

## **Aptitude as a Predictor of Second Language Achievement: An Investigation in the Saudi Arabian Context**

**Christo Moskovsky**

School of Humanities & Social Science,  
The University of Newcastle, Australia

**Merzin Alshahrani**

King Khalid University, Saudi Arabia

**Silvia Ratcheva**

School of Humanities & Social Science  
The University of Newcastle, Australia

**Stefania Paolini**

School of Psychology,  
The University of Newcastle, Australia

### **Abstract**

A two-wave longitudinal design was used to examine the extent to which aptitude predicts second language achievement. First-year Saudi university students ( $N = 56$ ), who were involved in a seven-month intensive English language course as part of their degree, completed Aptitude and English proficiency tests at the start and the end of the treatment period. Univariate and multivariate analyses found moderate relationships between aptitude and second language achievement. Results are discussed for their implications for theory and interventions.

**Keywords:** L2 aptitude, aptitude testing, aptitude stability, L2 proficiency, MLAT

## Introduction

Regardless of their theoretical persuasion, most second language (L2) researchers acknowledge that the learning of non-primary languages in adulthood involves a complex interplay among a set of often interrelated cognitive, psychological and environmental factors. As a consequence, adult L2 acquisition is a highly variable phenomenon, with plenty of empirical and anecdotal evidence of differential success even in situations in which learners are in essentially identical conditions. Hence, much attention has been devoted to individual learner differences—like age of onset, aptitude, attitudes, cognitive and learning styles, learning strategies and motivation—with the aim to ascertain their contribution to learning outcomes (see, e.g., Ehrman et al., 2003; Skehan, 1991). Among these learner-internal variables, beyond age of onset, language aptitude and motivation “have generated the most consistent predictors of second language learning success” (Dörnyei & Skehan, 2003, p. 589).

Historically, aptitude research is associated with Carroll’s work on the development of a language aptitude test in the 1950s—the so called ‘golden era’ of aptitude research. Over the following 30 years, however, aptitude excited no interest in the teaching profession. This talent was perceived as immutable and course materials were designed on the assumption that all learners were equally talented. Similarly, many L2 researchers did not see aptitude as a fruitful area of investigation either (but note Skehan’s research). The last 30 years or so (since Parry & Stansfield, 1990) have seen a remarkable resurgence in aptitude research, reflecting an increased recognition of the importance of aptitude as an L2 factor. To quote Dörnyei and Skehan again (2003, p. 591):

[T]he concept of aptitude, long regarded as out of date, has much to offer, but needs new conceptualizations to link it to insights and findings from SLA research. It also merits an active research program.

On the agenda for future research, Dörnyei and Skehan list ten questions—some pragmatic, some theoretical. Of these, in keeping up with the traditionally pragmatic orientation of the field of L2 acquisition, our study examines the question whether aptitude is a fixed trait (i.e., not sensitive to instruction and/or learning experience), and whether it has the capacity to predict L2 learning achievement.

## Conceptualization of Aptitude

Language aptitude is typically defined as a set of relatively fixed characteristics or talents—generally insensitive to language learning experience—which enable learners to master a new language faster and easier (Carroll & Sapon, 1959; Carroll, 1981; Gardner & MacIntyre, 1992; Skehan 2002). Within this line of thought, aptitude is a stable cognitive trait specific to (second) language learning and therefore different from, and independent of, general intelligence. Likewise, aptitude is seen as different from, and independent of, non-cognitive factors like motivation, attitudes, personality, etc. (Carroll, 1962; Sasaki, 1993; Wesche et al., 1982).

Aspects of this conceptualization of aptitude have been challenged. Not everyone accepts that aptitude exists as a cognitive trait distinct from general intelligence, and there is disagreement as to what aptitude is conceptually, what types of language-related cognitive traits it involves, how much it determines L2 achievement, whether it is specific to the learning of non-primary languages or also bears some relation to first language acquisition and first language competence, and finally how much it is sensitive to environmental factors (incl. whether it could

be developed via training) (Harley & Hart, 1997; Kiss & Nikolov, 2005).

### **Aptitude Components and Testing**

The whole discourse on aptitude is linked to aptitude testing because the existence of (different levels of) aptitude can only be established via some form of aptitude testing.

The best known aptitude test, the Modern Language Aptitude Test (MLAT), was constructed by Carroll and Sapon in the late 1950s. Its development was driven by practical considerations. The purpose was to produce a selective and diagnostic tool to allow language institutions to identify learners with different levels of language learning ability, thus enabling such institutions to streamline teaching and learning, and make the process more cost-effective. By comparing participants' performance on a range of cognitive tasks the researchers identified four language-related skills which together form the trait of language aptitude (Carroll, 1981, p. 105):

Phonetic coding: the ability to identify distinct sounds, to form association between these sounds and the symbols representing them, and to retain these associations.

Grammatical sensitivity: the ability to recognise the grammatical functions of words (or other linguistic entities) in sentence structure.

Inductive language learning: the ability to infer or induce the rules governing a set of language material, given samples of language material that permit such inference.

Rote memory: the ability to learn association between sounds and meanings rapidly and efficiently, and to retain these associations.

More recently Skehan (1989) suggested that 'grammatical sensitivity' and 'inductive language learning' belong to a higher order skill he called 'language analytic ability' that has been found to be most strongly correlated with L2 success (Ehrman & Oxford, 1995; Harley & Hart, 1997; Nagata et al., 1999).

Carroll (1990, p. 14) argued that research since 1959 "has not suggested any major change in the components of foreign language aptitude that have been recognised from the start," although others contend that aptitude is a richer and more complex construct than the one originally proposed by Carroll. According to Skehan (2002), in addition to the four components above, aptitude also includes auditory segmentation, attention management, working memory, etc. It is unclear, however, whether most of these can be conceptualized as strictly language aptitude components rather than as more general cognitive processes.

Another relatively recent proposal is Robinson's (2002) hierarchical cognitive model of aptitude complexes, according to which each learner possesses a unique combination of aptitude components. Robinson's idea is to categorize learners according to aptitude profiles and match these profiles to "effective instructional options" (p. 113). Appealing as Robinson's work may be theoretically, in terms of practical applications its usefulness may be limited, because such applications would require substantial resources well beyond the capacity of most language teaching institutions.

### **Aptitude's Stability and Malleability**

Another debated issue is aptitude's innateness and stability as a cognitive trait. Carroll (1993, p. 16) describes aptitude as "relatively stable and resistant to attempts to change through education or training." Similarly, Skehan (1988, p. 86) maintains that "foreign language aptitude is relatively fixed over a long period of the individual's life span, and relatively hard to modify in any significant way" (see also Harley & Hart, 1997; Politzer & Weiss, 1969). Others have

described aptitude as “a form of developing expertise rather than as an entity fixed at birth” (Eisenstein, 1980, p. 401) and have argued that aptitude can be enhanced through experience (Grigorenko et al., 2000, p. 173; McLaughlin, 1990; Sáfár & Kormos, 2008).

The issue of first language literacy’s impact on language aptitude is related to the question of aptitude stability. Skehan’s (1986; 1989) follow-up of Wells’ (1985) Bristol study found a strong correlation between early native language literacy skills and L2 aptitude in later years. Some recent studies (e.g., Sparks et al., 1995; Sparks et al., 1998; Sparks & Ganschow, 2001; and Sparks et al., 2008) also provide compelling evidence that early first language literacy skills, such as spelling and reading skills, are highly correlated with L2 achievement in later years. On the basis of such evidence one could argue that improving native language literacy rates via early childhood education would lead to a higher L2 aptitude later on. Even so, it is unclear from these studies whether aptitude can be changed after the onset of adulthood.

### **Aptitude as a Predictor of L2 Achievement**

Carroll and Sapon’s aptitude test, which remains the most widely used aptitude test today, systematically predicts L2 achievement in various learning contexts. According to Carroll (1981, p. 96):

[T]he predictive validity coefficients for foreign language aptitude batteries in representative samples are typically in the range .40 to .60 against suitable criterion measures of success in foreign language attainment, such as final course grades, objective foreign language attainment tests, or instructors’ estimates of foreign learning ability.

Numerous studies have detected similar correlation ranges between aptitude and L2 achievement (e.g., Ehrman & Oxford, 1995; Sparks et al., 1995; Nagata et al., 1999; Harley & Hart, 2002; Erlam, 2005; Kiss et al., 2005; Sparks, 2009). The latter is, in fact, noteworthy as the study which has established the highest aptitude-L2 achievement correlation (.75) we are aware of. Others, however, report much smaller correlations (e.g. Bialystok, 1978; Keitges, 1986; Hsieh, 2004; Rysiewicz, 2008), including null correlations (e.g., Headrick, 1984; Harley & Hart, 1997).

Aptitude’s ability to predict L2 achievement (via aptitude testing) has been challenged vigorously. Some have argued that aptitude, as well as the tests measuring it, are only relevant to skills developed via the audio-lingual teaching methodology, and that aptitude cannot predict achievement in relation to skills developed through the currently favoured communicative approach (see, e.g., Cook, 2001 and the sources cited therein). It has been suggested that the types of language-related cognitive traits measured with MLAT and other aptitude tests are relevant to what Cummins (1984 and elsewhere) calls ‘cognitive/academic language proficiency’ (CALP), but not to ‘basic interpersonal communication skills’ (BICS); the latter is the language of everyday communication, while the former could be described as the language of academic discourse and the product of formal classroom instruction.

Skehan (1989 and elsewhere) disagrees suggesting that analytical language ability should be equally relevant to formal and naturalistic contexts, as well as to spoken and non-spoken language skills, because it essentially involves the ability to deal with decontextualized language. Skehan refers to Reves (1983), whose study reportedly demonstrates that in both formal and naturalistic settings aptitude is more strongly correlated with achievement than a range of other

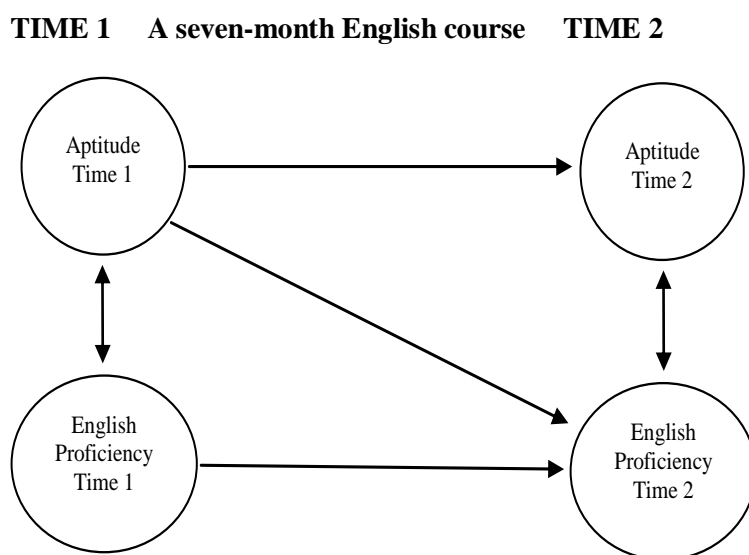
individual learner variables, incl. motivation and learning strategies. More recent studies by Harley and Hart (1997), Ehrman (1998), and Ehrman and Oxford (1995) found aptitude to be a strong predictor of L2 success even in a communicative teaching environment, but others found only a weak correlation between MLAT scores and achievement in communicative classrooms (Goodman et al., 1990; Ranta, 2002; Sáfár & Kormos, 2008). Hence, uncertainty remains— theoretically and empirically—about the extent to which aptitude has the capacity to predict L2 achievement.

This uncertainty partly reflects methodology. Most studies have used cross-sectional correlational designs (e.g. Harley & Hart, 1997; Sparks & Ganschow, 2004; Hsieh, 2004; Hummel, 2009; Gardner, 1997; etc.) which provide no firm basis for causal inferences. Also, many studies used the English version of MLAT to measure the language aptitude of speakers of languages other than English (Nagata et al., 1999; Robinson, 1997). As a result, their participants' performance would—at least in part—reflect their L2 competence in English, rather than language aptitude per se.

### Design and Objectives of the Present Study

The present research was specifically designed to avoid such methodological problems. To this end, we conducted an empirical study with a cross-sectional and two-wave longitudinal design. First year university students at King Khalid University in Saudi Arabia involved in a 7-month intensive English language course as part of their degree were administered aptitude and English proficiency tests at the beginning (Time 1, or T1) and at the end of their course (Time 2, or T2). To remove the English-proficiency confound of past assessments, we measured aptitude using a dedicated Arabic adaptation of MLAT.

We conducted a range of statistical analyses to isolate the relationships between the key constructs displayed in Figure 1.



**Figure 1. A cross-lagged design for L2 aptitude and achievement**

The combined cross-sectional and longitudinal design allowed us to pursue the following objectives:

- (1) *Testing growth in L2 proficiency overtime*: Establish to what extent our participants' English proficiency changed from T1 to T2 [the unidirectional link between T1 English proficiency and T2 English proficiency];
- (2) *Testing aptitude stability vs. malleability over time*: Establish to what extent aptitude is sensitive to L2 learning experience over a period of 7 months of intensive instruction [the unidirectional link between T1 aptitude and T2 aptitude];
- (3) *Testing aptitude's ability to predict L2 proficiency*: Establish to what extent aptitude covaries with, or is correlated to, L2 proficiency, cross-sectionally and longitudinally [the bidirectional links between aptitude and L2 proficiency at T1 and at T2, and the unidirectional link between T1 aptitude and T2 proficiency].

As part of the latter objective we also tested which aptitude components best predict L2 proficiency and which aptitude components best predict the three L2 skills included in the proficiency test.

## Method

### *Participants*

Our participants were male Saudi learners of English as a foreign language, between 18 and 20 years of age, who were first-year university students at the Languages and Translation College of King Khalid University in Abha (a city in the Southern region of Saudi Arabia). The initial sample involved 90 students: this number included the whole 2008 1<sup>st</sup> semester intake into the Languages and Translation College, although due to attrition this number eventually dropped to 56. Recruitment was conducted using a convenience sampling procedure (Mackey & Gass, 2005).

Socio-cultural and religious factors characteristic of the Saudi context did not allow us to include female learners in the present study. According to Carroll (2002, p. 26), MLAT's capacity to predict L2 achievement is essentially the same for males and females. Hence, we do not expect the gender homogeneity of our sample to have significantly affected the external validity of the study's findings.

All participants were native speakers of Arabic sharing a very similar social and cultural background. At the time of enrolment in the intensive language program, all learners had learned English in high school for six years, four 45-minute lessons per week, typically from non-native speakers of English and through grammar-based textbooks, with little focus on spoken language skills.

The language program at King Khalid University involved twenty 50-minute English classes per week. By the end of the two-semester program learners would have been learning English for over 650 hours. Overall, the participant sample can be regarded as typical of 1<sup>st</sup> year university students majoring in English in Saudi Arabia.

### *Instruments*

#### *Arabic aptitude test*

The Modern Language Aptitude Test (MLAT) was used to measure aptitude because of its highest predictive validity compared to similar aptitude instruments (Parry & Child, 1990, p.

52). We avoided the confounds of past aptitude assessments with non-English speaking learners by using an Arabic adaptation of this test, which we constructed following Stansfield and Reed's (2003) guidelines, and drawing from Rysiewicz's (2008) Polish adaptation.

The Arabic aptitude test was modelled on the short version of MLAT whose predictive power, according to the MLAT Manual (Carroll & Sapon, 2002), is comparable to the longer version. Our test included three sections:

(i) Vocabulary and written coding (VWC) (the equivalent of MLAT's Spelling Clues) – 20 items;

(ii) Grammar sensitivity (the equivalent of MLAT's Words in Sentences) – 20 items; and

(iii) Memory (the equivalent of MLAT's Paired Associates) – 24 items.

Participants were given 30 minutes to complete the test: 10 min. per section. The Aptitude test was pilot-tested on a small group of comparable EFL learners.

### *English proficiency test*

A short version of the Test of English as Foreign Language (TOEFL) was used to evaluate learners' English proficiency in Listening, Grammar and Reading. This test is regarded as the best predictor among other English proficiency tests for non-native speakers of English (Hsieh, 2004).

The TOEFL version of the test we used was sourced from a TOEFL practice book (Phillips, 2001). It consisted of 50 multiple-choice items, divided into three sections: 20 listening items (10 min.), 20 grammar items (20 min.), and 10 reading items (20 min.). Participants' L2 spoken skills were not assessed due to the limited time and attention given to the development of spoken skills in this program and in Saudi Arabian language classrooms generally.

### **Data collection**

The Arabic aptitude test and the TOEFL were administered twice in the same order—always aptitude first followed by L2 proficiency—at the start (T1) and the end (T2) of a seven-month period. Test administration was conducted in line with the respective manuals. Test answers were checked by two independent scorers. Discrepancies were resolved through accuracy check to ensure there is only one correct answer for each item.

### **Results**

#### *Analyses overview*

To address our key research objectives, the study focused on two main variables: the Arabic aptitude test scores as an independent variable consisting of three components (Vocabulary and Written Coding, Grammar Sensitivity and Memory) and acting as a measure of aptitude; and the TOEFL scores as the dependent variable, also consisting of three components (Listening, Grammar, and Reading), and acting as a measure of L2 proficiency.

To assess aptitude stability vs. malleability over seven month intensive English instruction, we performed paired sample t-tests between the aptitude scores at T1 and at T2, as well as a correlational analysis between the two sets of scores.

To assess the Arabic aptitude test's capacity to predict L2 achievement, Pearson's product-moment correlation tests were conducted between the aptitude scores and English proficiency.

Multiple regression analyses were used to quantify the capacity of each of the aptitude test's individual components to uniquely predict L2 achievement on the TOEFL subtests (i.e., while

controlling for the predictive ability of the other components). For all analyses, the significance level (or alpha) was set to .05.

**Preliminary analyses**

Because 34 out of the original 90 participants dropped out of the course, attrition analyses were conducted on the data. These analyses confirmed that the data obtained from the retained learners were not statistically different from those obtained from the non-retained learners (i.e., there was no significant selection bias in our data).

We ascertained normality of the aptitude and proficiency scores and sub-scores (all skewness and kurtosis < 1). Descriptive statistics for all variables at T1 and T2 for the 56 participants are reported in Table 1.

**Table 1. Descriptive statistics of the Arabic aptitude scores and English proficiency scores at T1 and T2 (N = 56)**

Variable s (metric)	<u>TIME 1</u>						<u>TIME 2</u>					
	Mea n	SD	Mi n	Max	Skewne ss	Kurtosi s	Mea n	SD	Mi n	Max	Skewne ss	Kurtos is
VWC (0 -20)	12.3 3	2.7 4	7	18	-.15	-.44	13.3 2	2.9 9	7	18	-.59	-.38
Gramma r Sensitivit y (0 – 20)	6.91	2.4 8	3	12	.37	-.78	7.16	2.3 6	2	13	.07	-.23
Memory (0 -24)	14.8 2	4.6 0	5	24	.02	-.63	14.0 7	5.1 4	3	23	-.28	-.45
Aptitude Total (0 -64)	34.0 1	7.1 0	20	51	.16	-.36	34.5 3	7.0 1	19	48	-.06	-.47
Listening (0 – 20)	8.69	3.0 6	2	16	-.07	-.42	9.21	2.7 1	4	15	.13	-.16
Gramma r	6.73	2.8 6	.00	12	.00	-.30	8.33	2.6 8	4	14	.73	-.01



	<u>TIME 1</u>						<u>TIME 2</u>					
(0 – 20)												
Reading	3.00	1.4	1	7	.52	.45	3.91	1.8	1	8	.44	-.58
(0 – 10)												
English Proficiency	18.4	5.4	6	29	-.09	-.14	21.5	5.7	7	37	.37	.30
(0 – 50)												

Note: Numbers in parentheses indicate the respective variable’s metric.

There was little difference between the means of the total scores for the Arabic aptitude test at T1 (M=34.01 , SD=7.10, 53.1%) and at T2 (M=34.53, SD=7.01, 53.95%), suggesting that overall L2 aptitude underwent little or no change in our group of participants over the 7-month treatment period. There was a small increase in the mean of English proficiency test between T1 (M=18.41, SD=5.43, 30.7%) and T2 (M=21.50, SD=7, 35.8%), indicating modest growth in English proficiency over the treatment period. We formally tested for the significance of these changes with inferential tests, and report these results in the sections below.

**Testing for changes in aptitude over time**

To examine the stability/malleability of aptitude over time, we carried out a series of paired sample t-tests on each of the variables and identified reliable differences over the T1-T2 period (see Table 2).

**Table 2. Mean differences in aptitude measures at T1 and T2 (N = 56)**

Variables	<u>Means</u>		T2-T1	SD	t	p
	<u>T1</u>	<u>T2</u>	diff			
VWC	12.33	13.32	0.99	2.86	2.56	.01
Grammar Sensitivity	6.91	7.16	0.25	2.39	.78	.43
Memory	14.82	14.07	-0.75	5.51	-1.01	.31
Aptitude—Total	34.01	34.53	0.42	7.04	.55	.58

Note: diff = difference; t = t-value from paired sample t-test.

There was no statistically significant difference between the means of the Arabic aptitude test total scores or the scores on the Grammar Sensitivity and Memory aptitude components at T1 and T2 ( $p > .05$ ). The only statistically significant change (in the direction of improvement) was detected in VWC,  $p = .01$  (Cohen's  $d = -.35$ ).

We used correlational analyses as an additional means of quantifying changes in aptitude between T1 and T2:

**Table 3. Correlations of aptitude measures between T1 and T2 measurements (N = 56)**

Aptitude variables	<i>r</i>	<i>p</i>
VWC T1 and VWC T2	.50	< .001
Grammar Sensitivity T1 and Grammar Sensitivity T2	.51	< .001
Memory T1 and Memory T2	.36	< .001
Aptitude T1 and Aptitude T2	.50	< .001

The total scores at the beginning and the end of the course were moderately correlated ( $r = .50$ ,  $p < .05$ ,  $N = 56$ ), adding support to the findings from the means difference test. Likewise, the scores for the three aptitude components were significantly and moderately correlated between T1 and T2, suggesting that little change occurred over the testing period. Overall, these results confirm that language aptitude is an internally stable trait.

***Testing for aptitude's capacity to predict L2 proficiency***  
***Cross-sectional correlations***

We examined the cross-sectional correlations between aptitude and proficiency separately at T1 and T2. The results are reported in Tables 4 and 5, respectively.

**Table 4. Zero-order correlations between aptitude and English proficiency components at T1 (N = 56)**

	Engl. proficiency	Listening	Grammar	Reading
Aptitude Total	.29*	.31*	.24	.02
VWC	.29*	.31*	.26*	-.13

**Aptitude as a Predictor of Second**

Moskovsky, Alshahrani, Ratcheva & Paolini

Grammar	.21	.21	.17	-.06
Sensitivity				
Memory	.16	.16	.11	.18

Note: \*  $p < .05$ .

We found a significant moderate positive correlation between the aptitude total score and the English proficiency total score at T1 ( $r = .29, p < .05, N = 56$ ). Hence, learners with a higher aptitude at T1 performed better on the English proficiency test at T1. Of the individual proficiency components, aptitude total score at T1 was significantly correlated with Listening ( $r = .31, p < .05$ ) but not with Grammar or Reading. Among the three individual aptitude components, only VWC displayed a significant correlation with the English proficiency total score ( $r = .29, p < .05, N = 56$ , Listening ( $r = .31, p < .05, N = 56$ ) and Grammar ( $r = .26, p < .05, N = 56$ ). Grammar sensitivity and Memory showed no significant correlation with either the English proficiency total score or any of its components.

Regression analyses revealed that Aptitude Total on its own predicted 8.7% ( $R^2 = .087$ ) of the variance in proficiency scores. We used multiple regression analyses to further investigate the unique relationships between the Aptitude components and the L2 Proficiency components. In a first set of analyses, we regressed each and all the scores of the English proficiency subtests at T1 separately onto each of the aptitude components at T1. The results revealed that none of the individual aptitude components uniquely predicted the scores on any of the English proficiency subsets (Listening, Grammar or Reading), all  $ps > .05$ , when controlling for the other components.

We carried out a set of parallel analyses on T2 data. These analyses aimed at assessing the cross-sectional relationships between aptitude and English proficiency, this time at the end of the 7-month training (see Table 5).

**Table 5. Zero-order correlations between aptitude and English proficiency components at T2 (N = 56)**

	Engl. proficiency	Listening	Grammar	Reading
Aptitude Total	.34**	.36**	.36**	.05
VWC	.21	.25(*)	.24(*)	-.04
Grammar Sensitivity	.26*	.24(*)	.20	.10

---

Memory	.23(*)	.24(*)	.13	.09
--------	--------	--------	-----	-----

---

Note: (\*)  $p < .07$  \*  $p < .05$ , \*\*  $p < .01$ .

We found a moderate, positive, significant correlation between the aptitude total score at T2 and the English proficiency total at T2 ( $r = .34$ ,  $p = .005$ ,  $N = 56$ ). As with the T1 findings, learners with a higher aptitude at T2 performed better on the English proficiency test at T2, confirming the Arabic aptitude test's capacity to predict proficiency.

Among the three individual aptitude components, only Grammar sensitivity showed a significant correlation with the English proficiency total score at T2 ( $r = .26$ ,  $p < .05$ ,  $N = 56$ ); VWC and Memory were not significantly correlated with the T2 English proficiency total score.

Of the three individual proficiency components at T2, the aptitude total score was significantly correlated with Listening ( $r = .36$ ,  $p < .01$ ,  $N = 56$ ) and Grammar ( $r = .26$ ,  $p < .01$ ,  $N = 56$ ).

As Table 5 shows, despite the significant correlation between the total scores for aptitude and proficiency at T2, the correlations between individual aptitude test components and individual English proficiency subtests, while close, never reached conventional levels of significance.

Overall, the correlation between T1 aptitude and T1 English proficiency totals ( $r = .29$ ) is similar in magnitude to that at T2 ( $r = .34$ ). These correlations are lower than Carroll's correlation range (.40 – .60), but are in line with the correlations reported for adapted versions of MLAT (Sasaki, 1993; Ottó, 2002; Kiss, et al., 2005; Rysiewicz, 2008). We return to this in our Discussion.

T2 Aptitude total explained 12% ( $R^2 = .12$ ) of the variance in Proficiency total. As was the case with the T1 analyses, we used multiple regression analyses to investigate the relationship between the aptitude and the proficiency components at T2. As at T1, none of the individual aptitude components uniquely predicted any of the English proficiency components (Listening, Grammar or Reading) or the L2 proficiency total, all  $ps > .05$ . However, unlike T1 results, at T2 the three aptitude components *together* significantly predicted Listening scores,  $R^2 = .14$ ,  $F(3, 52) = 2.91$ ,  $p = .04$ , and marginally predicted Proficiency total,  $R^2 = .14$ ,  $F(3, 52) = 2.76$ ,  $p = .052$ . None of the individual aptitude components, however, made a unique significant contribution to the regression models.

To summarize the results from the cross-sectional analyses, total aptitude and proficiency scores were significantly positively correlated at both T1 ( $r = .29$ ,  $p < .05$ ) and T2 ( $r = .34$ ,  $p < .01$ ). Total aptitude scores were significant, although moderate, predictors of total proficiency, explaining 8.7% (T1) and 12% (T2) of variability in overall proficiency. As far as the aptitude components, VWC was the only one that displayed significant correlations with proficiency and its components at T1. However, neither VWC nor any of the other aptitude components made a significant *unique* contribution to the prediction of total proficiency or its components at either T1 or T2.

### ***Longitudinal correlations***

We further examined the zero-order correlations between T1 aptitude scores and T2 English proficiency scores. These results are reported in Table 6.

**Table 6. Zero-order correlations between T1 aptitude and T2 English proficiency components (N = 56)**

	Prof. Total T2	Listening T2	Grammar T2	Reading T2
Aptitude Total T1	.27*	.26*	.23*	.04
VWC T1	.17	.27*	.16	-.10
Grammar Sensitivity T1	.14	.11	.16	.13
Memory T1	.24*	.18	.17	.13

Note: Prof. Total T2 = English proficiency total score at Time 2; \* $p < .05$ .

T1 Aptitude total correlated significantly although moderately (all  $ps < .05$ ) with T2 Proficiency total ( $r = .27$ ) and its Listening and Grammar components ( $r = .26$  and  $r = .23$ , respectively) while, of the aptitude T1 components, VWC correlated with Listening ( $r = .27$ ) and Memory correlated with total proficiency ( $r = .24$ ). Neither total aptitude nor any of its components were significantly correlated with Reading. Notably, the lack of correlation between T1 Grammar Sensitivity and T2 Proficiency is inconsistent with results reported in the aptitude literature (Carroll & Sapon, 1959; Skehan, 1989).

Multiple regression analyses conducted separately with the total aptitude and the three aptitude components at T1 as predictors of English proficiency and its individual components at T2 revealed that T1 aptitude total accounted for 7.6% of the variance in English proficiency total at T2 ( $R^2 = .076$ ;  $p = .04$ ). None of the individual T1 aptitude components made a significant unique contribution to T2 proficiency total or any of the T2 English proficiency sub-test scores.

### Discussion

The present study tackled issues central to the resurging debate on the nature of L2 aptitude—whether aptitude is stable or malleable over time and whether it predicts L2 proficiency.

Our study is one of few to use a two-wave design, allowing for both cross-sectional and longitudinal tests of aptitude-proficiency relationships. Also, to avoid the confound of participants' competence in English on their performance on the aptitude test, we developed and pilot-tested our own Arabic version of the aptitude test which we modelled on the most widely established and validated test, MLAT. We discuss below our key findings and put forward some ideas for future research.

*Evidence of aptitude stability*

We found evidence of aptitude stability over time. Our results showed no significant change in participants' performance on the language aptitude test over a period of seven months—i.e., between the beginning of the English course (T1) and its conclusion (T2); there was also a moderate correlation between T1 and T2 aptitude total scores. One could argue that the lack of change, rather than reflecting stable L2 aptitudes, reflects a plateau in aptitude development. Either way, the results suggest that the seven-month intensive language course (involving 20 hours per week of exposure to English instruction) was not sufficient, either in length or intensity, to produce a change in language aptitude—at least as measured with the Arabic aptitude test.

These findings go against Spark et al.'s (1998) study, but are consistent with Carroll's conceptualization of language aptitude as a language-specific cognitive trait that is stable and is generally unaffected by previous L2 learning (Skehan, 1991). Put differently, our results do not lend support to a view of aptitude as a dynamic trait amenable to training.

*Evidence of limited increase in L2 proficiency*

Our analyses reveal a statistically significant improvement in our participants' performance on the TOEFL at T2 (Cohen's  $d = -.55$ ). However, as the data in Table 1 show, in practical terms the change in English proficiency in our learners was small (averaged group performance ranging from 18.41 at T1, to 21.50 at T2, out of a scale maximum of 50): certainly much less than what one would hope to see after an intensive seven-month course. This relatively small achievement was particularly obvious in participants' listening skills, which changed little and at T2 were essentially at the same level as at T1.

Our study was not designed to investigate the reasons for the weak growth in English proficiency among our Saudi EFL learners. Low achievement in English has been reported for Saudi learners at all levels of the educational system (Zaid, 1993; Al-Seghayer, 2005; Alshahrani, 2007), and is attributed to a combination of learner-external and learner-internal factors. For example, it is acknowledged by L2 researchers and language teaching practitioners alike that language teaching/learning in the Saudi context involves limited exposure to L2 input, overpopulated classes, inferior syllabuses and substandard teaching/learning resources, inadequately qualified teachers, and unmotivated learners many of whom are only prepared to do the bare minimum to pass the course.

*Evidence of a slim aptitude–L2 proficiency link*

One issue at the forefront of aptitude research is the relationship between aptitude (as measured with aptitude tests) and L2 achievement. In practical terms, this is the question whether aptitude tests can predict L2 achievement. Our analyses reveal a positive significant correlation between our participants' performance on the Arabic aptitude test and their performance on the TOEFL. The moderate correlations we found ( $r = .27$  to  $.34$ ;  $R^2 = .07$  to  $.12$ ) are consistent with the correlations found in contexts where English is learned as a foreign language and non-English language versions of the MLAT are used (e.g., Rysiewicz, 2008:  $r = .31$ ; Sáfár & Kormos, 2008:  $r = .36$ ). Employing a non-English language aptitude test to predict English language proficiency eliminates the possibility that the scores of L2 English learners on an English MLAT are not only a function of learners' language aptitude, but their English language proficiency as well.

Another interesting finding of our research is that, while the total L2 aptitude score was significantly correlated with L2 proficiency, we did not find any significant correlations between

the individual L2 aptitude components and the total L2 proficiency. In other words, the capacity of aptitude as a whole to predict L2 achievement is different from the respective capacities of aptitude's individual component parts. This can be regarded as evidence that language aptitude is not a mechanical sum total of separate individual parts, but rather operates as an integrated cohesive trait fusing together a number of cognitive features.

### Conclusion

Our study contributes to the understanding of the nature (stable versus dynamic) and function of language aptitude in the context of Saudi university classrooms. In the process, we attempted to overcome problems previous aptitude research has had with design and the measurement of aptitude. Ours is one of few studies to examine the relationship between aptitude and proficiency both cross-sectionally and longitudinally. It is also one of a few studies that have employed a non-English language aptitude test, thus avoiding the confounds of English competence proficiency in the measurement of aptitude.

Our Arabic aptitude test is the first of such instruments developed for native speakers of Arabic. With our newly developed Arabic aptitude test, we were able to examine the construct of language aptitude from a pragmatic perspective and found that it is stable in nature and moderately effective as a predictor of L2 achievement. It is necessary to bear in mind the characteristics of the participant sample when considering these findings. The strong homogeneity of the sample may have limited our ability to detect a larger relationship. Future research will benefit from replicating the present research protocol with a sample that is more heterogeneous not only in relation to its social-demographic composition, but also in relation to participants' initial levels of English proficiency. Also, it would be worth examining the capacity of our Arabic aptitude test to predict L2 achievement with Arabic learners of languages other than English.

Recently published literature on aptitude (e.g., Robinson, 2005; Skehan, 2002) suggests that contemporary aptitude research may be shifting its focus towards the exploration of the component structure of aptitude, with a view to establishing learner aptitudinal profiles and matching them to dedicated instructional treatments. The position we take here, however, is that the idea to match language instruction to learner aptitudinal profiles—although appealing—may not come to fruition any time soon for at least two reasons: first, the number of combinations within the “constellation of individual differences” (Dörnyei & Skehan, 2003, p. 601) may be rather large and, second, batteries of tests will have to be developed in order to arrive at these profiles. In addition, the establishment of different aptitudinal profiles would not, in and by itself, automatically bring about improved learning outcomes unless dedicated curricula and language strategies/learning materials are developed to match each aptitudinal profile—a task well beyond the capacity of most language teaching institutions.

In relation to the role that individual aptitude components play with regard to L2 achievement, our study only examined the ones included in the short version of MLAT: ‘Vocabulary and Written Coding’ (our equivalent of ‘phonetic coding’), ‘grammatical sensitivity’, and ‘rote memory’. As far as these three are concerned, we found little evidence of a link between them individually and L2 achievement. At the same time we found a small, but statistically significant link between aptitude as a whole construct and L2 achievement. These findings point to the possibility that aptitude is a cohesive unified cognitive trait, not a mechanical sum of autonomously operating individual constituent parts.

Our study also showed that, similarly to other versions of MLAT, the Arabic aptitude test has the capacity to predict L2 achievement. For purely practical reasons, MLAT-like aptitude tests like the one we devised and used in this research will most likely remain a useful selection and diagnostic tool for many language teaching institutions and for some time to come.

**About the Authors:**

**Christo Moskovsky:**

Dr. Christo Moskovsky is a Senior Lecturer at the University of Newcastle, Australia, where he has been involved in postgraduate education and research in applied linguistics. He has conducted and/or supervised a wide range of projects on issues of second language acquisition. He has authored and co-authored a number of publications in this field of study, including in *TESOL Quarterly* and *Language Learning*.

**Merzin Alshahrani:**

Dr. Merzin Alshahrani is an Assistant Professor of linguistics at King Khalid University, Saudi Arabia. He currently holds the position of Dean of Academic Development and Quality at KKU. Dr. Alshahrani completed his BA in English in Saudi Arabia, and as a full scholarship student later did an MA and a PhD in Australia. Dr. Alshahrani is engaged in applied linguistics research, incl. second language acquisition, language teaching, and teaching strategies.

**Silvia Ratcheva:**

Dr. Silvia Ratcheva is a Lecturer in linguistics at the University of Newcastle. Her interests are in the social psychological aspects of second language acquisition and the use of multivariate statistics in survey-based language acquisition research. She has co-supervised a number of successful doctoral theses by international students.

**Stefania Paolini:**

Dr. Stefania Paolini is Senior Lecturer in Social Psychology at the University of Newcastle (Australia). She obtained a Laurea Degree in Social and Organizational Psychology at the University of Padova, Italy, and a PhD at Cardiff University, UK. Stefania has an international research profile. Her research interests are the social psychology of prejudice, stereotyping, and intergroup relations more broadly. Currently, her main research projects relate to the effects of intergroup contact and the learning of intergroup anxiety and stereotyping.

**References**

- Alshahrani, M. A. (2007). *Age and Second language acquisition among Saudi intermediate students*. Unpublished Masters thesis. The University of Sydney.
- Al-Seghayer, K. (Ed.). (2005). *English in the Kingdom of Saudi Arabia: Slowly but Steadily Changing*. Lawrence Erlbaum Associates.
- Bialystok, E., & Fröhlich, M. (1978). Variables of classroom achievement in second language learning. *The Modern Language Journal* 62 (7), 327-336.
- Carroll, J. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training and Research in Education*. University of Pittsburgh Press.



- Carroll, J. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual Differences and Universals in Language Learning Aptitude*. Newbury House.
- Carroll, J. (1993). *Human Cognitive Abilities: A Survey of Factor Analytic Studies*. Cambridge University Press.
- Carroll, J. (Ed.). (1990). *Cognitive Ability in Foreign Language Aptitude*. Prentice-Hall Regents.
- Carroll, J., & Sapon, S. (2002). *Modern Language Aptitude Test (MLAT): Manual*. MD, Second Language Testing, Inc.
- Cook, V. (2001). *Second Language Learning and Language Teaching*. Edward Arnold.
- Cummins, J. (1984). *Bilingualism and Special Education: Issues in Assessment and Pedagogy*. Multilingual Matters.
- Dörnyei, Z. (2005). *The Psychology of the Language Learner*. Lawrence Erlbaum Associates.
- Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition*. Blackwell.
- Ehrman, M. (1998). The modern language aptitude test for predicting learning success and advising students. *Applied Language Learning* 9 (1&2), 31-70.
- Ehrman, M., & Oxford, R. (1995). Cognition plus: correlations of language learning success. *The Modern Language Journal* 79 (1), 67-89.
- Ehrman, M., Leaver, B. L. & Oxford, R. (2003). A brief overview of individual differences in second language learning. *System* 31, 313-330.
- Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. *International Journal of Applied Psychology* 29, 159-174.
- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research* 9 (2), 147-171.
- Gardner, R. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal* 81 (3), 344-362.
- Gardner, R., & MacIntyre, P. D. (1992). Integrative motivation, induced anxiety, and language learning in a controlled environment. *Studies in Second Language Acquisition* 14, 197-214.
- Goodman, J., Freed, B., & McManus, W. (1990). Determining exemptions from foreign language requirement: Use of the modern language aptitude test. *Contemporary Educational Psychology* 15, 131-141.
- Grigorenko, E. L., Sternberg, R., & Ehram, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The canal-F theory and test. *The Modern Language Journal* 84, 390-405.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition* 19, 379-400.
- Harley, B., & Hart, D. (2002). Age, aptitude, and second language learning on a bilingual exchange. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. John Benjamins.
- Headrick, R. (1984). *Relationship of Aptitude and Attitude to Achievement in German*. Unpublished PhD dissertation, University of Missouri - Columbia, MO, USA.
- Hsieh, C. (2004). *Relationships between Non-Linguistic Variables in Second Language Acquisition and the Attained English Language Proficiency of Taiwanese University Students*. Unpublished PhD dissertation, Lynn University, Florida.

- Hummel, K. (2009). Aptitude, phonological memory, and second language proficiency in non-novice adult learners. *Applied Psycholinguistics* 30, 225-249.
- Keitges, S. (1986). *The Relationships between Foreign Language Aptitude, Attitudes and Motivations and Personality Traits and the Attained English Language Proficiency of Japanese University Students*. Unpublished PhD dissertation, University of Texas at Austin.
- Kiss, C., & Nikolov, M. (2005). Developing, piloting and validating an instrument to measure young learners aptitude. *Language Learning* 55 (1), 99-150.
- Mackey, G. S., & Gass, S. (2005). *Second Language Research: Methodology and Design*. Lawrence Erlbaum Associates.
- McLaughlin, B. (Ed.) (1990). *The Relationship between First and Second Languages: Language-Proficiency and Language Aptitude*. Cambridge University Press.
- Nagata, H., Aline, D. & Ellis, R. (1999). Modified input, language aptitude and the acquisition of word meanings. In R. Ellis (Ed.), *Learning a Second Language through Interaction*. John Benjamins.
- Ottó, I. (2002). *Magyar Egységes Nyelvérzékmérő-teszt* [Hungarian Language Aptitude Test]. Mottó-Logic Bt.
- Parry, T., & Stansfield, C. W. (1990). *Language Aptitude Reconsidered*. Prentice Hall.
- Parry, T., & Child, J. (Eds.) (1990). *Preliminary Investigation of the Relationship between VORD, MLAT and Language Proficiency*. Prentice Hall.
- Phillips, D. (2001). *Longman Complete Course for the TOEFL Test: Preparation for the Computer and Paper Tests*. Addison-Wesley Longman, Inc.
- Politzer, R., & Weiss, L. (1969). *An Experiment in Improving Achievement in Foreign Language Learning through Learning of Selected Skills Associated with Language Aptitude*. Stanford University Press.
- Ranta, L. (Ed.) (2002). *The Role of Learners Language Analytic Ability in the Communicative Classroom*. John Benjamins.
- Reves, T. (1983). *What Makes a Good Language Learner? Personal Characteristics Contributing to Successful Language Acquisition*. Unpublished PhD dissertation, Hebrew University, Jerusalem.
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning* 47, 45-99.
- Robinson, P. (2002). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. John Benjamins.
- Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics* 25, 46-73.
- Rysiewicz, J. (2008). Cognitive profiles of (un)successful fl learners: A cluster analytical study. *The Modern Language Journal* 92 (1), 87-99.
- Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL* 46, 113-136.
- Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning* 43, 313-344.
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing* 3, 188-221.

- Skehan, P. (1988). A comparison of first and foreign language learning ability: a follow-up to the Bristol Language Project. (ESOL Working Document No. 8.) London University, Institute of Education.
- Skehan, P. (1989). The relationship between native and foreign language learning ability: Educational and linguistic factors. In H. W. Dechert (Ed.), *Current Trends in European Second Language Acquisition Research*. Multilingual Matters.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisitions* 13, 275-298.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. John Benjamins.
- Sparks, R., Artzer, M., Ganschow, L., Siebenhar, D. & Plageman, D. (1998). Differences in native-language skills, foreign language aptitude, and foreign language grades among high-, average-, and low- proficiency foreign language learners: two studies. *Language Testing* 15 (2), 181-216.
- Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics* 21, 90-111.
- Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *Journal of Educational Psychology*, 100, 162–174.
- Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2009). Long-term cross-linguistic transfer of skills from L1 to L2. *Language Learning*, 59, 203–234.
- Sparks, R., Richard, L., Ganschow, L., & Patton, J. (1995). Prediction of performance in first-year foreign language courses: connection between native and foreign language learning. *The American Psychological Association*, 87, 638–655.
- Stansfield, C., & Reed, D. (2003). Adaptation of the Modern Language Aptitude Test and the Pimsleur Language Aptitude Battery for examinees whose first language is not English. Paper presented at the 2nd Annual Conference of the East Coast Organization of Language Testers, Washington, D. C. 20–21 March, 2003.
- Thompson, A. (2008). Prominent factors in the acquisition of Portuguese: Language aptitude versus previous language experience. Paper presented at the 10th Hispanic Linguistics Symposium, Somerville, MA.
- Wells, G. (1985). *Language development in the preschool years*. Cambridge: Cambridge University Press.
- Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics* 3, 127–140.
- Zaid, M. A. (1993). Comprehensive analysis of the current system of teaching English as a foreign language in the Saudi Arabian intermediate schools. Unpublished PhD, University of Colorado at Boulder.