening scores 0.449, and for the reading scores 0.584, all of which reached significance at a 0.01 level. Such correlations would be classified as “moderate” correlations; thus, it is useful to examine the data from another perspective: effect size (R^2), or “a measure of how much of the variance in one variable is accounted for by the other variable” (Larson-Hall, 2010. p. 161). To arrive at the percentage of variance accounted for by the other variable, the correlation coefficient is squared and then multiplied by 100. Looking at the correlation coefficients for the current study, R^2 for the CEFR and MSUFLT total scores comparison is 0.342, for the listening scores comparison is 0.202, and for the reading scores comparison is 0.341, meaning that the percentages are 34.2%, 20.1%, and 34.1%, respectively. Table 4 below illustrates the R^2 values for all of the correlations.

	MSUFLT: Total	CEFR: Total	MSUFLT: Listening	MSUFLT: Reading	CEFR: Listening	CEFR: Reading
MSUFLT: Total	1	.342	.805	.876	.255	.346
CEFR: Total		1	.235	.324	.843	.953
MSUFLT: Listening			1	.480	.202	.223
MSUFLT: Reading				1	.222	.341
CEFR: Listening					1	.656
CEFR: Reading						1

Table 4. R^2 values for the MSUFLT and CEFR scores

To compare these results to an oft-studied theme in SLA, age of onset and language proficiency, an R^2 value over 0.25 would be considered a large effect size (Larson-Hall, 2010). However, while the values here do indicate a large effect size in some areas of research, these results do not indicate that self-assessment can replace traditional language proficiency tests for evaluation purposes. In order to arrive at a proportion of shared variance of 80%, a correlation coefficient of 0.9 would have to be obtained. The percentages of shared variance for this data set are moderate; thus, it can be concluded that one measure does not measure the same skills as the other. Bachman (2004) gives an example of almost an identical correlation coefficient, a 0.58 correlation of a writing test and a teacher assessment, which makes the shared variance 34%. According to Bachman (2004), with these numbers, “...we might decide that the test and the classroom teacher rankings provide complementary information, and thus decide to use both” (p. 104). However, the purpose of the current study is to evaluate the potential use of self-assessment tools in SLA research, a topic further elaborated in the discussion section.

5. Discussion

The purpose of this study is to examine the role of self-assessment as a viable tool in assessing participants' language learning experiences in the context of SLA research. As stated in the above literature review, self-assessment is an oft-discussed concept in the literature. In fact, some researchers find the idea less than satisfactory, as stated by Brown (2004): "A conventional view of language assessment might consider the notion of self- and peer assessment as an absurd reversal of politically correct power relationships. After all, how could learners who are still in the process of acquisition, especially in the early processes, be capable of rendering an accurate assessment of their own performance?" (p. 270). Brown (2004), however, goes on to say in the next sentence: "Nevertheless, a closer look at the acquisition of any skill reveals the importance, if not the necessity, of self-assessment..." (p. 270). Additionally, self-assessment as part of a language assessment portfolio has been embraced by the Council of Europe, evidenced by the fact that it is a crucial component of the European Language Portfolio (Little, 2005).

The question of whether self-assessment is an adequate measure to investigate participants' language learning backgrounds for research purposes is the crucial inquiry for the current study. The current trend in empirical SLA research is the preference of standardized proficiency tests rather than self-assessments by the language learners. However, as Ross (1998) queried, why hasn't there been more discussion of the value of self-assessment as a replacement for costly proficiency assessment? To give an example to illustrate this point, one of the standardized tests that is offered in 37 languages is the Oral Proficiency Interview (OPI) and the Writing Proficiency Test (WPT), which are both under the purview of the American Council on the Teaching of Foreign Languages (ACTFL). These would be an ideal way to assess the productive language skills of speaking and writing, but the cost is prohibitively expensive - \$65 for the WPT and \$134 for the OPI.³ Researchers cannot ask participants to pay for their own tests, and research budgets rarely allow for such expenses on the part of the researcher.

Other than the expense, there are also reasons why giving proficiency tests in several different languages would be extraordinarily difficult. When conducting empirical research, unexpected events are commonplace. Imagine a situation in which a researcher had planned to go abroad to collect data and had carefully prepared language proficiency tests for the language backgrounds of participants as found in the pilot study. Upon arrival, the researcher discovered that the potential participants had more diverse linguistic profiles than the results of the pilot study showed. If this researcher did not have an alternate way of assessing linguistic background of the participants, many months of preparation would go to waste as the study might not be able to be carried out. In a situation such as this, one of the great advantages of self-assessment tools, such as the CEFR, is that they are language neutral. No matter the linguistic background of the participants, the same self-assessment tool can be used.

In an ideal world with no limits regarding funding and time involved for both the participant and the researcher, both traditional tests of language ability as well as self-assessments could be used. However, in many instances, including that of the current study, these tests are not available. The language backgrounds of the participants were English, Esperanto, French, German, Greek, Italian, Japanese, Latin, and Spanish and finding a comparable placement test would be difficult, if not impossible (or prohibitively expensive). In such a case, the linguistic neutrality of the CEFR can-do statements is ideal. Also, the way that

³ These were the costs at the time of the writing of this paper; however, the costs are subject to change.

this particular self-assessment is structured asks questions that help learners to reflect on different situations in which they might use the language in question. Asking language learners to reflect on language needed in a variety of situations, helps the language learners to be able to give valuable information to the researcher about their linguistic abilities. Alternatively, if the CEFR is too cumbersome to be administered in multiple languages, the alternative self-assessment technique used in Thompson and Lee (2014) could be used. In this study, the 148 participants completed a self-assessment for each language studied using a Likert-scale for reading, writing, speaking, listening, and grammar. An exploratory factor analysis using the Maximum Likelihood extraction method with an oblique direct oblimin rotation was employed. The results illustrated that the five skill ratings could be reduced into one single factor, which could explain 73.63% of the total variance (eigenvalue = 3.68). This single factor solution allowed the authors to average the self-assessment of the skills across each participant, using this average as a continuous independent variable. As such, there is more than one method to employ quantitative self-assessment tools in SLA research.

In fact, SLA researchers often do ask research participants for opinions that are then treated as “fact” in the form of quantitative statistics. Motivation questionnaires are prime examples of commonly used opinion-based questionnaires in SLA research. Two of the most prolific motivation researchers, Gardner and Dörnyei, use opinion-based questionnaires to research language learner motivation. Gardner (1985) developed the Attitude/Motivation Test Battery, a motivation questionnaire that has been used in SLA research since that time. Among other questionnaires used, Dörnyei et al. (2006) used an attitudinal questionnaire with a total of 13,391 participants in a 12-year longitudinal study to measure the dynamic nature of language learning attitudes and motivation and has more recently developed questionnaire for the L2 motivational self system (L2MSS, Dörnyei, 2009). Such questionnaires are widely used and accepted in the field of SLA as valuable research tools. If learner assessments of matters such as motivation and attitudes can be accepted, it is logical to assume that the judgments of these same learners be accepted when it comes to matters of linguistic competence. As Dewaele (2012) states, the participants “have no reason to lie.” In fact, the trend in empirical studies, especially those studies using large-scale questionnaires, is for self-assessment to be used in the background questionnaire of the participants in order to give the researchers a detailed language learning profile of the participants (e.g. Thompson, 2013a; Thompson & Lee, 2014; Dewaele, 2010; Dewaele & Stavans, 2012; Dewaele & Wei, 2012). This information is crucial in operationalizing multilingualism for a study, and for providing the researcher(s) with adequate information to isolate a potential independent variable for the study. Self-reported language background data could be used for grouping students in more traditional definitions of multilingualism (i.e. previous language experience) or newer ways of conceptualizing multilingualism, such as Perceive Positive Language Interaction (PPLI, e.g. Thompson, 2013a; Thompson & Aslan, 2014; Thompson & Erdil-Moody, 2014; Thompson & Khawaja, 2015). The correlations and the R^2 values for the current study (see Tables 3 and 4 above) support the notion that self-assessments provide an adequate portrayal of participants’ linguistic abilities for research purposes.

6. Conclusion

The results of this study and the arguments presented in favor of self-assessments should not be interpreted as a call to use self-assessments as a replacement for all language tests. As discussed above, self-assessments would be especially inappropriate in high-stakes situations such as satisfying language requirements or admission to degree programs. In these high-stakes situations, language learners would have a reason to assess their abilities higher than they actually are. However, in low-stakes situations, such as in a needs analysis or participation in empirical research, the language learners would not benefit by inaccurate answers. As discussed in the literature review, many SLA researchers do not adequately evaluate the linguistic background of their participants, except in the cases of high levels of proficiency (e.g., Clyne, 1997; Clyne & Cassia, 1999; Nayak et al. 1990; Williams & Hammarberg, 1998), but there have been recent arguments illustrating why this should not be the case (De Angelis, 1999, 2005a, 2005b; Thompson, 2013a). Using self-assessment tools have been shown to be an accurate measure of linguistic ability by several researchers (e.g. Birckbichler et al., 1993; Blanche & Merino, 1989; Brantmeier, 2005; Hargan, 1994; Heilenman, 1991; von Elek, 1987), and several studies have used self-assessment tools to investigate the language learning backgrounds of their participants (e.g. Thompson, 2013a; Thompson & Lee, 2014; Dewaele, 2010; Dewaele & Stavans, 2012; Dewaele & Wei, 2012). With moderate percentages of variance as well as correlation coefficients significant at the $p < 0.01$ level, this study provides empirical support to the body of research supporting the use of self-assessment measures in SLA research.

References

- Bachman, L.F. (2004). *Statistical analyses for language assessment*. Cambridge, UK: Cambridge University Press.
- Birckbichler, D., Corl, K., & Deville, C., (1993). *The dynamics of language program testing: Implications for articulation and program revision. The dynamics of language program direction*. Boston, MA: Heinle & Heinle.
- Blanche, P. & Merino, B. (1989). Self-assessment of foreign language skills: Implications for teachers and researchers. *Language Learning*, 39, 313-40.
- Brantmeier, C. (2005). Non-linguistic variables in advanced L2 reading: learner's self-assessment and enjoyment. *Foreign Language Annals*, 38(4), 493-503.
- Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System*, 34(1), 15-35.
- Brown, J.D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston, MA: Heinle & Heinle.
- Brown, H.D (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Longman.
- Clyne, M. (1997). Some of the things trilinguals do. *The International Journal of Bilingualism*, 1(2), 95-116.

- Clyne, M., & Cassia, P. (1999). Trilingualism, immigration and relatedness of languages. *I.T.L. Review of Applied Linguistics*, 123-124, 57-74.
- Cook, G. (2001). *Investigating growth trajectories on English as a second language listening and reading comprehension test*. Unpublished doctoral dissertation. Michigan State University, East Lansing.
- De Angelis, G. (1999). Interlanguage transfer and Multiple Language Acquisition: a case study. Paper presented at TESOL 1999, New York City.
- De Angelis, G. (2005a). Interlanguage transfer of function words. *Language Learning*, 55(3), 379-414.
- De Angelis, G. (2005b). Multilingualism and non-native lexical transfer: an identification problem. *International Journal of Multilingualism*, 2(1), 1-25.
- De Angelis, G. 2007. *Third or additional language acquisition*. Bristol: Multilingual Matters.
- Dewaele, J.-M. (2010). Multilingualism and affordances: Variation in self-perceived communicative competence and communicative anxiety in French L1, L2, L3 and L4. *IRAL*, 48(2-3), 105-29.
- Dewaele, J.-M. (2012) Do multilingualism and multiculturalism shape personality profiles? AAAL Conference, March 24, in Boston, MA.
- Dewaele, J.-M., & Stavans, A. (2012). The effect of immigration, acculturation and multicompetence on personality profiles of Israeli multilinguals. *International Journal of Bilingualism*, OnlineFirst. doi: 10.1177/1367006912439941 1-19.
- Dewaele, J.-M., & Wei, L. (2012). Is multilingualism linked to a higher tolerance of ambiguity? *Bilingualism: Language & Cognition*, first view, 1-10. doi. <http://dx.doi.org/10.1017/S1366728912000570>
- Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Z., Csizér, K., & Németh, N. (2006). *Motivation, Language Attitudes and Globalization: A Hungarian perspective*. Bristol: Multilingual Matters.
- Ekbatani, G. (2000). Moving toward learner-directed assessment. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL*. (pp. 1-11). Mahwah, NJ: Erlbaum.
- Gardner, R.C., (1985). *Social psychology and the role of language learning: The role of attitude and motivation*. London, UK: Edward Arnold.

- Gravatt, B., Richards, J., & Lewis, M. (1997). *Language needs in tertiary studies*. Auckland: Occasional Paper Number 10, University of Auckland Institute of Language Teaching and Learning.
- Hargan, N. (1994). Learner autonomy by remote control. *System* 22(4), 455-503.
- Heilenman, K. (1991). Self-assessment and placement: A review of the issues. In R.V. Teschner (Ed.), *Assessing foreign language proficiency of undergraduates: AAUSC issues in language program direction* (pp. 93-114). Heinle & Heinle, Boston.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.
- Little, D. (2005). The common European framework and the European language portfolio: Involving learners and their judgments in the assessment process. *Language Testing*, 22(3), 321-336.
- Nayak, H., Hansen, N., Krueger, N., & McLaughlin, B. (1990). Language learning strategies in monolingual and multilingual adults. *Language Learning*, 40, 221 – 244.
- Peirce, B. N., Swain, M., & Hart, D. (1993). Self-assessment, French immersion, and locus of control. *Applied Linguistics*, 14(1), 25-42.
- Richards, J.C. (2001). *Curriculum development in language teaching*. Cambridge, UK: Cambridge University Press.
- Ringbom, H. (1987). *The role of the first language in foreign language learning*. Bristol: Multilingual Matters.
- Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experimental factors. *Language Testing*, 15(1), 1-20.
- Roever, C. (2000). Web-based language testing. *Language Learning Technology*, 5(2), 84-94.
- Selinker, L., & Baumgartner-Cohen, B. (1995). Multiple Language Acquisition: ‘Damn it, why can’t I keep these two languages apart?’ *Language, Culture, and Curriculum*, 8(2), 115-121.
- Thompson, A. S. (2013a). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *Modern Language Journal*. 97(3), 685–70.
- Thompson, A. S. (2013b). English as a lingua franca and its effect on the motivational profiles of English language learners. In Y. Bayyurt & S. Akcan, (Eds.). *ELF5: The Proceedings of the Fifth International Conference of English as a Lingua Franca*. Istanbul: Bogazici University, 371-381.

- Thompson, A. S. & Aslan, E. (2014). Multilingualism, Perceived Positive Language Interaction (PPLI), and Learner Beliefs: What do Turkish Students Believe? *International Journal of Multilingualism*. Online first version, 1-17.
- Thompson, A. S. & Erdil-Moody, Z. (2014). Operationalizing Multilingualism: Language Learning Motivation in Turkey. *International Journal of Bilingual Education and Bilingualism*. Online first version, 1-18.
- Thompson, A. S. & Khawaja, A. (2015). Foreign Language Anxiety in Turkey: The Role of Multilingualism. *Journal of Multilingual and Multicultural Development*. Online first version, 1-16.
- Thompson, A. S. & Lee, J. (2013). Anxiety and EFL: Does multilingualism matter? *International Journal of Bilingual Education and Bilingualism* 16(6), 730-749.
- Thompson, A. S. & Lee, J. (2014). The impact of experience abroad and language proficiency on language learning anxiety. *TESOL Quarterly*, 48(2), 252–274.
- Wei, L. (2001). *The bilingualism reader*. New York, NY: Routledge.
- Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: implications for a polyglot speaking model. *Applied Linguistics*, 19(3), 295-333.
- von Elek, T. (1987). A test of Swedish as a second language: An experiment in self-assessment. In Y. Lee, A. Fok, R. Lord, & G. Low (Eds.), *New directions in language testing*. (pp. 47-57) Oxford: Oxford University Press.