



Received: 18 March 2019  
Accepted: 04 August 2019  
First Published: 13 August 2019

\*Corresponding author:  
Susanne M. Reiterer Department of  
Linguistics and Teacher Education  
Centre, Unit for Language Learning  
and Teaching Research, Universitat  
Wien, Austria  
E-mail: [susanne.reiterer@univie.ac.at](mailto:susanne.reiterer@univie.ac.at)

Reviewing editor:  
Juan de Dios Martínez Agudo,  
Didactics of Languages and  
Literatures, University of  
Extremadura, Badajoz, Spain

Additional information is available at  
the end of the article

## EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

# Developing and validating a second language pragmatics aptitude test

Nasrin Sedaghatgoftar<sup>1</sup>, Mohammad N. Karimi<sup>2</sup>, Esmat Babaii<sup>3</sup> and Susanne M. Reiterer<sup>3\*</sup>

**Abstract:** The present study aimed at developing a second language pragmatic aptitude test. To do so, the relevant literature was consulted, different components contributing to pragmatics aptitude were identified and tabulated and test items were developed for each component. The outcome was a test comprising four sections, i.e. memory for pragmatic rule learning, extroversion and cultural intelligence self-assessment questionnaire, mind-reading from films and mind-reading from voices. Three experts were invited to examine the face and content validity of the test. It was, then, administered to 40 native speakers of English. To establish the reliability of the test, the obtained data were subjected to Cronbach's Alpha Analysis. The results indicated that the newly developed test was a valid and reliable measure of aptitude for learning pragmatics. In order to ensure the construct validation of the test, it was administered to another 160 participants. The data gathered were analyzed using Factor Analysis. The results revealed that three sections of the test measured the same construct showing quite high correlations but extroversion and cultural intelligence self-assessment questionnaire did not. Consequently, it was removed from the test.

**Subjects:** Applied Linguistics; Psycholinguistics; Language Teaching & Learning

### ABOUT THE AUTHORS

Nasrin Sedaghatgoftar is a PhD candidate in applied linguistics, a university lecturer and an English language instructor. Her research interests include individual differences, language learning aptitude and pragmatics.

Mohammad N. Karimi holds a PhD in Applied Linguistics and currently works as an associate professor in the Department of Foreign Languages, Kharazmi University. His main areas of interest include Second Language Acquisition, Teacher Education/Development, and Cognition.

Esmat Babaii is professor of applied linguistics at Kharazmi University, Iran. Her research focus is on Systemic Functional Linguistics, Appraisal theory, test-taking processes, test validation, and critical approaches to the study of culture and language.

Susanne M. Reiterer, PhD in biological psychology, is a professor for cognitive neuroscience of second language acquisition with a special focus on language aptitude and individual differences in the psychology of language learning.

### PUBLIC INTEREST STATEMENT

Some individuals are better at doing something such as playing an instrument or learning a language, that is, they have a certain aptitude. Like other talents such as musical talent, the talent for foreign language learning is composed of various independent linguistic skills and is measured through performance on special psychological tests. It is widely believed that the current language learning aptitude tests do not measure the ability to learn the pragmatic aspects of a second or foreign language. The present study attempted to develop and validate a pragmatics subtest to the current language aptitude batteries. The test measures the ability to remember cross-cultural pragmatic rules, the ability to recognize the thoughts and feelings of people speaking a foreign language from their gestures and facial expression and from the tone of their voice.

**Keywords:** Language learning aptitude; second language pragmatics; aptitude tests; mind-reading; extroversion; cultural intelligence

### 1. Introduction

People have always been fascinated by the simple fact that some individuals are just better at doing something such as playing an instrument or learning a language. The gifted individuals seem to reach a high achievement or mastery level without putting much effort or time into acquiring the skills. It is believed that such individuals possess an innate potential for achieving high ability in a certain domain, that is, they have a certain aptitude for something (Dornyei & Skehan, 2003; Robinson, 2002a, 2002b; Turker, Reiterer, Schneider, & Seither-Preisler, 2018; Wen, Skehan, Biedron, Li, & Sparks, 2019; Wesche, Edwards, & Wells, 1982).

A number of studies (Dornyei, 2005; Granena & Long, 2013; Skehan, 2002; Wen et al., 2019) introduce aptitude as the most significant cognitive individual difference (ID). IDs, the learner qualities exerting the greatest amount of consistent influence on the SLA process, affect L2 learning. Simply put, individuals are different regarding the cognitive (and affective) resources, aptitudes and abilities they bring to the context of learning, so their learning is subject to variation (Dornyei, 2010; Reiterer, Hu, Sumathi, & Singh, 2013). Similarly, Skehan (2002) admits the centrality of aptitude in relation to L2 acquisition processes.

#### 1.1. The structure of language aptitude

Aptitude has traditionally been considered as stable and monolithic (Singleton, 2017). However, recent studies (e.g. Dornyei, 2010; Doughty, 2018; Granena, 2013; Hu et al., 2013; Jilka, 2009; Suzuki & DeKeyser, 2017; Wen, Biedron, & Skehan, 2017) have indicated that the construct is componential. Dornyei (2010) contends that “the concept covers a range of different cognitive factors making up a composite measure that can, in turn, be referred to as the learner’s overall capacity to master a foreign language” (p. 3). Similarly, Ellis (2015) reiterates that language aptitude is best seen as a complex construct comprising a number of distinct abilities. It is further suggested by Sternberg (2002) that language aptitude involves multiple aspects rather than a single fixed trait.

Carroll (1959; cited in Wen et al., 2017) made the first attempt to conceptualize the construct of language aptitude as containing multiple components. The classical model of language aptitude is a four-component model developed by Carroll and Sapon as a result of a series of factor analytic studies done on the whole range of variables operationalizing abilities thought to be important in FL learning (Dornyei, 2005; Ellis, 2015; Gonzalez, 2011; Singleton, 2017; Wen et al., 2017). Table 1 (from Ellis, 2015) summarizes the four abilities measured by the MLAT (Modern Language Aptitude Test by Carroll & Sapon, 1959; cited in Ellis, 2015).

Due to the serious criticisms directed at the MLAT, such as being outdated (Dornyei, 2005; Ellis, 2015) and not measuring the ability to learn the pragmatics of a second language, attempts were later made to introduce new conceptualizations of the construct of language

**Table 1. Carroll’s four-component model of aptitude**

Ability	Definition
Phonetic coding ability	The ability to code unfamiliar sounds in a way that they can be remembered later
Grammatical sensitivity	The ability to recognize the grammatical functions of words in sentences
Inductive language learning ability	The ability to identify patterns of correspondence and relationships between form and meaning
Rote-learning ability	The ability to form and remember associations between L1 and L2 vocabulary items

learning aptitude (e.g. Robinson's Aptitude Complex Model, 2005; Skehan's Processing Stage Model of Aptitude, 2002; Doughty's Aptitude for High-Level Language Proficiency, 2013). In what follows two of the most recent models of aptitude which have the potential to meet a number of the concerns raised against the traditional models are described: Robinson's Aptitude Complex Model and Skehan's Processing Stage Model of Aptitude.

Drawing on Snow's (1987, 1994; cited in Robinson, 2002a, 2005) interactionist Robinson (2001, 2002a, 2007, 2012) identifies a number of aptitude-complexes, or combinations of cognitive abilities. Robinson, like other recent views of aptitude, sees it as a multiple construct and suggests that learners have sets of aptitudes or what following Snow (1994; cited in Robinson, 2002a) he calls aptitude complexes for learning from classroom instruction. Robinson holds that primary abilities (e.g. pattern recognition, speed of processing in phonological working memory, grammatical sensitivity) combine to define sets of higher second order abilities (e.g. noticing the gap, memory for contingent speech, deep semantic processing, memory for contingent text, metalinguistic rule rehearsal) which are hypothesized to support language learning. The second order abilities can group up to form aptitude complexes (Snow, 1987; cited in Robinson, 2002a, 2005) representing hypothesized combinations of aptitude variables that jointly influence learning in a particular situation (Snow, 1994; cited in Robinson, 2002a). Dornyei and Skehan (2003) assert that the Aptitude Complexes Hypothesis is consistent with a cognitive view of SLA.

In another attempt to define the structure of aptitude, Skehan (2002) relates different constituent components of aptitude complexes to different stages of second language acquisition. Skehan (2002, also see Robinson, 2012) observes that the phonetic sensitivity component of MLAT as well as the measures of attentional management and working memory can measure the first stage in his model that is noticing and concerns the registration of the input. Further, Skehan proposes that the grammatical sensitivity and inductive language learning subtests of MLAT can capture some of the abilities contributing to stages two to five which generally concern pattern analysis. On the other hand, stage five named integrating (the analyzed knowledge) which involves the capacity for restructuring, stages six through eight concerning the control of analyzed knowledge and the cumulative proceduralization of knowledge in fluent performance over time, and the last processing stage (stage nine), that is lexicalizing involving going beyond rule-based processing to build a fluent lexical repertoire, are not well matched by any existing aptitude tests or subtests.

### **1.2. Current aptitude test batteries**

Several test batteries have been designed to date to measure aptitude for second language learning such as MLAT (Carroll & Sapon, 1959), PLAB (Pimsleur, 1966; cited in Dornyei, 2005), DLAB (Petersen & Al-Haik, 1976), VORD (Parry & Child, 1990), CANAL-F (Grigorenko, Sternberg, & Ehrman, 2000), LLAMA (Meara, 2005) and Hi-LAB (Linck et al., 2013) to mention a few (for more information see Skehan, 2002).

Carroll's four-factor model of FL aptitude was an empirically established model which served as the basis for the construction of the Modern Language Aptitude Test. The four components of Carroll's model of FL aptitude were operationalized in five specific performance tasks in such a way that the postulated construct reflected itself best in the nature of the abilities called for to complete it. This test-writing process of turning the theoretical construct into its concrete ability-based form resulted in the following tasks which made up the final MLAT test battery (Gonzalez, 2011; Rysiewicz, 2008; Singleton, 2017; Wen et al., 2017): **Part I.** Number learning, **Part II.** Phonetic Script, **Part III.** Spelling clues, **Part V.** Paired Associates, **Part IV.** Words in Sentences.

Following MLAT, there has been a steady flow of aptitude research, although firmly within the framework established by Carroll. First, it is important to mention that there has been large-scale work aimed at the production of aptitude test batteries other than the MLAT produced by Carroll and Sapon. However, other attempts to produce complete aptitude batteries have had a more restricted efficiency (Dornyei & Skehan, 2003; Wen et al., 2017).

### **1.3. Problems with the current foreign language aptitude batteries**

As the MLAT measure of aptitude is the measure most widely used, to date, in SLA research (Ellis, 2015; Rysiewics, 2008) and the later studies have not reconceptualized aptitude in any significant manner (Skehan, 2002), it will be scrutinized here, along with the problems associated with it.

Robinson (2013) finds two major problems with the MLAT and the other current aptitude tests, which undermines the construct validity of such tests. First, he criticizes the tests for not measuring the whole range of different abilities drawn on at different stages of language learning. There are numerous abilities contributing to speaking and interacting in an L2 none of which can be measured by the MLAT. In other words, the current aptitude tests do not measure the ability to learn pragmatics and the appropriate formal and informal phrases and expressions needed when socializing, to learn how to manage interactions with others in the L2 and how to behave appropriately in the L2. Second, the MLAT and other current aptitude tests do not measure the whole range of different abilities relied on in different settings of language learning. Simply put, aptitude tests such as the MLAT do not actually measure the full range of abilities pertaining to language learning across settings in which L2 exposure can occur. In other words, the abilities contributing to learning the language inside or outside a classroom or through explicit or implicit exposure to language are not measured by the current aptitude batteries.

Other criticisms of the MLAT include being outdated (Gonzalez, 2011; Sparks, Javorsky, & Ganschow, 2005). As Ellis (2015) states the MLAT was designed to predict which learners would succeed in language learning when taught by the pattern-drilling within the Audiolingual method of language teaching popular at the time such tests as the MLAT were designed. In other words, the current tests of language aptitude are not informed by recent research and theories in the field of SLA. There has been a great deal of SLA research since the 1950s and 1960s when aptitude batteries such as the MLAT were first researched, piloted, and then published. These early aptitude batteries were developed without the benefit of findings from recent research (Ellis, 2008; Granena, 2013; Wen, 2012).

Further, a number of studies have indicated that the MLAT is not a good indicator of oral proficiency and speaking ability although it is a good predictor of reading and writing abilities (Brecht, Davidson, & Ginsberg, 1995; Ehrman, 1998; Ellis, 1986; Gonzalez, 2011).

Later, researchers (e.g. Dornyei, 2010; Ellis, 2015; Robinson, 2012, 2013) wondered whether a new model of language aptitude and a different test battery was needed, a new battery or new sub-tests that take into account the abilities relied on in more communicative approaches to teaching and learning in more naturalistic settings. Similarly, Skehan (2002) argues that if the relationship between the existing aptitude sub-tests and the possible aptitude components is explored, it will result in a new research agenda and new aptitude sub-tests could beneficially be developed.

Moreover, the changes in the field of second language teaching over the last few decades emphasize the need to revise and modify the tests of aptitude for language learning such as the MLAT although it is found that they have predictive validity as measures of learning in predominantly audiolingual classrooms in the 1960s. Sternberg (2002) concedes that taking the new conceptualizations of the construct of aptitude into account in language teaching and assessment will help language learners to work to their potential.

There is, therefore, a clear need to update the current theories and measures of aptitude, accommodating, where necessary, the recent findings from SLA and cognitive psychology and cognitive neuroscience research (Reiterer, 2018). Researching the issues raised by these more recent theories of aptitude is to be encouraged for the light this can cast on explanations of SLA phenomena, as well as for its potential relevance to pedagogy, and the issue of matching learner aptitudes to optimal conditions of instructional exposure, Robinson (2013) suggests.

#### **1.4. Pragmatics aptitude**

Although aptitude is one of the most important individual differences in SLA, cognitive characteristics have been investigated only scarcely in L2 pragmatics research. As the mechanisms behind pragmatics learning, like any other type of learning, are governed by cognition, the effect of cognitive abilities on learning pragmatics is worth investigating (Taguchi & Roeber, 2017).

To date, there have been few studies in the area of pragmatics test development, and almost no attempts to develop a pragmatic aptitude test, compared to the large number of speech acts studies. One reason is that it is not easy to obtain a comprehensive picture of what is an appropriate measure of pragmatics and pragmatic ability. However, there is a growing interest in this area due to the important role of pragmatic competence in the development of communicative competence and thus its importance in language teaching itself (Yamashita, 2008).

The theoretical direction for the measurement of pragmatics was guided by the components introduced in Bachman's (1990) model of pragmatic competence. Bachman (1990) divides language competence into organizational competence, subdivided into grammatical and textual competence, and pragmatic competence, subdivided into illocutionary and sociolinguistic competence. Pragmatic competence includes organizational competence as well as the types of knowledge employed in the contextualized performance and interpretation of socially appropriate illocutionary acts in discourse. In describing illocutionary competence, Bachman (1990) refers to the theory of speech acts and language functions all of which seem to be important for the learners to acquire as part of their pragmatic competence tested as pragmatic competence.

Researchers working in L2 pragmatics have used a number of different types of instruments such as written discourse completion tasks (WDCT), multiple-choice discourse completion tasks (MDCT), oral discourse completion tasks (ODCT), roleplays, self-assessments and role-play self-assessments (RPSA) to test pragmatic ability (Brown, 2008).

Yamashita (2008) holds that present tests measure a general language ability which includes knowledge of grammar, morphology, semantics, syntax and phonology or skill categories such as listening, speaking, reading and writing. However, pragmalinguistics is not yet regularly included in these tests. The reason is that the theories of communicative competence and communicative language teaching have not been fully developed and rigorous empirical studies need to be undertaken.

## **2. Empirical studies**

It is evident that second language acquisition (SLA) research over the last two decades has considerably expanded our knowledge of the cognitive processes and constraints implicated in instructed SLA (Robinson, 2002a, 2012). What follows reviews some of the research undertaken in the area of language learning aptitude and pragmatics testing.

### **2.1. Studies on aptitude**

Winke (2013) investigated the plausibility of a model of language learning that included cognitive (rote memory, phonemic coding ability, grammatical sensitivity and phonological working memory), cognitively oriented (strategy use) and affective (motivation) variables as learning predictors. The study was accomplished via structural equation modeling (SEM), which has rarely been used in L2 aptitude research, to investigate L2 learning aptitude. How the factors affect each other within the model was also examined. SEM was used to evaluate the conjectured causal relations among the various variables investigated in the study. Winke (2013) then hypothesized a model of L2 learning that included both cognitive and affective traits.

Wen and Skehan (2011) propose and make a case for incorporating the working memory (WM) construct as a component of foreign language aptitude. They try to demonstrate that the concept of foreign language aptitude is still a viable and necessary concept for language learning and SLA

research. More importantly, it is suggested that the prospect of incorporating WM as a key component in foreign language aptitude is possible, feasible and promising.

Tellier and Roehr-Brackin (2013) examined whether the teaching and learning of either Esperanto or French would facilitate the development of language learning aptitude and metalinguistic awareness in 8- to 9-year-old English children. The results indicated that the development of language learning aptitude and metalinguistic awareness could be enhanced in children aged 8 to 9 years old. The authors found that in addition to progressing in the L2 they were taught, both treatment groups showed significant gains on measures of language learning aptitude and metalinguistic awareness.

Sparks, Patton, Ganschow, Humbach, and Javorsky (2006) studied the effects of NL (native language) skills including NL literacy (reading, spelling), NL oral language (receptive vocabulary, listening comprehension) and general intelligence (IQ) on oral and written FL (foreign language) proficiency and FL aptitude were examined. Among the variables, NL literacy measures were the best predictors of FL proficiency, and NL achievement and general (verbal) intelligence were strong predictors of FL aptitude.

Safar and Kormos (2008) investigated the relationship between foreign language aptitude and working memory and phonological short-term memory capacity, the role of foreign language aptitude in predicting success in the framework of focus-on-form foreign language instruction, and the stability of language aptitude and phonological short-term memory in the course of language learning. In line with previous research (e.g. Dornyei, 2005; Robinson, 2005), the results revealed that the Carrollian concept of language aptitude needs to be revisited considerably.

## **2.2. Studies on pragmatics testing**

Brown (2008, see also Roever, 2011) introduces Hudson, Detmer and Brown (1992, 1995; cited in Brown, 2008) as the first attempt by language testers to systematically develop and examine the effectiveness of tests of pragmatic ability. In their study, six types of tests were developed: written discourse completion tasks, multiple-choice discourse completion tasks, oral discourse completion tasks, self-assessments, role-play discourse tasks and role-play self-assessments. These tests were administered to English as a second language students. Then, descriptive, reliability and validity statistics for all the measures were calculated. The results indicated that all the measures except for the MDCT worked reasonably well from a psychometric standpoint.

One of the pioneering, most influential studies in the area of pragmatics is a large international project undertaken by Blum-Kulka and Olshain (1986). They studied the cross-cultural realization of requests and apologies and presented lists of the strategies. They suggested a number of coding categories for naming realization of requests and apologies. The findings of this empirical study highlighted speech acts as unique items for testing pragmatic ability.

Roever's (2006) web-based test of language (pragmatics) included three components of speech acts, implicatures and routines. It was intended to elicit test-takers' knowledge of commonly used strategies and beliefs in regard with what practices are acceptable in the target speech community, rather than their individual preferences. Messick's (1989; cited in Roever, 2006) validation framework was relied on to ensure the validity of the test.

To investigate different measures such as DCTs, multiple-choice tests, self-assessment and role-play tests, Yamashita (1996; cited in Yamashita, 2008) utilized three distinctive speech acts of requests, apologies and refusals in the test items. The findings indicated that all these measures were reasonably reliable and valid, except for the multiple-choice test. Yamashita (1996; cited in Yamashita, 2008), however, admitted that these tests can cover a very limited area as they only involved three speech acts, while around 1000 to 10000 illocutionary forces or speech acts have been estimated to be present in English language (Austin, 1962; cited in Yamashita, 2008).

Yamashita (2008) suggests that it does not seem appropriate to use only a few speech acts in testing learners' pragmatic ability and claim that the learner is at a certain level of pragmatic ability. Other speech acts such as giving orders, making promises, giving thanks and so forth need to be included in testing pragmatics.

### **2.3. Rationale for the present study**

Due to the above-mentioned shortcomings of the current aptitude tests along with the important role the identification of the aptitude profiles of the learners can play in their language learning success, updating the present language learning aptitude tests to accommodate the findings from communicative language teaching research as well as the new conceptualizations of the construct of aptitude and new cognitive concepts such as the working memory seems essential (Robinson, 2007).

The time has come to expand our notion of language aptitude and to move beyond conventional notions of testing it. Therefore, there is a clear need to update our current theories and measures of aptitude, accommodating, where necessary, the recent findings from SLA and cognitive psychology research (Robinson, 2013; Skehan, 2002; Wen, 2012; Winke, 2013).

Consequently, this study attempts to design a pragmatic aptitude subtest which has been lacking in the current aptitude tests to compensate for the above-mentioned shortcomings. It is hoped that adding a pragmatic subtest would give us a clearer and more complete image of the abilities of the learners in learning a foreign language.

As noted, the objective of the study is to design a test which can measure the aptitude for learning the pragmatics of a second language. To do so the following research question was formulated:

Q: Does the pragmatic aptitude test predict achievements in learning second language pragmatics?

## **3. Methodology**

### **3.1. Participants**

A total of 57 (33 female and 24 male) adult native speakers of English language participated in the first phase of the study for Alpha Chronbach's Analysis. It should be admitted that most of the data were gathered through sharing the link of the test on social media such as Facebook as the instrument was an online test. To ensure the data were obtained from native speakers of English, following large-scale studies such as Linck et al. (2013), the demographic part of the test included an item enquiring about the native language of the participants and their answers were trusted in, which must be admittedly regarded as a drawback. However, as the data were gathered online, it was really hard for the researchers to find a more trustable alternative way of ensuring that the participants were native speakers. Since the study deals with pragmatics and one section of the test measures aptitude for cross-cultural pragmatic rule learning, the demographic part of the test also asked the participants what other languages and cultures they were familiar with. The participants who were familiar with one of the cultures introduced and assessed in the test (i.e. Japanese, Korean, Greek, Arabic and Hebrew) were excluded. As a result, 11 female and 6 male participants were removed, leaving 40. Another 160 native speakers took the test in the second phase of the study for construct validation and Factor Analysis. The age of the participants ranged from 18 to 38.

### **3.2. Procedure**

To devise the pragmatic aptitude test, first, various theories of pragmatic ability (e.g. Bachman, 1990; Bachman & Palmer, 2010) and language aptitude including pragmatic aptitude (e.g. Robinson, 2002a, 2005) were consulted. Then the components of pragmatic ability were recognized and tabulated. This phase of the study was one of the most difficult and most complicated as the fluid and context-dependent nature of pragmatics did not lend itself to a clear-cut definition. A list of the most cited components, nevertheless, was extracted: speech acts, functions, appropriateness of utterance in

a context (from Bachman, 1990; Bachman & Palmer, 2010), nonverbal factors (Celce-Murcia, 1995), mind-reading, extroversion (Robinson, 2005) and cultural intelligence (Ang et al., 2007; Sternberg & Grigorenko, 2005). As developing a test with all of these components included would be too long, the researchers decided to choose only the most determining factors and exclude the rest of the components. As a result, speech acts, appropriateness of utterance in a context, mind-reading (from films and from voices) and extroversion and cultural intelligence were included. Then each component was operationalized through a number of test items. The items were divided to four sections: 1) memory for pragmatic rule learning consisting of 20 items to measure the ability to remember unfamiliar pragmatic rules of another language and to determine the appropriateness of speech acts in a context, 2) extroversion and cultural intelligence self-assessment questionnaire composed of 10 items to measure the degree of extroversion and 10 items to examine cultural intelligence that is the capability to function effectively in culturally diverse settings (Ang et al., 2007), 3) mind-reading from films comprising 10 items to investigate the ability to figure out what an interlocutor means through their body language when you do not know the language they speak and mind-reading from voices consisting of 10 items to measure the ability to understand what someone says through the prosodic features of language when you do not know the language they speak. The following steps were taken to construct the four sections of the test.

### 3.2.1. First section

The first section of the test named “memory for pragmatic rule learning” involves measuring the ability to remember pragmatic rules from other languages which are culturally different from English. In doing so, some pioneering studies, such as Gass and Neu (1996), Wierzbicka (1985), Trosborg (2010) and Han (1992) were consulted. Trying to single out some of the most striking differences, the researchers came up with the following list:

- Koreans are most likely to disagree with compliments. Even if the compliments are accepted, it would be in the form of a downgrade (Han, 1992).
- Greek compliments are often seen as information seekers. So the complimentee would provide information about how to obtain the object of the compliment (Jaworski, 1995).
- Arabs believe complimenting on good things may bring evil to the complimentee, so they answer like “Praise God instead” (Nelson, El Bakary & Al Batal., 1996).
- Arabs would offer the complimented item in response to a compliment (Nelson et al., 1996).
- Unlike American English speakers who tend to follow an apology with an explanation, Hebrew speakers are likely to give an explanation only (Murphy & Neu, 1996).
- Japanese tend to use apologies in requests while Americans would use thanks (Lee, Park, Imai, & Dolan, 2012).

Then some test items were developed according to each cross-cultural difference comprising 20 items in total. In the exposure phase of this section of the test, the testees were exposed to the above-mentioned pragmatic differences for 60 seconds. They were supposed to read and remember the sentences containing information on the cross-cultural verbal behavior of people with different nationalities and cultures regarding three speech acts of request, compliment-response and apology. Subsequently, in the test phase, the questions were displayed, each on a separate page with a time limit of 40 seconds. Every page presented some short scenarios, each followed by a brief incomplete dialog. For each item the participants needed to check the answer that best completed the dialog considering the information given in the exposure phase. For each item, all the choices would be correct but given the culture referred to in each scenario, one answer was the most probable.

It is worth noting that although the cross-cultural differences are derived from languages and cultures other than English, the test items are presented through the medium of English since the participants are native speakers of English none of whom know any of the languages or cultures the study aimed at in this section.

The counterbalance between sociopragmatics and pragmalinguistic was achieved through a meticulous description of the setting, participants, purpose, and content in each of the constructed scenarios in this study. The role relationship between the interlocutors in each scenario was clearly described in some detail as professor, student, friend and classmate.

Care was also exercised to limit the content of the prompts to the situations that were familiar to the test takers. The findings indicate that the procedure that was used for the construction of the scenarios in this study was highly effective.

### 3.3. Sample items

1. A is a 22 year-old Japanese female. B is a 22 year-old Japanese female. They are classmates. A missed the previous session of the class and wants to ask B for her notes. They are in the cafeteria.

*What would A most probably say?*

- a. Can I borrow your notes for a few days?
- b. Sorry, can I borrow your notes for a few days?
- c. Would you please give me your notes for a few days?

2. A is a 45-year-old Greek female. B is a 44 year-old Greek female. The conversation took place in a tea party.

A: "Nice brooch!"

B: " ... .. "

*What would B most probably say?*

- a. Do you like it? Have it.
- b. Oh, you think so? Thanks.
- c. Is it? I bought it at the Macy's.

#### 3.3.1. Second section

The second section of the test measures the degree of extroversion and cultural intelligence through a self-assessment questionnaire. To develop the items of this section, "cultural intelligence assessment questionnaire" developed by Ang et al. (2007) and "Five Factor Inventory" designed by Goldberg (1992) were relied on and adapted. Ten items were selected from each and some minor modifications were applied. This section is composed of twenty items asking the participants to assess themselves on a five-point Likert Scale (Strongly agree = 5 to Strongly disagree = 1).

### 3.4. Sample items

1. I know the cultural values and religious beliefs of other cultures.  
1. strongly disagree 2. disagree 3. neutral 4. agree 5. strongly agree
2. I don't talk a lot.  
1. strongly disagree 2. disagree 3. neutral 4. agree 5. strongly agree

### 3.4.1. Third section

Various studies in the area of mind-reading (e.g. Baron-Cohen & Cross, 1992; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Golan, Baron-Cohen, Hill, & Golan, 2006; Golan, Baron-Cohen, Hill, & Rutherford, 2006) were the models for the third and fourth sections of the test. Golan, Baron-Cohen, Hill and Golan, (2006) focused on the recognition of emotions and mental states in others. They concede that emotion recognition strongly depends on the ability to integrate multimodal information in context. They developed a new “ecological” task for assessing recognition of complex emotions and mental states, using social scenes from films.

However, these studies were all examining autistic people. The third section was modelled after mind-reading from films. This section was entitled “mind-reading from films” following the studies (e.g. Baron-Cohen & Cross, 1992; Baron-Cohen et al., 2001; Golan, Baron-Cohen, Hill, & Golan, 2006; Golan, Baron-Cohen, Hill, & Rutherford, 2006) it was modelled on. The only difference was that the film clips used in this study were muted unlike the film clips utilized in research dealing with the autistic. The clips were muted so the participants had to figure out what the actors or actresses meant only through their body language, facial expressions and gestures; thus, representing the real life situation of communicating with people speaking an unfamiliar language.

First a collection of movies was created and carefully watched by the researchers. Then, ten fragments from seven of the movies containing ten scenes considered as proper for the purposes of this study were selected. The films were chosen due to their dramatic value and occurrence of emotion-bearing scenes. The target scenes (11 to 20 seconds long) were extracted and muted. In each sampled scene two to five characters interacted revealing their emotions or mental states through their facial expressions. In each scene, one of the characters was selected and their emotional or mental state at the end of the scene was labeled. The scenes were, then, played to three native speakers to ensure their perceptions and interpretations of the scenes matched those of the researchers. In order to build the four choices of each test item, Golan, Baron-Cohen, Hill & Golan, (2006) was consulted. They relied on an emotion taxonomy by Baron-Cohen, Golan, Wheelwright and Hill (2004; cited in Golan, Baron-Cohen, Hill, & Golan, 2006) to match the verbal difficulty of the correct answer and the distractors.

### 3.5. Sample items



1. At the end of the scene, how is the man feeling?

- a. pleased                      b. smug                      c. wanting                      d. kind



2. At the end of the scene, how is the woman feeling?

- a. serious                      b. sad                      c. relaxed                      d. reflective

### 3.5.1. Fourth section

As mentioned above, this section of the test was modelled after “mind-reading from voices” studies for the autistic (e.g. Golan, Baron-Cohen, Hill, & Rutherford, 2006) and was entitled “mind-reading from voices” accordingly. The spectrum of emotions used in Golan, Baron-Cohen, Hill, & Rutherford, (2006) was considered as the basis. Ten of the emotions were selected, the proper utterances to convey the feelings were decided upon, verbalized and recorded. The recordings were unisex (feminine). Three native speakers of Persian, in addition to the three Persian-speaking researchers of the study, were invited to listen to the recordings and confirm if the intended emotions were conveyed through the recorded utterances. Persian (the native language in Iran) was chosen as the language of the recorded voices as the researchers believed it was highly improbable that the participants knew Persian. Nevertheless, to make sure that the testees were not familiar with Persian, the demographic part of the test requested the participants to specify the languages they were familiar with, if any, in addition to English as their mother tongue. The purpose in doing so was to identify the participants who were familiar with Persian (or the other languages presented in the first section) and to exclude them from the study. Further, as Persian was the mother tongue of three of the researchers, it was more convenient to use Persian than find another language most probably unknown by the participants.

### 3.6. Sample items

The transcript of the Persian recording:/Che jaleb! Man kheili khosham oomad/

The English equivalent (not included in the test): How interesting! I really liked it.

1. How is she feeling?

- a. concerned                      b. interested                      c. hopeful                      d. scared

The transcript of the Persian recording:/Goftam ke, Na. Nemisheh/

The English equivalent (not included in the test): I said no. No way.

2. How is she feeling?

- a. friendly                      b. stern                      c. resentful                      d. excited

After the different sections of the test had been constructed, a platform for making online tests named “Flexiquiz” was used to present the test online. This platform, compared to other platforms, had a number of qualities that made it uniquely appropriate for the purposes of this study. First of

all, it was possible to include text as well as videos and audios in the test. Second, it was possible to allocate a separate page to each and every question and set time limits to every page. Third, every time a new participant took the test, the researchers received an e-mail notifying the new test results and the researchers could easily check the results on the platform.

Subsequently, data were gathered from 40 native speakers of English. To establish the reliability of the test, the obtained data were subjected to Alpha Cronbach's analysis. Further, two rounds of Factor Analyses were administered for construct validation.

#### 4. Results

This study aimed at developing a test that could measure the aptitude for learning the pragmatics of a second language. SPSS 22 was used to analyze the data.

When the online test was constructed, it was distributed on social media like Facebook. As mentioned above, whenever a new participant did the test, the researchers received an e-mail notifying that a new test result has been received. All in all, 57 people took the test. Seventeen test results were excluded because either English was not the first language of the participant or s/he was familiar with one of the languages or cultures used in the test.

As the reliability of the scale is not only important in and of itself but it is crucial to the establishment of construct validity; first, the reliability of the test was established through Alpha Chronbach's analyses (For a summary of Alpha Chronbach's Analyses see Table 2).

As mentioned before, in the first section of the test, the participants were exposed to a number of unfamiliar cross-cultural pragmatic differences and then were required to choose from among three choices the best response regarding each situation and the culture in question. The data collected from this section were subjected to Cronbach's alpha. The results showed that the items of the first section of the test are 84% internally consistent.

In the second section the participants were supposed to assess themselves through choosing the statements which described them regarding the two qualities of extroversion and cultural intelligence. Since the ten items of the questionnaire measure extroversion and the other ten assess cultural intelligence which are two separate qualities, an independent Cronbach's Alpha for each quality and an overall Cronbach's Alpha for all the items altogether were conducted. As can be seen in Table 2, the degree of reliability is 83% which is quite high. The results of the Cronbach's Alpha for the ten items measuring cultural intelligence was 78%.

The third series of calculations for the second section of the test revealed that all the items, that is, items on extroversion and items on cultural intelligence altogether, were 85% reliable.

The third section constituted of ten muted film clips sampled from feature movies. The testees had to watch the short scenes and choose from among four choices the one that described the emotional or mental state of the character at the end of the scene. The Cronbach's Alpha turned out to be 80% for this section.

**Table 2. Summary of reliability statistics for all the four sections of the test**

Test section	Cronbach's Alpha	N of Items
Memory for pragmatic rule learning	.895	20
Extroversion-cultural intelligence self-assessment questionnaire	.857	20
Mind-reading from films and voices	.895	

The fourth section, like the third section, measures mind-reading ability but through voices. The participants were required to listen to ten recorded utterances in Persian and then determine the emotional or mental state of the speaker. The results of this section were subjected to Cronbach's alpha and produced a high reliability estimate of 81%. The results can be seen in Table 2.

Since sections three and four are subscales of mind-reading, an overall alpha Cronbach was administered for the two sections altogether. The result turned out to be 89%.

Table 2 summarizes the reliability statistics for all the four sections of the test.

Having ensured the reliability of the test, the researchers set out to establish the construct validation of the test. Since a far greater sample was required for Factor validation, data were gathered from 160 other native speakers of English, in addition to the 40 participating in the first phase, making up a pool of 200 participants. The results from this phase of the study were subjected to Factor Analysis. Due to the nature of the data which were partly dichotomous and partly ordinal, first, the data were computed and reduced to four variables, i.e. memory for pragmatic rule learning, extroversion and cultural intelligence self-assessment questionnaire, mind-reading from films and mind-reading from voices. It is worth noting that computing the data was necessary to unify the nature of the data into interval data so SPSS could do the Factor Analysis correctly. Tables 3, 4 and 5 and Figure 1 display the results from the first round of the Factor Analyses.

As can be seen in the Correlation Matrix, three of the variables, namely, memory for pragmatic rule learning, mind-reading from films and mind-reading from voices, are highly correlated with one another ( $r > .7$ ) except for the last one, which is extroversion and cultural intelligence self-assessment questionnaire, showing a very low correlation with the other three variables ( $r < .11$ ).

The result for Kaiser-Meyer-Olkin test turned out to be .72 which was well above .6 indicating that the sample was quite suited for Factor Analysis. Further, the result for Bartlett's test was

**Table 3. Correlation matrix**

Correlation		Memory for pragmatic rule learning	Mind-reading from films	Mind-reading from voices	Self-assessment questionnaire
	Memory for pragmatic rule learning	1.000	.779	.918	.113
	Mind-reading from films	.779	1.000	.827	.005
	Mind-reading from voices	.918	.827	1.000	.104
	Self-assessment questionnaire	.113	.005	.104	1.000

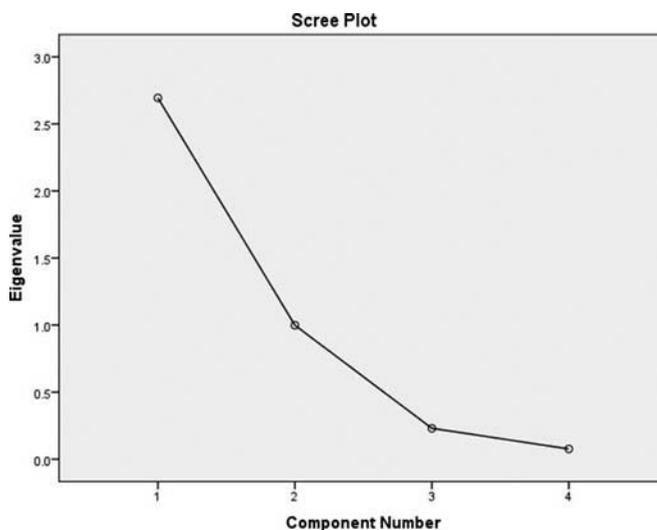
**Table 4. KMO and Bartlett's test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.720
Bartlett's Test of Sphericity	Approx. Chi-Square	600.217
	df	6
	Sig.	.000

**Table 5. Component matrix**

	Component
	1
Mind-reading from voices	.969
Memory for pragmatic rule learning	.953
Mind-reading from films	.912
Self-assessment questionnaire	

**Figure 1. Scree plot showing the number of components.**



significant ( $.000 < .005$ ) suggesting that the variables did not have equal variances and were suitable for structure detection.

According to Table 5, all the variables displayed loadings on one component (assumed as language pragmatics aptitude) except for extroversion and cultural intelligence self-assessment questionnaire which did not. The Scree Plot in Figure 1 confirms this result.

Since extroversion and cultural intelligence self-assessment questionnaire showed neither a significant correlation with the other three variables nor a loading on the extracted component, this variable was removed and a second round of Factor Analysis was administered. The results are displayed through Tables 6, 7 and 8 and Figure 2.

**Table 6. Correlation matrix**

		Memory for pragmatic rule learning	Mind-reading from films	Mind-reading from voices
Correlation	Memory for pragmatic rule learning	1.000	.779	.918
	Mind-reading from Films	.779	1.000	.827
	Mind-reading from voices	.918	.827	1.000

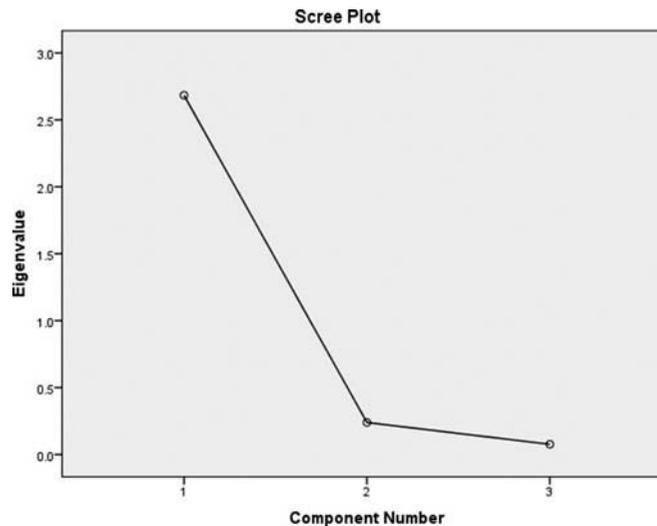
**Table 7. KMO and Bartlett's Test**

<b>Kaiser-Meyer-Olkin Measure of Sampling Adequacy</b>		<b>.724</b>
Bartlett's Test of Sphericity	Approx. Chi-Square	594.275
	df	3
	Sig.	.000

**Table 8. Component matrix**

	<b>Component</b>
	1
Mind-reading from voices	.969
Memory for pragmatic rule learning	.953
Mind-reading from films	.912

**Figure 2. Scree plot showing the number of components.**



Correlations were computed among the three variables. The results suggest that the variables are highly correlated ( $r > .7$ ). The high internal consistency estimates indicate that the items function together to consistently measure the pragmatics aptitude construct.

KMO and Bartlett's Test results show that the data were suitable for Factor Analysis.

The component Matrix (Table 8) and the Scree Plot (Figure 2) confirm that all the three variables load on one component which should most probably be regarded as language pragmatics aptitude suggesting that the three sections of the test measure the same construct.

## 5. Discussion

A number of studies have criticized the current language aptitude batteries for lagging behind the state of the art in the field of SLA and falling short of measuring aptitude for learning the communicative aspects of language such as pragmatics and have called for a new reconceptualization of aptitude and aptitude measurement taking pragmatics into consideration (Ellis, 2015; Robinson, 2013; Skehan, 2002; Sparks et al., 2005; Wen, 2017).

Consequently, this study aimed at making the first steps into developing and validating a pragmatic aptitude test. To accomplish the goals of this research, an online test comprising four sections, i.e. memory for pragmatic rule learning, extroversion and cultural intelligence self-assessment questionnaire, mind-reading from films and mind-reading from voices, was developed. To ensure the validity of the test, three experts were invited to examine if the test can really measure what it is supposed to measure that is, the cognitive ability to learn the pragmatics of a second language. All three of the experts contended that the face and content of the test is valid and it can be used to tap into pragmatics aptitude. The next step was to make sure the newly developed test is reliable. In doing so, 40 native speakers of English took the test. The data obtained from different parts of the test were subjected to Alpha Cronbach's analysis. All sections of the test indicated quite high estimates of reliability. To establish the construct validation of the test, a series of Factor Analyses were conducted. As a result of these analyses, the extroversion and cultural intelligence self-assessment questionnaire was removed from the test since it showed neither a significant correlation with the other sections of the test nor any loadings on the only component extracted assumed as pragmatics aptitude.

Comparison to other studies is rather difficult, because pragmatic aptitude tests are so far not available and the ones existing are usually employed for special populations only such as the autistic (e.g. Baron-Cohen & Cross, 1992; Baron-Cohen et al., 2001; Golan, Baron-Cohen, Hill, & Golan, 2006; Golan, Baron-Cohen, Hill, & Rutherford, 2006).

A number of studies have recently attempted to develop new, theoretically updated aptitude test batteries (e.g. CANAL-F by Grigorenko et al., 2000; LLAMA by Meara, 2005; Hi-LAB by Linck et al., 2013). However, aptitude for learning pragmatics does not seem to have been dealt with in any significant way in these batteries. The most significant contribution to aptitude conceptualization and testing in the last few years has been the development of the Hi-LAB (Linck et al., 2013; Wen et al., 2017). Table 9 summaries the components measured by this language aptitude test (from Wen et al., 2017).

Although Hi-LAB has provided the first empirical evidence of the potential cognitive predictors of successful learning to advanced proficiency levels, i.e. plausible candidate components of the construct of high-level language aptitude, it does not engage in measuring pragmatics aptitude in any significant way given the constructs presented in Table 9. However, access to this battery is usually not provided; therefore, it cannot be really scrutinized and criticized (Ameringer, Green, Leisser, & Turker, 2018).

Another recent attempt to develop a new language aptitude battery (i.e. LLAMA) was made by Meara (2005). The LLAMA aptitude test is loosely based on the MLAT (Meara, 2005). The LLAMA comprises four sub-tests, that is, LLAMA B, a test of vocabulary learning, LLAMA D, a test of sound

**Table 9. Constructs and measures in Hi-LAB (from Wen et al., 2017)**

constructs	
Working Memory	Running Memory Span
Executive Functioning	Antisaccade
Updating	Task Switches Numbers
Inhibitory Control	Letter Span
Task Switching	Non-word Span
Phonological Short-term Memory	
Associative Memory	Paired Associates
Long-term Memory Retrieval (Priming)	ALTM Synonym
Implicit Learning	Serial Reaction Time
Processing Speed	Serial Reaction Time
Auditory Perceptual Acuity	Phonemic Discrimination: Hindi, English
	Phonemic Categorization: Russian

recognition, LLAMA E, a test of sound-symbol associations, and LLAMA F, a test of grammatical inferencing (Ameringer et al., 2018; Granena, 2013; Meara, 2005). A series of exploratory factor analyses revealed that the test measures two different aptitude dimensions interpreted as analytic ability and sound sequence learning ability (Granena, 2013). It is evident that pragmatics aptitude is not dealt with in the LLAMA language aptitude battery.

Due to the fact that almost none of the present language aptitude tests involves measuring pragmatics aptitude and attempts to develop more recent and more comprehensive language aptitude batteries failed to yield significant results, it was not an easy task to compare the present study with previous work in the field. Nevertheless, the reliability and validity results indicate that the instrument developed through the course of this study can probably be regarded as a reliable and valid measure of the cognitive ability to learn the pragmatics of a second language. Simply put, the test, as predicted, appears to assess the construct of language pragmatics aptitude indicating that the theoretical latent variable is a reasonable explanation for the performance variance among language learners. The findings suggest that the pragmatics aptitude test shows appropriate psychometric properties. Despite its novelty, this new measure is characterized by high internal consistency and a unitary structure. However, it might need to be further explored and developed.

## 6. Conclusion

Investigating the effects of cognitive aptitudes on the acquisition of language allows for uncovering the nature of learning by making inferences about a mental process that is facilitated or hindered by different aptitude components (DeKeyser, 2012).

Developments in language pedagogy and SLA research indicate that cognitive processes are fundamentally important to L2 success, and that variation between learners in the cognitive abilities that such processes draw on can in part explain variation between them in L2 success and these differences are measurable. However, the currently used aptitude batteries do not measure the ability to learn pragmatics (Robinson, 2002a, 2012, 2013). Additionally, Ellis (2008) concedes that the early aptitude batteries were developed without the benefit of findings from recent research in the field of SLA. There is, therefore, a need to revise and modify them.

The focus of this research was to develop and validate an instrument that would measure the cognitive ability to learn the pragmatics of a second language. A handful of studies (Ellis, 2015; Robinson, 2013; Skehan, 2002; Sternberg, 2002; Taguchi & Roever, 2017; Wen, 2012; Winke, 2013) have called for an update of the present language aptitude tests to include the new concepts and theories in the area of second language acquisition, especially pragmatics, that were not at the time the most prevalent language learning aptitude batteries such as the MLAT were developed. The results indicate that the instrument developed in this study has met the criteria for reliability and validity. Therefore, it can be used to tap into pragmatics aptitude but needs to be further developed and administered to larger samples.

### Acknowledgements

We would like to thank Professor Peter Robinson, Professor Karen Bryan, Dr. Shuai Li and Luke Green for their kind help.

### Funding

This work was supported by the Ministry of Science Research and Technology [no number travel fellowship].

### Author details

Nasrin Sedaghatgoftar<sup>1</sup>

E-mail: [n.s.goftar@gmail.com](mailto:n.s.goftar@gmail.com)

Mohammad N. Karimi<sup>2</sup>

E-mail: [n.s.goftar@gmail.com](mailto:n.s.goftar@gmail.com)

Esmat Babaii<sup>3</sup>

E-mail: [babai@khu.ac.ir](mailto:babai@khu.ac.ir)

ORCID ID: <http://orcid.org/0000-0001-9998-8247>

Susanne M. Reiterer<sup>3</sup>

E-mail: [susanne.reiterer@univie.ac.at](mailto:susanne.reiterer@univie.ac.at)

<sup>1</sup> Student of Applied Linguistics, Kharazmi University, Teheran, Iran.

<sup>2</sup> Applied Linguistics, Kharazmi University, Tehran, Iran.

<sup>3</sup> Linguistics, University of Vienna, Vienna, Austria.

### Citation information

Cite this article as: Developing and validating a second language pragmatics aptitude test, Nasrin Sedaghatgoftar, Mohammad N. Karimi, Esmat Babaii & Susanne M. Reiterer, *Cogent Education* (2019), 6: 1654650.

### References

- Ameringer, V., Green, L., Leisser, D., & Turker, S. (2018). In S. M. Reiterer (Ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (pp. 1–15). Heidelberg: Springer-Nature.
- Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural

- intelligence: Its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, 3(3), 335–371. doi:10.1111/j.1740-8784.2007.00082.x
- Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Baron-Cohen, S., & Cross, P. (1992). Reading the eyes: Evidence for the role of perception in the development of a theory of mind. *Mind and Language*, 7(1&2), 172–186. doi:10.1111/j.1468-0017.1992.tb00203.x
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in Films” Task: Complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251.
- Blum-Kulka, S., & Olshtain, E. (1986). Requests and apologies: A cross-cultural study of Speech Act Realization Patterns (CCSARP). *Applied Linguistics*, 5(3), 196–213. doi:10.1093/applin/5.3.196
- Brecht, R. D., Davidson, D. E., & Ginsberg, R. B. (1995). Predictors of foreign language gain during study abroad. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37–66). Philadelphia: J. Benjamins.
- Brown, J. D. (2008). Raters, functions, item types and the dependability of L2 pragmatics tests. In E. Alcon-Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 201–223). Britain: Cromwell Press Ltd.
- Celce-Murcia, M. (1995). The elaboration of sociolinguistic competence: Implications for teacher education. In J. E. Alatis, C. A. Strahle, & M. Ronkin (Eds.), *Linguistics and the education of language teachers: Ethnolinguistic, psycholinguistic, and sociolinguistic aspects* (pp. 699–710). Washington D.C.: Georgetown University Press.
- DeKeyser, R. M. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, 62, 189–200. doi:10.1111/lang.2012.62.issue-s2
- Dornyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 442–471). Oxford, UK: Blackwell Publishing Ltd.
- Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Dornyei, Z. (2010). The relationship between language aptitude and language learning motivation: individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum Companion to Second Language Acquisition* (pp. 247–267). London: Continuum.
- Doughty, C. J. (2013). Optimizing post-critical-period language learning. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 153–175). Philadelphia: John Benjamins.
- Doughty, C. J. (2018). Cognitive language aptitude. *Language Learning*, 68, 1–26.
- Ehrman, M. E. (1998). The modern language aptitude test for predicting learning success and advising students. *Applied Language Learning*, 9, 31–70.
- Ellis, R. (1986). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2015). *Understanding second language acquisition*. UK: Oxford University Press.
- Gass, S. M., & Neu, J. (Eds.). (1996). *Speech Acts across Cultures: Challenges to Communication in a Second Language*. New York: Mouton de Gruyter.
- Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The “Reading the Mind in Films” Task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience*, 1(2), 111–123. doi:10.1080/17470910600980986
- Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2006). The ‘Reading the Mind in the Voice’ Test-Revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-006-0252-5
- Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, 4(1), 26. doi:10.1037/1040-3590.4.1.26
- Gonzalez, D. R. (2011). *The effects of aptitude and motivation on the acquisition of EFL in young learners* (Unpublished MA Thesis). University of Barcelona, Barcelona.
- Granena, G. (2013). Reexamining the robustness of aptitude in second language acquisition. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 179–204). Philadelphia: John Benjamins.
- Granena, G., & Long, M. (Eds.). (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Philadelphia: John Benjamins.
- Grigorenko, E. L., Sternberg, R. J., & Ehrman, 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(3), 390. doi:10.1111/0026-7902.00076
- Han, C. (1992). A comparative study of compliment responses: Korean females in Korean interactions and in English interactions. *Working Papers in Educational Linguistics*, 8(2), 17–31.
- Hu, X., Ackermann, H., Martin, J., Erb, M., Winkler, S., & Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain and Language*, 127(3), 366–376. doi:10.1016/j.bandl.2012.11.006
- Jaworski, A. (1995). This is not an empty compliment! Polish compliments and the expression of solidarity. *International Journal of Applied Linguistics*, 5(1), 63–94.
- Jilka, M. (2009). Talent and proficiency in language. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 1–16). Berlin, Germany: Mouton de Gruyter.
- Lee, H. E., Park, H. S., Imai, T., & Dolan, D. (2012). Cultural Differences between Japan and the United States in Uses of “Apology” and “Thank You” in favor asking messages. *Journal of Language and Social Psychology*, 31(3), 263–289. doi:10.1177/0261927X12446595
- Linck, J., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., ... Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, 63(3), 530–566. doi:10.1111/lang.2013.63.issue-3
- Meara, P. (2005). *LLAMA language aptitude tests: The Manual*. Swansea: Lognostics.
- Murphy, B., & Neu, J. (1996). My grade’s too low: The speech act set of complaining. In S. M. Gass & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 191–216). New York: Mouton de Gruyter.

- Nelson, G. L., El Bakary, W., & Al Batal, M. (1996). Egyptian and American compliments: Focus on second language Learners. In S. M. Gass & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 109–128). New York: Mouton de Gruyter.
- Parry, T. S., & Child, J. R. (1990). Preliminary Investigation of the Relationship between VORD, MLAT and Language Proficiency. In T. S. Parry & C. W. Stansfield (Eds.), *Language Aptitude Reconsidered. Language in Education: Theory and Practice* (pp. 30–67). Washington, DC: Eric Clearinghouse on Languages and Linguistics.
- Petersen, C. R., & Al-Haik, A. R. (1976). The development of the defense language aptitude battery (DLAB). *Educational and Psychological Measurement*, 36, 369–380.
- Reiterer, S. M. (Ed.). (2018). *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience*. Heidelberg: Springer-Nature.
- Reiterer, S. M., Hu, X., Sumathi, T. A., & Singh, N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology (cognitive Science)*, 782(4), 1–13.
- Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes, and learning conditions in SLA. *Second Language Research*, 17, 268–392. doi:10.1177/026765830101700405
- Robinson, P. (2002a). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113–133). Amsterdam: Benjamins.
- Robinson, P. (Ed.). (2002b). *Individual differences and instructed language learning*. Amsterdam: Benjamins.
- Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 46–73. doi:10.1017/S0267190505000036
- Robinson, P. (2007). Aptitudes, abilities, contexts, and practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 256–286). New York: Cambridge University Press.
- Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 57–75). Springer-Verlag: Berlin Heidelberg.
- Robinson, P. (2013). Aptitude in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 4998–5003). Oxford: Blackwell Publishing Ltd.
- Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, 23(2), 229–256. doi:10.1191/0265532206lt329oa
- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, XX(X), 1–19.
- Rysiewics, J. (2008). Measuring foreign language learning aptitude. Polish adaptation of the modern language aptitude test by Carroll and Sapon. *Poznań Studies in Contemporary Linguistic*, 44(4), 569–595.
- Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics*, 46, 113–136. doi:10.1515/IRAL.2008.005
- Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching*, 7(1), 89–103. doi:10.14746/ssllt.2017.7.1.5
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). Amsterdam/Philadelphia: John Benjamins.
- Sparks, L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, 56(1), 129–160. doi:10.1007/s11881-006-0006-2
- Sparks, R., Javorsky, J., & Ganschow, L. (2005). Should the modern language aptitude test be used to determine course substitutions for and waivers of the foreign language requirement? *Foreign Language Annals*, 38(2), 201–210. doi:10.1111/flan.2005.38.issue-2
- Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language aptitude testing. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 13–44). Amsterdam: Benjamins.
- Sternberg, R. J., & Grigorenko, E. L. (2005). Cultural explorations of the nature of intelligence. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 226–241). Washington: American Psychological Association.
- Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, 67(4), 747–790. doi:10.1111/lang.12241
- Taguchi, N., & Roever, C. (2017). *Second Language Pragmatics*. Oxford: Oxford University Press.
- Tellier, A., & Roehr-Brackin, K. (2013). The development of language learning aptitude and metalinguistic awareness in primary-school children: A classroom study. *Essex Research Reports in Linguistic*, 62(1), 1–28.
- Trosborg, A. (Ed.). (2010). *Pragmatics across languages and cultures*. New York: Mouton de Gruyter.
- Turker, S., Reiterer, S. M., Schneider, P., & Seither-Preisler, A. (2018). The neuroanatomical correlates of foreign language aptitude. In S. M. Reiterer (Ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (pp. 119–148). Heidelberg: Springer-Nature.
- Wen, Z. (2012). Foreign Language Aptitude. *ELT Journal*, 66(2), 233–235. doi:10.1093/elt/ccr068
- Wen, Z. (2017). Testing foreign language aptitude: Insights from a working memory perspective, Assessing World Languages, Macau Polytechnic Institute, Nov. 8–11, Macau: University of Macau.
- Wen, Z., Biedron, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1–31. doi:10.1017/S0261444816000276
- Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude: Building and supporting a case for “Working Memory as Language Aptitude”. *Ilha Do Desterro: A Journal of English Language, Literatures and Cultural Studies*, 60, 15–44.
- Wen, Z., Skehan, P., Biedron, A., Li, S., & Sparks, R. (Eds.). (2019). *Language aptitude: Advancing theory, testing, research and practice*. UK: Routledge.
- Wesche, M. B., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics*, 3, 127–140. doi:10.1017/S0142716400006664

Wierzbicka, A. (1985). Different cultures, different languages, different speech acts: Polish vs. English. *Journal of Pragmatics*, 9, 145–178. doi:10.1016/0378-2166(85)90023-2

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning.

*The Modern Language Journal*, 97(1), 109–130. doi:10.1111/j.1540-4781.2013.01428.x

Yamashita, S. (2008). Investigating interlanguage pragmatic ability: What are we testing? In E. Alcon-Soler & A. Martinez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 201–223). Britain: Cromwell Press Ltd.



© 2019 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format.

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



**Cogent Education (ISSN: 2331-186X) is published by Cogent OA, part of Taylor & Francis Group.**

**Publishing with Cogent OA ensures:**

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

**Submit your manuscript to a Cogent OA journal at [www.CogentOA.com](http://www.CogentOA.com)**

