

Language Aptitude and Language Experience: Are They Related?¹

Mark Sawyer
University of Hawai'i at Manoa

Introduction

If there is one generalization that can be safely made about second language acquisition (SLA) among adults, it is that there is striking variability both in rate of learning and in ultimate attainment. Accounting for this variability, on the other hand, remains an area of much controversy. Of the many affective, cognitive, and social factors which have been offered as causes of individual differences in SLA, one that has been given relatively little attention in recent years is that of *language aptitude*. Besides the difficulty of defining it and specifying its relationship to SLA, one major reason why research on aptitude has become somewhat dormant is that language aptitude tests have come to be associated with the misuses and misinterpretations of intelligence testing. This association is largely based on the commonly held assumption that aptitude is fundamentally a fixed, innate capacity. Research on bilingualism and multilingualism, however, has produced interesting results suggesting that aptitude is not an immutable capacity at all, but rather is somehow enhanced during the process of learning additional languages. If language learning capacity can thus be increased through experience, the pedagogical implications are likely to be far-reaching.

The first half of this paper will review the major literature on aptitude, and then proceed to examine relevant studies dealing with the language learning capabilities of bilinguals and multilinguals. It will be shown that there is growing evidence consistent with the hypothesis that aptitude may increase naturally through language learning experience. The second half of the paper will report on a study designed to explore aspects of this hypothesis directly, by correlating scores on the Modern Language Aptitude Test (Carroll and Sapon, 1959) with various quantitative aspects of learners' previous language learning experience and subsequent performance in foreign language classes.

What is language aptitude?

J.B. Carroll, the psychologist and psychometrician whose name is most closely associated with the concept of language aptitude, has made various attempts at defining it. For him, the key aspect of aptitude is rate of learning--hence the general definition of aptitude as the "amount of time needed to master a task under optimal conditions" (1963), and the specification of language learning aptitude as "some characteristic of an individual which controls, at a given point of time, the rate of progress that he will make subsequently in learning a foreign language" (1974). Carroll believes that aptitude is innate, but in later writings has revealed some ambivalence:

Aptitude as a concept corresponds to the notion that in approaching a particular learning task or program, the individual may be thought of as possessing some current state of capability of learning that task--if the individual is motivated, and has the opportunity of doing so. That capability

¹ I would like to thank Mike Long for his suggestions on the original research design, and J. D. Brown for his assistance in every subsequent stage of the project. The study suffers to the extent that I have not yet been able to incorporate all their ideas.

is presumed to depend on some combination of more or less enduring characteristics of the individual. (1981, p. 81)

His use of "current state" and "more or less enduring characteristics" may be indications of Carroll's realization that the research is not conclusive on these issues.

Concerning the content of language aptitude, Carroll (1965) posited that it is not a unitary capacity but rather a complex of several abilities: (1) phonetic coding ability; (2) grammatical sensitivity; (3) rote-learning ability; and (4) inductive language learning ability. He offered an array of factor, multiple regression, and canonical regression analyses in support of the construct validity of these components.

Although various researchers on cognitive factors in SLA have differed with Carroll on the precise nature and effects of aptitude, they have tended to concede that something like language aptitude does exist. The notable exception is Neufeld (1978, 1979), who argued that "everyone, despite their age, is innately equal in their language learning ability" (1978, p. 17). Neufeld did not take seriously what some other researchers consider the "indisputable evidence linking performance on aptitude tests to classroom achievement in a new language" (Wesche, 1981, 119), because: (1) this evidence has come from "semi-empirical" observations rather than being grounded in studies of child language; and (2) the data is likely to be better explained by non-cognitive variables, such as "social pressures and norms, individual psychological traits, cultural biases, and student-teacher relationships. (Neufeld, 1979, p. 232)

At the other pole is Skehan (1986), whose ideas are grounded in studies of child language. In a follow-up study of subjects participating in the Bristol Language Project (Wells, 1981, 1985), Skehan (1986) found that language aptitude could be predicted by two aspects of variation in first language development: (1) syntactic complexity measures; and (2) "early measures which tap influences in calculating an ability to use language in a decontextualized way" (p. 199). Thus, Skehan, who basically agrees with Carroll's position on the ways in which aptitude predicts foreign language achievement, goes one step further and makes claims about what variations in child language development can predict later foreign language aptitude.

The following three sections will address specific areas of controversy concerning how aptitude is to be defined.

Measures of language aptitude

Ellis (1986) states that aptitude "is usually defined in terms of the tests that have been used to measure it" (p. 112). This situation is quite understandable, given the fact that while cognitive science is still in its infancy, aptitude tests with high predictive validity have been around for almost thirty years. By far the most widely used measure of language aptitude is the Modern Language Aptitude Test, or MLAT (Carroll and Sapon 1959). It was painstakingly developed, extensively corroborated in terms of predictive validity, and has been subject to continuing research (for review see McInnis, 1985). Other well known aptitude tests are the Pimsleur Language Aptitude Battery, or PLAB (Pimsleur, 1966), which was developed primarily for adolescents, and the Defense Language Aptitude Battery, or DLAB (Peterson and Al-Haik, 1976).

Carroll (1981), in his explanation of how parts of the MLAT correspond to the various components of aptitude, admits that the fourth component, "inductive language learning ability," is not specifically measured by any of his subtests, but is well represented on Subtest 4 of PLAB, *Linguistic Analysis*.

Aptitude and intelligence

Oller and Perkins (1978) claimed that most of the reliable variance in a wide variety of language proficiency measures could be attributed to a single global language proficiency factor, which they argued is identical to intelligence. Carroll (1981) did not deny that there is overlap between the components of aptitude and the components of intelligence, but pointed out that they can not be the same thing because they do not share the same patterns of correlation with foreign language achievement. For example, Gardner and Lambert (1965), in a factor analysis of 24 variables relating to the L2 achievement of 96 high school students, found that measures of intelligence were relatively independent of both aptitude and L2 achievement. Genesee (1976), studying children in Grades 4, 7, and 11, found that I.Q. was related to reading and language usage scores, but not to listening comprehension or interpersonal communication. In a large-scale study of 793 English-speaking Canadian public servants studying intensive French, Wesche, Edwards, and Wells (1982) concluded that the MLAT was far from independent of intelligence measures, but that its subtests did measure learning abilities that were distinct from general mental abilities.

Although it is clear from the above studies that the extent of overlap between intelligence and aptitude is likely to continue to generate disagreements, there is little doubt that the two concepts are by no means identical.

Aptitude, attitude, and acquisition setting

Based on his Monitor Model, Krashen (1981) claimed that aptitude can only predict language learning achievement in formal (classroom) situations, and as measured by tests permitting Monitor use. Attitude, conversely, will be more predictive in informal settings when the focus is on communication, and when Monitor-free measures of proficiency are used. Gardner (1978, 1986) has made similar statements. One question that arises here is how to determine if there is sufficient time for Monitor use. On the MLAT, for example, all sections are speeded, and two of the sections require very quick response to aural cues, which are not repeated. Without a clearer operationalization of Monitor use than has ever been offered, it would be unjustified to claim that aptitude tests allow Monitor use in some sense that informal situations do not.

Bialystok and Fröhlich (1978) also provided some empirical contradiction to Krashen's claims. In a study of the relative effects of cognitive and affective variables on the French language achievement of 157 high school students on both formal and functional tasks, they found, contrary to their expectations, that aptitude predicted achievement on both types of tasks, whereas attitude did not predict achievement at all.

Skehan (1986) argued that the decontextualization ability that he found important in classroom language learning should take on a greater role in informal settings, where learners have to abstract from experience without the organized clues that a formal situation provides. In support of this position, Skehan also offered the findings of Reves (1983), who studied the role of several predictors of L2 success in both formal and informal settings. She found that, while prediction is not as effective in an informal setting, aptitude still accounts for more variance than any of the measures of motivation, cognitive style, or learning strategies.

Another position sometimes taken is that part or much of the variance indicated by aptitude tests can be explained by attitudes and/or motivation. McInnis (1985) and Neufeld (1979), *inter alia*, have represented this view. Empirical evidence against this position has come from Gardner, Lalonde, and

Moorcroft (1985), who found the attitudinal/motivational characteristics of 170 college students to be virtually independent of language aptitude.

Thus far, various points of controversy concerning the nature of language aptitude have been presented. The following sections will examine evidence bearing at least indirectly on the question of whether or not aptitude is an innate, fixed quality.

Language abilities of child bilinguals

If bilingual children are found to process language differently from monolingual children, the implication is that language learning abilities can be affected by environmental factors. Ben-Zeev (1977), with a sample of 96 elementary school children, found that the bilingual children in her study, despite their lower L1 vocabulary level than the monolingual controls, showed more advanced processing of verbal material, more discriminating perceptual distinctions, more propensity to search for structure in perceptual situations, and more capacity to reorganize their perceptions in response to feedback. Interestingly, they showed no advantage over monolinguals in reorganizing visual (nonverbal) material.

Bialystok (1987), studying the effects of bilingualism on the metalinguistic awareness of 5, 7, and 9 year-old children, concluded that control of linguistic processes (one of two components of metalinguistic awareness) was enhanced by bilingualism. She also suggested that biliteracy may be a catalyst or precondition for the facilitating effects of bilingualism to be demonstrated.

Effects of bilingualism and previous language training on aptitude

Carroll (1959) reported that there is very little direct evidence concerning the effect of previous language training on MLAT scores, but he went on to report a study of a Foreign Service group of adults in an intensive training course whose number of years of previous foreign language training correlated moderately with MLAT (.44), and with instructors' ratings of "estimated ability" (.55). He dismissed the results as easily being accountable to self-selection, but nevertheless was finally forced to conclude that previous L2 experience probably does contribute to predictions of success in that "it may give the student a better idea of how to go about learning a new language." (p. 21)

Whereas Bialystok's (1987) research reported in the previous section contributed to specifying the advantage enjoyed by bilingual children with regard to metalinguistic knowledge, Masny and d'Anglejan (1985) showed that metalinguistic awareness is a reliable indicator of developing L2 competence and is significantly related to language aptitude. In their investigation of the statistical relationship between selected cognitive and linguistic variables and second language grammaticality judgments of 74 Francophone college ESL students, they found that L2 proficiency, L2 classroom achievement, and language aptitude were all significant predictors of their subjects' ability to make grammaticality judgments. Thus, the studies of Bialystok and of Masny and d'Anglejan, taken together, provided an indication of the possible route by which experience may lead to L2 success.

Consistent with these findings were those of Eisenstein (1980), who studied the relationship between the language aptitude of 93 college students and their past language learning experiences. She found that both bilingualism and previous (generally high school) language training led to greater language aptitude as measured by MLAT (short form). Polylinguals (defined as bilinguals who had learned more than one additional language before age 10) showed a trend toward advantage over simple bilinguals, though the difference fell short of

statistical significance. She also claimed that among bilinguals there was a trend toward an advantage for those with formal education in a second language.

Toward a more exact specification of what advantages bilinguals and multilinguals possess, Werker (1986) investigated the effect of multilingualism on phonetic perceptual flexibility, *viz.* whether multilinguals or bilinguals were able to discriminate novel speech contrasts (in Hindi) better than monolinguals. In her experiment, there was no advantage whatsoever for either group. Nevertheless, Werker concedes that the contrasts used may have simply been too difficult for differences to appear during a limited number of trials; multilinguals might have shown an advantage sooner for a less difficult contrast. Werker also speculates that multilingualism perhaps does not improve perception itself but rather facilitates the learning of particular perceptual contrasts.

In a study of 91 college students enrolled in German and Japanese courses, Clark (1978) investigated (among other factors) the effect of previous language training on MLAT, with the results that among languages previously studied, only study of Latin made a difference. This led her to wonder if perhaps the MLAT measures what Latin courses teach. Also of interest in Clarke's study is the result that aptitude-achievement correlations were higher for Japanese than for German. Since the former but not the latter program featured a tutorial method using native-speaking tutors, this could be an indication that aptitude measurements have validity beyond the traditionally formal classroom setting.

Effects of multilingualism on later language learning

Besides Werker (1986) and Eisenstein (1980), which were reported in the previous section, there are two additional interesting studies which consider the effects of multilingualism on later language learning ability. In the earlier of these, Ramsay (1980) compared the performance of adult multilinguals (11) and monolinguals (9) on the task of learning a new language (Basque) under controlled conditions. The subjects were asked to learn as much Basque as possible during three sessions of approximately 40 minutes each, through self-study with an array of materials: videotaped conversations, tapescripts including free translations and a phonetic transcription, vocabulary cards, audio cassettes of the video soundtrack, grammar cards containing basic grammatical rules, and a programmed instruction "primer." The subjects were given a battery of cognitive/psychological tests during and after the sessions (but unfortunately no language aptitude test), as well as tests of target language skills. On the basis of rank orders on the final exam and a test of lexical memory, subjects were regrouped as successful language learners (SLs) and unsuccessful language learners (USLs); five of the six SLs were multilinguals.

Although there were many idiosyncratic features of Ramsay's study which weakened the generalizability of her results, some of her conclusions can be interpreted in relation to studies mentioned earlier. Ramsay concluded that:

The emerging picture of the SL is of an early and avid reader, conscious of style and vocabulary, who draws information from any available source. (p. 93)

Thus, the effects of literacy, metalinguistic awareness, and the ability to abstract knowledge from available experience may be crucial qualities in successful language learners. A final intriguing suggestion in Ramsay's conclusion is that researchers should perhaps shift in focus from language ability *per se* to the construct of "communicative systems plasticity." Although she does not state how

this concept could be operationalized, it may make more sense in the light of the next study to be discussed.

Nation and McLaughlin (1986) conducted a study which is similar to Ramsay's (1980) in that they both investigated the differential performance of multilinguals and monolinguals on the task of learning a new language under controlled conditions. Both studies were also specifically concerned with differences in how their subjects approached the learning tasks. However, whereas Ramsay interprets her results without reference to a particular theoretical framework, Nation and McLaughlin interpreted theirs in information processing terms. Other differences included Nation and McLaughlin's use of an artificial miniature linguistic system rather than a natural language, and their inclusion of a bilingual experimental group (making three groups of 11 subjects).

Nation and McLaughlin's tasks involved making grammaticality judgments about possible strings of symbols in the artificial language, under two different conditions: implicit and explicit learning. They found no significant differences on the explicit learning task but superior multilingual performance when the learning task was implicit. They interpreted this finding to mean that the multilinguals are better able to abstract structural information from linguistic stimuli that they are exposed to when there is no instruction or obvious reason to do so. In information processing terms, this means that multilinguals may "habitually exert more processing effort in making sense of verbal stimuli." (p. 52) This finding is consistent with Skehan's (1986) idea of the good language learner being able to abstract from experience in order to handle decontextualized material.

Nation and McLaughlin's results were disappointing in that : (1) the multilinguals did no better than the monolinguals when the learning task included instructions (the explicit condition); and (2) the bilinguals showed no significant difference from the monolinguals on either task. Nevertheless, the researchers offered plausible explanations for these results. As for (1), the multilinguals may have reached a ceiling level on performance on the implicit task, since it was first; and regarding (2), the study was simply not sensitive enough to reveal these differences. Nation and McLaughlin concluded that multilinguals excel in language learning because they maintain a flexibility between controlled and automatic processing, and because they have a "readiness to explore different routines and heuristic strategies." (p. 53) These two features combined may be close to what Ramsay (1980) has in mind with her concept of "communicative systems flexibility."

THE PRESENT STUDY

The literature on aptitude provides intriguing possibilities for conjecture, but as yet no clear empirically sound characterization of the relationship between language aptitude and previous foreign language experience has emerged. The present study seeks to contribute toward clarifying this relationship by directly addressing the research question: What aspects of previous language learning experience can predict current language aptitude?

HYPOTHESES

To specify more accurately what aspects of language experience may have an effect on language learning capacity, a number of hypotheses were formulated, all of them consistent with findings in one or more of the studies described above:

Language aptitude scores will correlate positively with:

- 1) Variables associated with previous instruction in foreign languages
- 2) Variables associated with previous naturalistic exposure to foreign languages
- 3) Level of attained proficiency in foreign languages
- 4) Self-rating of language learning ability.
- 5) Earliness of exposure to foreign languages
- 6) Subsequent success in foreign language learning

METHOD

Subjects

The 129 subjects in this study were students attending the 1988 South East Asian Studies Summer Institute (SEASSI) at the University of Hawaii at Manoa. The testing took place at the beginning of the Institute, prior to the subjects beginning language courses in Indonesian, Khmer, Tagalog, and Thai, and Vietnamese. There were 67 males (51.9%) and 62 females (48.1%), a large majority being native speakers of American English (80%). The subjects were neither randomly sampled nor self-selected, but were intact groups who participated as a consequence of their particular instructors' cooperation.

Materials

Measure of aptitude. The Short Form of the Modern Language Aptitude Test was Administered. The Short Form takes about 30 minutes and consists of three parts: *Spelling Clues*, *Words in Sentences*, and *Paired Associates*. *Spelling Clues* involves choosing the best among possible synonyms for words which have been spelled in a very idiosyncratic way. *Words in Sentences* requires recognizing the grammatical function of an underlined segment of a sentence, and then choosing the segment in a second sentence which has a parallel function. *Paired Associates* is a memorization task; subjects have a short time during which to memorize a list of Kurdish words and their English equivalents. Administering the Short Form does not require a tape player. The Long Form, which could not be used in this study, includes two additional sections, *Number Learning* and *Phonetic Script*, which do require a tape recorder. *Number Learning* involves listening to a short lesson on the number system in an artificial language, and then recognizing novel combinations of the numbers as they are spoken rapidly. *Phonetic Script* similarly features an aural teaching phase, this time of phonetic sound-symbol correspondences, and then a multiple-choice testing phase in which the correct graphemic representations for a variety of sounds must be selected.

Previous language learning experience. A "Language Experience Questionnaire," designed by the researcher, was administered. The questionnaire elicited information on age of onset, length and intensity of both formal instruction and informal exposure, and self-ratings of proficiency, for every language the subjects had had any contact with. It also asked for self-ratings of language aptitude, and current age. The full questionnaire appears in Appendix 1.

Current proficiency in the SEASSI language. As reported by Brown *et al.* (1990), tests designed to be equivalent for each of the SEASSI languages were developed. The sections of the test battery included *Listening*, *Oral Interview*, *Dictation*, and *Cloze*.

Measures of achievement. SEASSI language course grades were also obtained to corroborate the predictive validity of the MLAT.

Procedures

One hour was available for the entire procedure of completing the questionnaire and taking the aptitude test. Subjects were tested in groups according to the language they were studying. The language experience questionnaire was administered first, and then the MLAT. The cassette tape accompanying the MLAT was used for giving directions and timing the sections, so conditions for all groups can be considered equivalent, even though the test was administered by several different people.

The SEASSI tests were administered in a subsequent session by the same administrators; the procedure is described in detail in Brown *et al.* (1990).

The procedures for arriving at course grades was at the discretion at the individual instructors; no information is available.

Analyses

The total and part scores from the MLAT, the biographical variables, the SEASSI proficiency results, and course grades were analyzed (pairwise) for their Pearson product-moment correlation coefficients. Using the same correlation matrix, a principal components analysis was performed. All analyses were done on a Macintosh computer using the Systat 5.0 statistical package (Wilkinson 1989), with all data being imported from the Microsoft Excel spreadsheet program (Microsoft 1989).

RESULTS

Summary descriptive statistics are presented in Tables 1-5. Table 1 shows relevant statistics for the Modern Language Aptitude Test. It will be recalled that the parts of the short form of the test were as follows: (3) Spelling Clues; (4) Words in Sentences; and (5) Paired Associates. In the row beneath the labels for the parts, the K symbol indicates how many correct answers were possible for each section. The N of cases row shows that 129 subjects took all parts of the test. Then the minimum and maximum scores obtained by any subject are given; the range is the minimum subtracted from the maximum. The standard deviation then gives a preliminary idea of how much the individual subjects varied from each other. Skewness and kurtosis are two measures which indicate if the sample meets the assumptions for performing Pearson product-moment correlations. The table clearly shows that there was a large amount of variance among subjects, with some attaining nearly perfect scores, others answering very few items correctly.

Table 1. Descriptive Statistics (1)
Modern Language Aptitude Test (Short Form)

	TOTAL (K=119)	PART 3 (K=50)	PART 4 (K=45)	PART 5 (K=24)
N OF CASES	129	129	129	129
MINIMUM	15.000	4.000	6.000	0.000
MAXIMUM	106.000	48.000	44.000	24.000
RANGE	91.000	44.000	38.000	24.000
MEAN	66.682	22.031	27.202	17.457
VARIANCE	240.781	89.593	46.053	34.109
STANDARD DEV	15.517	9.465	6.786	5.840
SKEWNESS (G1)	-0.040	0.351	-0.248	-0.877
KURTOSIS (G2)	0.571	0.030	0.252	0.142

Table 2 provides the same statistics for the biographical variables concerning experience with what subjects considered to be their currently strongest second language. Again there is a large amount of variance, with zeros in the Minimum row usually not meaning that subjects had no second language, but rather that what they considered to be their second language was learned either exclusively in classrooms or naturalistically.

Table 2. Descriptive Statistics (2)
Previous language learning experience (L2)

	AGE OF ONSET	YEARS OF STUDY	HOURS/ WEEK	YEARS OF EXPOSURE	PERCENTAGE OF L2 USE
N OF CASES	127	127	127	127	100
MINIMUM	1.000	0.000	0.000	0.000	1.000
MAXIMUM	58.000	25.000	52.000	39.000	100.000
MEAN	17.126	3.863	8.612	2.485	46.490
STANDARD DEV	9.069	4.066	9.421	5.775	32.573
SKEWNESS (G1)	1.136	1.986	2.164	4.126	0.180
KURTOSIS (G2)	2.977	5.768	4.807	19.800	-1.368

Table 3 gives summary statistics indicating additional aspects of the subjects' multilingualism, as well as their current age. In fact, more detailed data was collected for each additional language that the subjects reported experience with, but this data was dropped from the analysis when preliminary results indicated that it added little new relevant information to the study.

Table 3. Descriptive Statistics (3)
Previous language learning experience (other)

	AGE OF L3 ONSET	TOTAL FL STUDY	TOTAL FL EXPOSURE	TOTAL # OF FLS	CURRENT AGE
N OF CASES	112	127	127	129	128
MINIMUM	1.000	0.000	0.000	0.000	19.000
MAXIMUM	33.000	30.000	57.000	9.000	66.000
MEAN	18.857	6.976	4.329	2.814	29.758
STANDARD DEV	7.122	5.414	9.253	1.327	8.609
SKEWNESS (G1)	-0.164	1.500	3.769	1.169	1.632
KURTOSIS (G2)	-0.412	3.062	16.090	3.935	3.301

Since it was impractical to attempt to obtain actual L2 proficiency measures from the subjects, self-ratings were gathered so as to at least get a general idea of their strengths and weaknesses in their L2s. These are presented, by skill area, in Table 4.

Table 4. Descriptive Statistics (4)
Self-ratings of L2 proficiency

	LISTENING	SPEAKING	READING	WRITING
N OF CASES	127	127	127	127
MINIMUM	1.000	1.000	0.000	0.000
MAXIMUM	6.000	6.000	6.000	6.000
MEAN	3.693	3.413	3.720	3.197
STANDARD DEV	1.504	1.438	1.522	1.517
SKEWNESS (G1)	-0.121	0.108	-0.500	0.052
KURTOSIS (G2)	-0.889	-0.731	-0.320	-0.643

It was also of interest to see how aware subjects were of their own aptitude; thus they were asked to rate their language learning ability in both classroom and naturalistic settings, 1 being "always near the bottom of the class" or "I pick up languages very slowly," respectively, and 6 being "always near the top of the class," and "I pick up languages very quickly."

Table 5: Descriptive Statistics (5)
Self-assessments of language aptitude

	CLASSROOM	NATURALISTIC
N OF CASES	129	129
MINIMUM	1.000	1.000
MAXIMUM	6.000	6.000
MEAN	4.286	3.626
STANDARD DEV	1.207	1.336
SKEWNESS (G1)	-0.439	0.026
KURTOSIS (G2)	-0.272	-0.750

Table 6 provides descriptive statistics for the SEASSI language proficiency tests. What is important to note here is that only slightly over half of the subjects took the listening and oral interview components of the test battery, and slightly less than half took the dictation and cloze components. This means that correlations involving these tests will require a much higher criterion for significance than those involving only variables which have been discussed so far. Another important consideration is that the statistics presented here represent aggregates for what are actually the components of five different tests: one for each of the SEASSI languages.

Table 6. Descriptive Statistics (6)
SouthEast Asian language proficiency tests

	LISTENING	ORAL INTERVIEW	DICTATION	CLOZE
N OF CASES	74	72	56	56
MINIMUM	3.000	1.000	3.000	2.000
MAXIMUM	32.000	9.000	50.000	27.000
MEAN	18.716	5.472	32.518	13.500
STANDARD DEV	6.183	1.736	11.050	5.520
SKEWNESS (G1)	-0.422	-0.295	-0.514	0.468
KURTOSIS (G2)	-0.014	-0.204	-0.008	0.170

Table 7 presents the all the patterns of intercorrelations involving the MLAT scores, SEASSI proficiency measures, and SEASSI course grades, along with the questionnaire variables in relation to them. In general, the experience variables were shown to be weakly or not related to proficiency, aptitude, and course assessments; suggestions in the literature that aptitude is enhanced by language experience were not borne out in this set of data.

One point of explanation regarding the criterion for significance is needed here. Since almost all of the variables, if related, would be logically related in only one direction, positive for all but age of onset, which would be supposed to be negative (see Long 1991 for review of maturational constraints in SLA), a directional (one-tailed) test was chosen. However, there were conflicting considerations of the direction of the relationship between any of the other variables and age, so for age only nondirectional (two-tailed) tests were chosen.

Table 7. Correlation matrix (partial): All variables in relation to MLAT and Course Grade
(Pearson Product-Moment Correlations)

	Course Grade	MLAT-SF Total	MLAT Part 3	MLAT Part 4	MLAT Part 5
MLAT Short Form Total	.431*				
MLAT 3 (Spelling Clues)	.308*	.828*			
MLAT 4 (Gramm. Sens.)	.341*	.648*	.291*		
MLAT 5 (Rote Learning)	.274*	.562*	.242*	.089	
Total Number of FLs	.234*	.234*	.161	.159	.176*
L2 Years of Exper. (C)	.035	.001	-.072	.138	-.032
Total FL Exper. (C)	.042	.048	-.067	.146	.071
L2 Hours/Week (C)	-.011	-.064	-.005	-.098	-.042
Total FL Exposure (N)	-.082	.175*	.121	.168	.070
L2 Years of Exposure (N)	-.032	.131	.081	.154	.035
L2 % of Use (N)	.124	.040	.060	-.108	.125
L2 Self-Rating (Listening)	.192*	.118	.028	.129	.120
L2 Self-Rating (Speaking)	.155	.091	-.001	.133	.090
L2 Self-Rating (Reading)	.109	.070	.004	.143	.015

L2 Self-Rating (Writing)	.047	.021	-.059	.143	-.011
SEASSI Listening	.533*	.269*	.202*	.162	.205*
SEASSI Oral Interview	.404*	.310*	.225*	.248*	.184*
SEASSI Dictation	.366*	.160	.034	.103	.221*
SEASSI Cloze Test	.361*	.111	.084	.007	.127
Aptitude Self-Rating (C)	.281*	.293*	.165	.233*	.240*
Aptitude Self-Rating (N)	.230*	.196*	.108	.070	.264*
L2 Age of Onset	-.042	-.102	-.052	-.145	-.018
Current Age	-.294*	-.194*	-.072	-.142	-.238*

* $p < .05$ (one tailed); $df = 127$, critical $r = .1726$

(C) = classroom settings

(N) = naturalistic settings

Not surprisingly, the strongest correlations are between the whole of the MLAT and its parts. Also as could be expected, the next strongest correlations are between course grade and the various components of the proficiency tests; they can be interpreted as measuring how much of a head start each learner has.

There are also moderately high correlations between MLAT scores and subsequent course grades. These come as no surprise either, but are important, because they tend to corroborate the MLAT's predictive validity at the same time as other biographical variables are shown not to be related to either aptitude or course grade.

Self-assessment of classroom language learning ability also showed itself to be related relatively strongly to measured (MLAT) aptitude and to eventual course grade; but with all correlations falling under .30, this finding, as most of the others in this study, must be interpreted with caution. A correlation of .30 means that only .09 of the variance is accounted for.

The next step in the study was to take a deeper look at the structure of the data, to determine if a smaller number of underlying factors could account for the data better than the many individual variables. To do this, a principal components analysis was performed. Principal components analysis is a variation of factor analysis, with important theoretical differences but generally with similar and more reliable results (see discussion in Wilkinson 1989, 79-82).

The first step was deciding on how many factors to extract. By using the conventional eigenvalue setting of 1, seven factors were extracted. To obtain a simpler and more economical analysis, five and six factor solutions were also tried, but they resulted in compressing the data too far, so that in addition to losing some of the variance accounted for, the factors became less rather than more interpretable. Thus, the seven-factor orthogonal solution with varimax rotation was decided upon. It accounted for 71.937% of the total variance, and produced the factor structure shown in Table 8.

The variables which load most heavily on each factor in Table 8 are enclosed in boxes, and those which exceed the criterion of .30 but have considerably lower loadings are marked by parentheses. Based on the component structure illustrated in the table, the patterns of loading can be interpreted as revealing the following underlying factors in the data:

Factor 1: L2 experience. The classroom experience measures, along with self-ratings of proficiency, load heaviest on this factor. It makes sense that learners should feel more proficient with more instruction, especially the relatively

successful language learners who comprise this sample. Naturalistic exposure also loads on this factor, but not as heavily, probably reflecting the predominance of instructed over naturalistic learners in the sample.

Factor 2: *Proficiency in SEASSI language.* All of the SEASSI test measures load heavily on this factor, with nothing else but the subsequent course grade in the same language. As discussed earlier, the relationship between proficiency and grade is to be expected.

Factor 3: *Language aptitude.* Parallel to the SEASSI test pattern, all of the MLAT scores load on this factor with course grade as the only other variable present. Thus, the pattern reasserts itself: Course grade is strongly related to both proficiency and aptitude, but to none of the other variables. Interestingly on this factor, course grade is more strongly related to the rest of the MLAT than is Part 4 (Paired Associates, requiring rote memorization). What may be reflected here is that rote memorization was much more crucial to success in language classes in the 1950s and 1960s, when the test was developed and validated, than in the classrooms of the 1990s, and especially for languages that do not have a long foreign language classroom tradition.

Factor 4: *Naturalistic language experience.* The two measures of years of naturalistic language experience load heavily on this factor, with the variable of age also appearing. In general, opportunities to visit target language settings accumulate with age.

Table 8. Principal components analysis with Varimax
Rotated loadings (Varimax), 7 factor solution

	1	2	3	4	5
	L2 Exp. (C)	SEASSI Proficiency	Language Aptitude	For. Lang. Exp. (N)	?Early Multiling.
L2 Years of Instruction (C)	.820	-.194	.029	.184	.059
L2 SELF-RATING (Reading)	.807	.098	-.009	.079	-.105
L2 SELF-RATING (Writing)	.781	.173	-.109	.103	-.102
TOTAL Years of FL (C)	.759	-.049	-.028	.326	.250
L2 SELF-RATING (Speaking)	.741	.232	-.022	.165	-.003
L2 Age of Onset	-.739	-.014	-.128	.082	-.197
L2 SELF-RATING (Listening)	.722	.268	.007	.188	.020
SEASSI Cloze Test	-.123	.812	.017	-.005	.118
SEASSI Dictation Test	.101	.777	-.077	.135	.082
SEASSI Listening Test	.152	.751	.161	-.049	.005
SEASSI Oral Interview	.074	.725	.197	.113	.099
SEASSI Course Grade	.096	.533	.474	-.254	-.004
MLAT TOTAL	-.011	.137	.932	.103	.130
MLAT 3 (Spelling Clues)	-.090	.039	.839	.085	-.056
MLAT 4 (Gramm. Sens.)	.226	.176	.616	.049	-.125
MLAT 5 (Rote Learning)	-.139	.088	.410	.070	.577
TOTAL Years FL Exp. (N)	.251	.050	.131	.835	.066
L2 Years of Exp. (N)	.350	.078	.100	.824	.051
Current Age	-.238	-.113	-.224	.519	-.595
L3 Onset Age	-.155	-.219	.297	-.167	-.758
APT. SELF-RATING (N)	.141	.178	.060	-.088	.119
APT. SELF-RATING (C)	.245	.157	.195	-.113	.061
TOTAL Number of FLs	.236	.118	.130	.150	.002
L2 % of Use (N)	.407	.035	.025	.117	.272
L2 Hours per Week (C)	-.012	.204	.015	-.058	-.260
	6	7			
	For. Lang. Confidence	L2 Intensity			
YEARS L2 (C)	-.039	-.138			
SELF READ L2	.253	.193			
SELF WRITE L2	.270	.126			
TOTAL YEARS FL	.173	-.132			
SELF SPEAK L2	.357	.308			
ONSET AGE L2	.047	.161			
SELF LISTEN L2	.338	.323			
CLOZE	-.204	.142			
DICTION	.306	-.090			
LISTENING	.275	.142			

ORAL INTERVIEW	.182	.027
COURSE GRADE	.138	-.051
MLAT TOTAL	.175	-.040
MLAT PART 3	.092	.146
MLAT PART 4	.014	-.440
MLAT PART 5	.301	.168
TOTAL FL (N)	-.009	-.029
YEARS L2 (N)	-.033	.058
CURRENT AGE	.017	.045
ONSET AGE L3	-.039	.067
SELF APTITUDE (N)	.784	.125
SELF APTITUDE (C)	.683	-.066
TOTAL # FLs	.601	-.144
% USE L2 (N)	.182	.618
HOURS/WK L2 (C)	-.270	.534

(C)=CLASSROOM SETTINGS,
(N)=NATURALISTIC SETTINGS

Factor 5: *Early multilingualism.* This factor is the least clear of the seven. There are negative loadings, fairly and very strong, respectively, for age and age of onset of an L3, and a moderately strong loading for MLAT 5 (Paired Associates, again). It can be inferred that young people, particularly young multilinguals, have good memorization abilities. The strange thing is that MLAT 5 loads heavier on this factor than on the aptitude factor.

Factor 6: *Self assessment of aptitude.* The two types of aptitude self-rating load heaviest on this factor, with total number of foreign languages also making a substantial contribution. This is a logical combination, with learners being more likely to attempt more languages if they perceive themselves as good at learning them. Self-ratings of listening and speaking also appear weakly here, with the other two self-ratings not far below criterion. Finally, MLAT 5 shows up once again, indicating perhaps that, when learners perceive themselves as good as languages, it is memorization ability that is most salient for them.

Factor 7: *Intensity of L2 experience.* The two intensity measures, of classroom and naturalistic language experience respectively, have the heaviest loadings on this factor. Additionally, self-ratings of speaking and listening load rather lightly, whereas there is a moderately strong negative loading for MLAT 4, the Words in Sentences section, measuring grammatical sensitivity. It could be that intensive courses and periods of exposure to a foreign language are conducive to promoting (at least perceived) speaking and listening ability, and are preferred by learners who are less concerned about the more systematic aspects of language learning.

Admittedly, a large part of the interpretation of these factors is conjecture. However, the patterns of loadings are fairly clear, and there is surprisingly little overlap, *i.e.*, there are few variables that load on multiple factors.

DISCUSSION

Going back to the original hypotheses, what can these results tell us about the relationship between language aptitude and previous language learning experience? The Pearson product-moment correlations tell us that

demonstrated success on tests of language proficiency (at least listening, oral interview, and dictation) are related to aptitude in our sample, as is self-assessment of aptitude, but none of the biographical background variables from the questionnaire entered into strong, consistent correlations with aptitude. Since it is difficult to reconcile such a result with the literature on the capabilities of bilinguals and multilinguals, it is tempting to claim that the Modern Language Aptitude Test does not effectively measure language learning capability. This may be the case, and arguments against the face validity of the Short Form as an adequate measure of aptitude could easily be made, but the fact that the aptitude scores did correlate significantly with subsequent course grades strengthens the evidence that the null findings for previous experience were not a fluke. Moreover, the obtained correlation of .431 between MLAT Short Form and subsequent course grade in a South East Asian language is nearly identical to the .42 that Carroll (1959) obtained with non Indo-European language course assessments in his original validation studies.

The results of the factor analysis further corroborate the trustworthiness of the correlational findings. The biographical variables clustered into several different plausible factors, fairly independent of one another, and certainly independent of the proficiency factor and the aptitude factor. These in turn were also clearly independent of one another, bridged only by SEASSI language course grade. Not only is the double loading of course grade far from being problematic, but doubt would be cast on one or both of the test batteries if this doubling up did not occur. Another interesting discovery of the factor analysis is the multiple loadings of MLAT 5, which involves short-term rapid rote memorization. Weakest in loading on the aptitude factor, it also loads (negatively) with age and late multilingualism, and positively with self-rating of aptitude and total languages reported. This pattern is simply too idiosyncratic to be explained on the basis of this data set alone. More data is needed, especially with all five MLAT components being represented, before more specific interpretations can be ventured.

Nevertheless, this anomaly, together with the conflicting findings of previous research on language learning experience in relation to capability for future language learning, points to the need for considering that various aspects of aptitude may be quite independent of each other in different profiles of learners, as suggested in the Wesche (1981) study mentioned earlier. A fruitful direction for research, therefore, may be in seeking out aspects of aptitude not measured by the currently available language aptitude tests, and then pinpointing how these and previously identified components interact in different learners under different learning conditions.

The present study suffers from a number of limitations. As mentioned, for a study dealing directly with language aptitude, the Short Form of the MLAT may be inadequate. Although the sample size was reasonable to begin with, there are problems due to missing values on many of the variables, and the fact the students were treated as one large group when in fact they constituted several groups, studying different languages, at different levels, with different teachers; some even had different native languages. Also, there is no way to take account of the teaching process which mediated between the initial proficiency and aptitude scores, on the one hand, and the final course grades, on the other. Nor is it possible to know the range, weighting and consistency of the criteria involved in grading. Finally, the Language Experience Questionnaire may not have been sensitive enough to register aspects of experience which may be crucial to aptitude.

CONCLUSION

Although the present study did not turn up any strong connections between previous language learning experience and language aptitude, the case is far from closed. Instrumentation problems could have led to the null findings, especially considering the heterogeneity of the group. More refined analysis of the present data is called for, as are further studies in which the variables are controlled better. Obvious improvements would be use of the long form of the MLAT, independent corroboration of self-report variables and refinement of the items on the Language Experience Questionnaire, and specification of how course grades are arrived at in the subsequent language courses.

There are many additional paths to be pursued in language aptitude research as well. One, briefly mentioned earlier, is to explore the possible interaction of learner aptitude profile and type of instruction; Wesche (1981) and Skehan (1991) are advocates of this approach. Another important line of inquiry is that which seeks to find either additional components of aptitude not measured by the MLAT, or more explanatory cognitive characterizations for the abilities which are currently measured. For example, working memory (McLaughlin, 1990; Harrington & Sawyer, in press), may account for a large part of the variance in language learning achievement not accounted for by the MLAT, and/or it may be the capacity that underlies successful performance on parts of the MLAT, especially Parts 1, 2, and 5. Working memory represents but one of many cognitive constructs that have come to be better understood through research in cognitive psychology during the 30 years since Carroll and Sapon developed the MLAT.

A related area of needed research is that which seeks to clarify the relationship between language aptitude and learner strategies. There is currently an effusion of enthusiasm in the language teaching field over learner strategies, yet there is no solid empirical evidence that they can be taught to learners so as to influence learning outcomes. If whatever variables that are associated with language experience are of no use in affecting learning capability, why should learner strategies be expected to help? On the other hand, however, learner strategies may be exactly the kind of thing which can compensate for deficiencies in aptitude, or perhaps serve to trigger enhancement of aptitude. This area of research most definitely offers a wealth of fascinating work to be undertaken in the future.

References

- Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*, 48, 1009-1018.
- Bialystok, E. (1987). Influences of bilingualism on metalinguistic development. *Second Language Research* 3, 2, 154-166.
- Brown, J., H. Cook, C. Lockhart, and T. Ramos (1990). SouthEast Asian Language proficiency examinations. Paper presented at 1990 RELC International Conference on Language Testing and Programme Evaluation, Singapore.
- Carroll, J.B. (1963). A model of school learning. *Teachers College Record*, 64, 723-733.
- Carroll, J.B. (1965). The prediction of success in foreign language training. In R. Glaser (Ed.), *Training, research, and education*. New York: Wiley, 87-136.
- Carroll, J.B. (1974). The aptitude-achievement distinction: the case of foreign language aptitude and proficiency. In D.R. Green (ed.), *The aptitude-achievement distinction*. Monterey, CA: CTB/McGraw-Hill, 289-303.

- Carroll, J.B. (1981). Twenty-five years of research on foreign language aptitude. In K.C. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House, 83-118.
- Carroll, J.B., and Sapon, S.M. (1959). *Modern Language Aptitude Test (MLAT): Manual*. New York: The Psychological Corporation.
- Clarke, S. (1978). The correlation between aptitude and achievement in Japanese and German. In J.E. Redden (ed.), *Occasional papers in linguistics number 3: proceedings of the Second International Conference on Frontiers in Language Proficiency and Dominance Testing*. Carbondale, IL: Southern Illinois University.
- Eisenstein, M. (1980). Childhood bilingualism and adult language learning aptitude. *Revue Internationale de Psychologie Appliquee* 29, 1-2, 159-172.
- Ellis, R. (1986). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Gardner, R. C. (1978). Social psychological aspects of second language acquisition. In H. Giles and R.N. ST. Clair (eds.), *Language and social psychology*. Baltimore, MD: University Park Press, 193-220.
- Gardner, R. C., Lalonde, R.N., and Moorcroft, R. (1985). The role of attitudes and motivation in second language learning: correlational and experimental considerations. *Language Learning* 35, 2, 207-227.
- Gardner, R.C. (1986). *Social psychology and second language learning*. London: Edward Arnold.
- Gardner, R.C. and Lambert, W.E. (1965). Language aptitude, intelligence, and second-language achievement. *Journal of Educational Psychology* 56, 191-199.
- Genesee, R. (1976). The role of intelligence in second language learning. *Language Learning*, 26, 267-280.
- Harrington, M. & Sawyer, M. (in press). L2 working memory capacity and L2 reading skill. *Studies in Second language Acquisition*.
- Krashen, S.D. (1981). Aptitude and attitude in relation to second language acquisition and learning. In K.C. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House, 155-175.
- Masny, D. and d'Anglejan, A. (1985). Language, cognition, and second language grammaticality judgments. *Journal of Psycholinguistic Research*, 14, 2, 175-197.
- McInnis, C.E. (1985). Review of the Modern Language Aptitude Test. In D.J. Keyser and R.C. Sweetland (eds.), *Test critiques. Volume 5*. Kansas City, MO: Test Corporation of America, 271-277.
- McLaughlin, B. (1990). Another look at aptitude. Paper given at Ohio State University, October 14, 1990.
- Microsoft (1989). *Microsoft Excel: Complete Spreadsheet with business graphics and database (Version 2.2)*. Microsoft Corporation.
- Nation, R. and B. McLaughlin (1986). Novices and experts: and information processing approach to the "good language learner" problem. *Applied Psycholinguistics* 7, 41-56.
- Neufeld, G. (1978). A theoretic perspective on the nature of linguistic aptitude. *International Review of Applied Linguistics*, 16, 15-25.
- Neufeld, G. (1979). Towards a theory of second language learning ability. *Language Learning*, 29, 227-241.
- Oller, J. & K. Perkins (1978). Intelligence and language proficiency as sources of variance in self-reported affective variables. *Language Learning*, 28, 85-97.

- Pimsleur, P. (1966). *The Pimsleur Language Aptitude Battery*. New York: Harcourt Brace Jovanovitch.
- Ramsay, R. (1980). Language learning approach styles of adult multilingual and successful language learners. *Annals of the New York Academy of Sciences*, 345, 73-96.
- Reves, T. (1983). What makes a good language learner? Unpublished Ph.D. dissertation, Hebrew University of Jerusalem.
- Skehan, P. (1986). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3, 2, 188-221.
- Skehan, P. (1987). Cluster analysis and the identification of learner types. In V. Cook (ed.), *Experimental approaches to language acquisition*. London: Pergamon.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition*, 13, 275-298.
- Skehan, P. (in press). Where has language aptitude come from? In P. Meara (ed.), *Proceedings of the 1985 BAAL conference*. Center for Information on Language Teaching.
- Wells, C. (1981). *Learning through interaction*. Cambridge: Cambridge University Press.
- Wells, C. (1985). *Language development in the pre-school years*. Cambridge: Cambridge University Press.
- Werker, J.F. (1986). The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics*, 7, 2, 141-155.
- Wesche, M.B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In D.C. Diller (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House, 119-153.
- Wesche, M.B., Edwards, H., and Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics*, 3, 2, 127-140.
- Wilkinson, L. (1989). *SYSTAT: The system for statistics*. Evanston, IL: SYSTAT Inc.