Appendix A Methodological Features of the Included Studies

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Study ID*** | ***Aptitudea*** | ***Covariateb*** | ***Scoresc*** | ***Criteriond*** | ***Learnere*** | ***Agef*** | ***L1g*** | ***L2h*** | ***Ni*** | ***Settingj*** |
| Abrahamsson & Hyltenstam ’08 | Swansea LAT |  |  | G | N | 38.3/  33/39 | S | SW | 42 | SL |
| Alderson et al. ’97 | MLAT 4 |  |  | G, P | U | 19 | E | F | 121 | FL |
| Bain et al. ’10 | MLAT-S |  |  | P | U | 20 | E | S | 95 | FL |
| Bell ’09 | MLAT 4, PLAB 4 | PSTM, EWM |  |  | MIX | NR | E | F | 36 | SL |
| Bell & McCallum ’12 | MLAT-S | ANX | Reported | P | U | 21.5 | E | S | 95 | FL |
| Bialystok & Frohlich ’78 | MLAT-S | MOT |  | G, L, R | H | 16 | E | F | 157 | FL |
| Biedron & Szczepaniak ’12 | MLAT | INT, EWM |  |  | MIX | 24 | P | MIX | 64 | FL |
| \*Bond ’11 | MLAT-S |  | Reported | G, P | U | 19.35 | E | S | 30 | FL |
| \*Burgos-Kohler ’91 | MLAT-S |  |  | P | U | 19 | E | S | 143 | FL |
| \*Carpenter ’08 | MLAT-S | INT, EWM |  |  | MIX | 23 | E | NR | 29 | FL |
| Carroll & Sapon ’02k | MLAT/  MLAT-S |  | Reported | P | H, U | NR | E | MIX | 3835 | FL |
| Cochran et al. ’10 | MLAT 1, 2, 3 |  |  | P | U | 21.5 | E | S, F, G | 648 | FL |
| Cooper ’64 | MLAT-S |  |  | P | U | 18 | E | G | 95 | FL |
| Curtin et al. ’83 | PLAB |  | Reported | P | H | 16.5 | E | MIX | 563 | FL |
| DeKeyser ’00 | HUNLAT 4 |  |  | G | N | 55/60/43.2 | H | E | 57 | SL |
| DeKeyser ’93 | Dutch MLAT 4 |  |  | G, S | H | 17 | D | F | 35 | FL |
| Ehrman & Oxford ’95 | MLAT |  |  | R, S | A | 39 | E | MIX | 855 | FL |
| Ehrman ’98 | MLAT |  |  | R, S | A | 39 | E | MIX | 343 | FL |
| Eisentein ’80 | MLAT-S |  | Reported |  | U | 20 | E | NR | 93 | NR |
| Erlam ’05 | MLAT 4, PLAB 5 | PSTM | Reported |  | H | 14 | E | F | 92 | FL |
| \*Fujii ’05 | LABJ 2 | PSTM, EWM |  |  | U | 18 | E | J | 15 | FL |
| Ganschow & Sparks ’96 | MLAT | ANX |  |  | H | 14.5 | E | F, G, S | 154 | FL |
| \*Garcia ’96 | MLAT-S | MOT |  | P | U | 25 | E | S | 135 | FL |
| Gardner & Lambert ’65 | MLAT |  |  | G, L, P, R, S, V | H | 16 | E | F | 96 | FL |
| Gardner & Lysynchuk ’90 | MLAT 1, 3, 4 |  |  | G, V | H | 14 | E | F | 128 | FL |
| Gardner et al. ’89 | MLAT-S |  |  | S | I | NR | E | F | 89 | SL |
| Gardner et al. ’97 | MLAT-S |  |  | P | U | 19 | E | F | 102 | SL |
| Goodman et al. ’90 | MLAT-S |  | Reported |  | U | NR | E | MIX | 587 | FL |
| \*Gordon ’80 | MLAT-E | ANX, MOT |  | P, W | E | 12.9 | MIX | E | 129 | SL |
| Granena & Long ’13 | LLAMA |  |  | G, S, V | N | 24.5/  31.78/24.19 | C | S | 77 | SL |
| Harley & Hart ’97 | MLAT 5, PLAB 4 |  | Reported | L, P, S, V, W | H | 16 | E | F | 65 | SL |
| Harley & Hart ’02 | PLAB 4 |  | Reported | L, P, S, V, W | H | 15.5 | E | F | 31 | SL |
| Hauptman ’71 | MLAT-E | INT |  |  | E | 10.5 | E | J | 69 | FL |
| \*Horwitz ’80 | MLAT-S |  | Reported | G, S | H | 15.9 | E | F | 79 | FL |
| \*Hsieh ’04 | MLAT | MOT |  |  | U | 21 | C | E | 207 | FL |
| Hummel ’09 | TALV | PSTM |  | G, P, R, V | U | 21.5 | F | E | 27 | FL |
| \*Keitges ’86 | GTT | ANX, MOT |  | L, G, R, P | U | 20.1 | J | E | 168 | FL |
| Kormos & Trebits ’12 | HUNLAT |  |  | S, W | H | 16.5 | H | E | 44 | SL |
| \*Lado ’08 | MLAT | PSTM, EWM |  |  | U | 21.2 | E | S | 151 | FL |
| Lalonde & Gardner ’84 | MLAT-S | ANX |  |  | U | NR | E | F | 88 | FL |
| \*Lee ’08 | LAT | EWM |  |  | H | 16.93 | K | E | 154 | FL |
| Li ’13 | MLAT 4 | EWM | Reported |  | U | 20.8 | E | C | 78 | FL |
| Mackey et al. ’02 | LABJ | PSTM, EWM |  |  | U | 25.7 | J | E | 30 | SL |
| \*McMullin ’88 | MLAT 1, PLAB 3-6 |  | Reported |  | M | 12.6 | E | F | 95 | FL |
| \*Myer ’85 | PLAB 5, 6 |  |  | L | M | 13 | E | F, G, S | 54 | FL |
| \*Natelson ’75 | MLAT |  |  | S | A | NR | E | MIX | 914 | FL |
| Parry & Child ’90 | MLAT, VORD |  |  | R, S | A | 39 | E | MIX | 36 | FL |
| \*Phillips ’98 | MLAT |  | Reported | P | U | 22.42 | E | S | 98 | FL |
| \*Rachlin ’05 | PLAB |  |  | P | M | 12 | E | S | 221 | FL |
| Robinson ’02 | LABJ | INT, EWM |  |  | U | 21.5 | J | E | 54 | FL |
| Roehr & Ganem-Gutierrez ’09 | MLAT | EWM |  | G | U | 23.85 | E | G, S | 39 | FL |
| Ross et al. ’02 | LABJ 3, PLAB 4 |  |  | P | U | 19.5 | J | E | 129 | FL/SL |
| \*Sachs ’10 | MLAT-S |  | Reported |  | U | 21 | E | J | 80 | FL |
| Safar & Kormos ’08 | HUNLAT | PSTM, EWM |  | L, P, R, S, W | H | 15.5 | H | E | 40 | FL |
| Sparks et al. ’95 (1) | MLAT |  |  | P | H | 14.7 | E | MIX | 154 | FL |
| Sparks et al. ’95 (2) | MLAT |  |  | P | H | 14.6 | E | MIX | 100 | FL |
| Sparks et al. ’98 (1) | MLAT |  |  | P | H | 16.1 | E | F, G, S | 60 | FL |
| Sparks et al. ’98 (2) | MLAT |  |  | P | H | 16.1 | E | F, G, S | 36 | FL |
| Sparks et al. ’09 | MLAT | INT, MOT |  | P, R, S, W | H | 16.3 | E | F, G, S | 54 | FL |
| Sparks et al. ’11 | MLAT | ANX, INT, MOT |  | P, R, S, V, W | H | 16.3 | E | F, G, S | 54 | FL |
| Sparks et al. ’12 | MLAT |  |  | S, P | H | 16.3 | E | F, G, S | 54 | FL |
| Sparks & Patton ’13 | MLAT | ANX, INT |  |  | H | 16.5 | E | F, G, S | 128 | FL |
| Wesch et al. ’82 | MLAT | INT |  |  | A | 37 | E | F | 793 | FL |
| \*Winke ’05 | MLAT |  |  | L, R, S, V, W | U | 18.71 | E | C | 42 | FL |
| Winke ’13 | MLAT | EWM, MOT |  |  | A | 25 | E | C | 96 | FL |
| \*Yoshizawa ’02 | LABJ |  |  | G, L, R | U | 19 | J | E | 528 | FL |

*Note.* \*Ph.D. or M.A. thesis.

1. Measures of aptitude: MLAT-S = short form of MLAT, MLAT-E = MLAT Elementary, HUNLAT = Hungarian version of the MLAT, LABJ/GTT = adapted MLAT in Japanese, LAT= Language Analysis Test, TALV = French version of MLAT
2. PSTM = phonological short-term memory, EWM = executive working memory, ANX = anxiety, MOT = motivation, INT = intelligence;
3. Whether descriptive statistics of aptitude scores were reported: some studies reported subtest scores but only full MLAT scores were analysed for the question on the stability of aptitude
4. The criterion variable refers to outcome measures: G = grammar, L = listening, P = proficiency, R = reading, S = speaking, V = vocabulary, W = writing
5. Learner type: A = adult (learners from government sponsored programs), E = elementary, H = high school, , I = immersion, M = middle school, Mix = mixed N = naturalistic, U = university,
6. Learners’ age: NR = not reported, multiple numbers separated by slashes meaning the study included learners of several age groups;
7. Learners’ first language: C = Chinese, D = Dutch, E = English, F = French, H = Hungarian, J = Japanese, K = Korean, MIX = mixed L1 backgrounds, P = Polish, S = Spanish;
8. Learners’ second language: abbreviations same as those for learners’ first language except that SW = Swedish, NR = not reported, G = German;
9. Number of participants,
10. Research setting: SL = second language, FL = foreign language, NR = not reported;
11. The MLAT manual contributed data based on 38 aptitude scores (33 total MLAT and 5 short MLAT scores) and 34 correlation coefficients for the associations between overall aptitude and general L2 proficiency.

Appendix B Evaluation of Publication Bias

-2.0

-1.5

-1.0

-0.5

0.0

0.5

1.0

1.5

2.0

0

10

20

30

40

50

Precision

Fisher's Z

The funnel plot in Figure 1, which displays the variation of effect sizes as a function of precision (higher precision represents a larger sample size), shows that publication bias was present in the dataset. As can be seen, the distribution of the effect sizes (the empty circles) is asymmetrical: At the bottom there are more data points on the right side of the mean effect size than on the left, which means studies with smaller sample sizes and smaller effect sizes are missing. A trim-and-fill analysis showed that a total of 22 studies were needed to make the distribution symmetrical, and imputing these studies (the solid circles) would change the mean effect size from .41 to .33 (in Fisher’s Z). A Fail-Safe N analysis demonstrated that altogether 31,668 studies were needed in order for the *p* value for the population effect to exceed .05. However, such analysis only identifies the number of “null” studies that would make the population effect nonsignificant and is therefore less informative than a trim-and-fill analysis, which projects the change of the mean effect size if the missing values were to be included in the analysis.