

Supplementary materials 1.

Information, reliability and validity measures for the different subtests.

KAIT. The American KAIT, developed in 1993 by A.S. Kaufman and N.L. Kaufman, was translated by Dekker, Dekker, and Mulder in 2004 and norms were collected on a standardization sample in the Netherlands and Flanders. The main goal of the KAIT is to evaluate analytic intelligence in individuals from 14 to 85 years old. The complete test was administered which consists of 10 subtests categorized into two types of intelligence: fluid and crystallized intelligence. The crystallized scale consists of four subtests: Word Definitions, Double Meanings, Auditory Comprehension, and Personalities. It reflects how well a person has learned concepts and knowledge that are part of the cultural and scholar context. It is influenced by verbal conceptual development and education. The fluid intelligence scale gives an indication of the person's potential and flexibility to solve new problems. The four subtests are Symbol Learning, Logical Reasoning, Secret Codes and Block Patterns. Additionally, there are two measures of long term memory, namely Delayed Auditory Memory and Delayed Symbol Learning. The combination of the fluid IQ score, the crystallized IQ score and the delayed subtests results in a total IQ-score. All three scores have a mean of 100 and a standard deviation of 15 points.

The KAIT was administered instead of the Wechsler Adult Intelligence Scale III (Wechsler, 2001) to avoid retest effects on the WAIS. Many students with dyslexia had been tested previously with the WISC or the WAIS as part of their assessment. Other reasons for choosing the KAIT were the less rigorous time constraints, which can be considered as an advantage for students with learning disabilities, and the inclusion of two subtests of delayed memory, namely Delayed Symbol Learning and Delayed Auditory Memory. Both subtests are considered valid measures of long term memory capacities.

The combination of fluid and crystallized IQ results in a total IQ-score. All three scores have a mean of 100 and a standard deviation of 15 points. Psychometric information can be found in the following table.

	Internal consistency Chronbach's alpha for age groups 16-19	Test-retest reliability for age group 14-24	Content validity: correlation with WAIS -R Total IQ scores
CIQ	.92	.80	.79
Definitions	.82	.81	
Double Meanings	.81	.72	
Auditory Comprehension	.81	.71	
Famous People	.76	.87	
FIQ	.93	.84	.76
Symbol Learning	.93	.85	
Logical Reasoning	.81	.66	
Secret Codes	.80	.61	
Block Patterns	.80	.82	
Delayed Auditory Comprehension	.55	.49	
Delayed Symbol Learning	.93	.81	
TIQ	.95	.89	.84

GL&SCHR (De Pessemier & Andries, 2009). The GL&SCHR, a Dutch reading and spelling test battery for (young) adults (De Pessemier & Andries, 2009) was also administered. This test includes many of the tasks frequently administered in dyslexia assessment (see above). There are three main tests specifically designed to evaluate reading and writing skills, namely Word Spelling, Proofreading, and Text Reading. Seven additional tests focus on associated language deficits such as phonological processing, rapid naming, short term memory and working memory, morphology and syntax, automation, text comprehension and vocabulary.

Information about reliability can be found in the table below. For different subtests different methods were used, namely KR20, Guttman split-half, and a test-retest correlation.

	KR20	Guttman split half (γ)	test-retest
Text Reading		.77 < r < .90	
Word Spelling (Word Spelling and Proofreading)		.69 < r < .80	
Reading Comprehension		.61	
Morphology and Syntax		.65	
Short Term Memory		.54 < r < .77	
Vocabulary	.90		
Phonological Awareness (Spoonerisms and Reversals)			.78 < r < .90
Rapid naming			.62 < r < .84

IDAA (Van der Leij et al., 2012). The IDAA is a new, standardized diagnostic instrument for dyslexia in young adults. Norms have been collected on secondary school children (final two years, ages from 16 to 18). This test battery was developed by The University of Amsterdam, Lessius College for Higher Education (Antwerp), and Muiswerk.

The five subtests we used in this study form the core of the IDAA, namely Reversals, Lexical decision, Flash typing words, Flash typing pseudowords and Flash typing English words¹. For this test the participant is seated in front of a computer screen wearing headphones. The test battery is fully computer administered. Instructions are given visually on the computer screen and auditorially through headphones. For the registration of reactions a standard computer keyboard is used. The sequence of the tasks is identical for each participant. During administration, no interaction takes place between the participant and the test leader.

This computer based assessment tool for diagnosing dyslexia has been validated in Flanders and the Netherlands. In Flanders test-retest reliability varies from .74 up to .90 for the five subtests. As for validity, the correlation with the OMT and the Klepel was .80.

¹ There are two more subtests in the IDAA (all administered) that were not included in this study: a questionnaire rating print exposure and a test to measure baseline reaction speed. This second subtest is only used to rule out significant problems with a computer based administration. None of the participants exhibited any problems on this domain.

EMT (Brus & Voeten, 1991). A classic word reading test in the Dutch-speaking countries is the *EMT* [One Minute Test]. Parallel-form reliability ranges from .89 to .97 in various studies, whereas test-retest reliability lies between .82 and .92. For more psychometric information about the EMT we refer to the test's manual.

OMT (Kleijnen & Loerts, 2006). The English version of the EMT, namely the One Minute Test or OMT was used as a measure for English word reading skill. Validity and reliability data of the OMT have been collected by Kleijnen, Steenbeek-Planting, and Verhoeven. Test-retest reliability varies between 0.87 and 0.92.

Tick Bite (Henneman, Kleijnen, & Smits, 2004). The test that was used -"Hoe gevaarlijk is een Tekenbeet? [How Dangerous Can a Tick Be?]" - is part of a screening instrument published. It provides an indication of silent reading speed and the ability to retain information. There are no norms for Flanders. To obtain further information about the validity of the test, the correlation with the EMT word reading test was calculated for this sample. A Pearson correlation coefficient of .66 (N = 200) was found.

Klepel (van den Bos, Spelberg, Scheepsma, & de Vries, 1999). The standard Dutch nonword reading test is *De Klepel*. The parallel-forms correlation varies between .89 and .95. In various studies, the results of the Klepel correlate between .74 and .91 with those of the EMT. For more psychometric information about the Klepel we refer to the test's manual.

WRAT III (Wilkinson, 1993). We used a standardized English test for word spelling: the *WRAT-III English Word Dictation*. The internal consistency coefficients for the English age groups 17-18 and 19-24 are both .90. For more information on validity and reliability in English we refer to the manual. Because this test has not yet been validated for bilinguals with Dutch as mother tongue, the Pearson correlation was calculated with the English flash typing test of the IDAA ($r = 0.72$; $N = 200$).

AT-GSN (Ghesquière, 1998). This test has been used in a number of scientific studies. Further information about the validity was obtained by correlating the scores with those of the Word Spelling test of the GL&SCHR ($r=.79$) and with the Dutch flash typing test of the IDAA ($r=.70$).

TTR (de Vos, 1992). The Tempo Test Rekenen, a Dutch standardized test for mental calculations was administered. The psychometric value of the test has been demonstrated on a

sample of 10,059 children (Ghesquière & Ruijsenaars, 1994). Cronbach's alpha for the current study was .89 for all groups.

CDT (Dekker, Dekker, & Mulder, 2007). to measure the participants' speed of processing, we used the CDT or Cijfer Doorstreept Test [Digit Crossing Test]. This is a standardized Dutch test to detect attentional deficits and measure the speed and accuracy of processing in a task of selective attention involving task-switching. It is one of the 23 tests of the DVMH [Differential Aptitude Tests for Middle and Higher Level], a test battery published in 2003 by Dekker and De Zeeuw. This test battery was developed according to Carroll's Three Stratum Model in order to assess a large variety of cognitive skills such as verbal and numerical reasoning, attentional skills and language skills. The test – retest reliability scores vary between 0.79 and 0.95.

- Brus, B., & Voeten, M. (1991). *Een-minuut-test vorm A en B, schoolvorderingstest voor de technische leesvaardigheid bestemd voor groep 4 tot en met 8 van het basisonderwijs. Verantwoording en handleiding*. Lisse, The Netherlands: Swets & Zeitlinger.
- De Pessemier, P., & Andries, C. (2009). *GL&SCHR Dyslexie bij +16-jarigen. Test voor Gevorderd Lezen en Schrijven*. Antwerpen, Belgium: Garant.
- de Vos, T. (1992). *Tempo Test Rekenen*. Amsterdam: Pearson Education.
- Dekker, R., Dekker, P. H., & Mulder, J. L. (2007). *De ontwikkeling van vijf nieuwe Nederlandstalige tests*. Leiden: PITS.
- Ghesquière, P. (1998). *Algemene toets gevorderde spelling van het Nederlands (AT-GSN): verantwoording en handleiding. Rapport van een specialisatiejaar: onderzoek At-GSN-dictee.*: Unpublished bachelor thesis, University Leuven, Leuven, Belgium.
- Ghesquière, P., & Ruijsenaars, A. (1994). *Vlaamse normen voor studietoetsen rekenen en technisch lezen lager onderwijs [Flemish standards for study evaluation of mathematics and technical reading in primary schools]*. Catholic University of Leuven, Center for Educational and Professional Guidance. Leuven, Belgium.
- Henneman, K., Kleijnen, R., & Smits, A. (2004). *Protocol Dyslexie Voortgezet Onderwijs: Deel 2- Signalering, diagnose en begeleiding*. : KPC Groep, Expertisecentrum Nederlands, Werkverband Opleidingen Speciaal Onderwijs.
- Kaufman, A. S., & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test. Manual*. Circle Pines MN: American Guidance Service.
- Kleijnen, R., & Loerts, M. (2006). *Protocol Dyslexie Hoger Onderwijs*. Antwerpen/Apeldoorn: Garant.
- van den Bos, A., Spelberg, H., Scheepsma, A., & de Vries, J. (1999). *De Klepel vorm A en B: een test voor leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling*. Lisse, The Netherlands: Swets & Zeitlinger.
- Van der Leij, A., Bekebrede, J., Geudens, A., Schraeyen, K., G.M., S., Garst, H., . . . Schijf, T. J. (2012). *Interactieve Dyslexietest Amsterdam-Antwerpen: Handleiding*. Uithoorn.
- Wilkinson, G. S. (1993). *Wide Range Achievement test*. Lutz, Florida: PAR.

Supplementary materials 2.

Variables ranged from large effect size to small effect size, t-values and exact p-values.

Variable °	Variable	M (dys)	SD (dys)	M (contr)	SD (contr)	Effect sizes	t-value	p-value
53	Word spelling (GL&SCHR)	91.650	15.692	121.400	12.843	1.440	-14.710	0.000
52	Word reading correct (EMT)	77.064	14.112	100.420	10.577	1.670	-13.430	0.000
51	Flash typing English words (IDAA)	28.720	5.551	36.810	2.748	1.570	-13.620	0.000
50	Sentence spelling (AT-GSN)	51.619	18.885	23.200	11.648	1.430	12.080	0.000
49	Flash typing pseudowords (IDAA)	15.165	4.118	22.120	4.073	1.950	-12.080	0.000
48	Lexical decision (IDAA)	27.906	3.479	32.930	2.508	1.760	-11.150	0.000
47	Pseudoword reading correct (KLEPEL)	41.300	10.407	59.580	12.751	1.470	-11.270	0.000
46	Flash typing words (IDAA)	34.400	3.169	38.190	1.587	1.070	-10.940	0.000
45	English word spelling (WRAT)	16.660	4.804	24.270	5.418	1.940	-10.510	0.000
44	English word reading correct (OMT)	41.300	10.407	59.580	12.751	1.870	-10.150	0.000
43	Text reading time (GL&SCHR)	308.490	44.017	258.530	25.260	1.440	9.440	0.000
42	Spoonerisms time (GL&SCHR)	176.334	54.014	115.439	34.896	1.140	9.470	0.000
41	Reversals (IDAA)	43.400	7.004	51.300	6.260	1.240	-8.410	0.000
40	Reversals time (GL&SCHR)	102.645	26.483	76.610	16.181	1.220	8.890	0.000
39	Mental calculation mix (TTR)	22.930	4.450	28.330	4.981	0.994	-8.850	0.000
38	Silent Reading (Tickbite)	343.895	87.392	263.13	58.722	0.955	7.710	0.000
37	Proofreading (GL&SCHR)	51.230	10.957	63.490	11.686	0.953	-7.530	0.000
36	Letter naming (GL&SCHR)	25.620	5.669	20.620	3.994	0.910	7.210	0.000
35	Digit naming (GL&SCHR)	23.745	5.123	19.280	3.635	0.899	7.080	0.000
34	Mental calculation division (TTR)	19.730	5.822	26.290	7.269	0.893	-7.440	0.000
33	Reversals accuracy (GL&SCHR)	15.670	2.322	17.755	1.862	0.889	-7.060	0.000
32	Text reading substantive errors (GL&SCHR)	14.460	8.003	7.810	5.187	0.886	6.730	0.000
31	Mental calculation addition (TTR)	30.500	3.403	33.810	3.410	0.875	-6.870	0.000
30	Morphology and syntax (GL&SCHR)	50.475	10.272	59.570	9.862	0.825	-6.870	0.000
29	Mental calculation multiplication (TTR)	21.740	5.022	26.780	6.187	0.818	-6.250	0.000
28	Word reading percentage error (EMT)	2.299	1.947	0.896	1.081	0.815	6.990	0.000
27	Pseudoword reading percentage error (KLEPEL)	10.689	6.649	5.922	4.763	0.763	5.280	0.000
26	Color naming (GL&SCHR)	32.450	5.821	28.250	4.314	0.760	5.970	0.000
25	English word reading percentage error (OMT)	5.551	3.789	3.074	2.707	0.705	5.170	0.000
24	Text reading time consuming errors (GL&SCHR)	13.620	7.373	9.170	4.907	0.671	5.240	0.000
23	Spoonerisms accuracy (GL&SCHR)	16.795	2.297	18.190	1.674	0.657	-4.080	0.000
22	Mental calculation subtraction (TTR)	27.550	3.658	30.080	3.860	0.639	-4.570	0.000
21	Vocabulary (GL&SCHR)	7.830	4.144	10.830	4.770	0.638	-4.480	0.000
20	Phonological STM (GL&SCHR)	20.130	4.683	23.230	4.561	0.637	-4.420	0.000
19	Speed of processing correct (CDT)	119.250	22.855	134.020	21.312	0.635	-4.260	0.000
18	Definitions (KAIT)	20.900	1.892	22.165	1.966	0.624	-4.360	0.000
17	Writing speed (GL&SCHR)	24.760	3.434	26.500	3.404	0.494	-3.980	0.000
16	Working memory (GL&SCHR)	39.460	4.606	41.575	4.181	0.469	-3.400	0.001
15	Text comprehension (GL&SCHR)	19.440	4.912	21.590	4.397	0.450	-3.610	0.001
14	Double meanings (KAIT)	14.440	3.914	16.100	3.710	0.426	-3.780	0.002
13	Delayed auditory comprehension (KAIT)	4.990	1.403	5.540	1.500	0.373	-2.770	0.008
12	Personalities (KAIT)	7.250	3.109	8.350	3.073	0.351	-2.170	0.013
11	Speed of processing percentage errors/missed (CDT)	1.837	1.235	1.435	1.045	0.347	2.880	0.014
10	Verbal STM (GL&SCHR)	35.540	5.386	37.355	5.003	0.345	-2.690	0.014
9	Automation (GL&SCHR)	10.642	17.455	5.706	11.532	0.330	2.360	0.019
8	Visual STM (GL&SCHR)	10.360	3.723	11.585	4.387	0.298	-2.290	0.035
7	Object naming (GL&SCHR)	39.515	7.056	37.785	6.639	0.251	1.860	0.076
6	Block patterns (KAIT) *	12.230	2.715	11.710	2.965	-0.183	1.293	0.197
5	Logical reasoning (KAIT)	11.350	3.415	11.810	3.093	0.141	-0.998	0.319
4	Secret codes (KAIT)	26.980	4.826	27.510	4.770	0.111	-0.781	0.436
3	Auditory comprehension (KAIT)	13.335	2.763	13.625	2.740	0.106	-0.745	0.457
2	Symbol learning (KAIT)	80.790	11.719	81.395	11.887	0.051	-0.362	0.717
1	Delayed symbol learning (KAIT)	51.210	9.790	51.690	9.565	0.050	-0.351	0.726

* Positive effect sizes mean the control group performed better than the group with dyslexia. Block patterns is the only variable where the group with dyslexia scores higher than the control group.

Supplementary materials 3.

Stage 2 EFA Factor Loadings, Uniqueness and Explained Variance of the 48 Variables Ordered from Large Effect Size to Small Effect Size.

		Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	Uniqueness
53	Word spelling (GL&SCHR)	-0.184	0.544	0.227	0.118	0.020	0.051	-0.112	0.033	-0.034	0.059	0.194
52	Word reading correct (EMT)	-0.372	-0.013	0.089	0.067	-0.083	-0.022	-0.097	0.034	0.030	0.652	0.054
51	Flash typing English words (IDAA)	-0.434	0.148	0.531	0.016	-0.020	0.001	-0.070	0.060	0.009	-0.048	0.170
50	Sentence spelling (AT-GSN)	0.216	-0.522	-0.200	-0.167	0.017	0.011	0.035	-0.013	-0.001	-0.051	0.186
49	Flash typing pseudowords (IDAA)	-0.435	0.086	0.376	0.027	0.030	0.154	0.086	-0.107	-0.097	0.053	0.364
48	Lexical decision (IDAA)	-0.255	0.261	0.510	-0.043	0.099	-0.017	0.158	0.035	-0.056	0.065	0.359
47	Pseudoword reading correct (KLEPEL)	-0.569	-0.058	-0.036	0.272	0.077	0.130	-0.157	-0.030	-0.004	0.317	0.212
46	Flash typing words (IDAA)	-0.349	0.090	0.614	0.010	-0.071	-0.003	0.030	-0.074	-0.021	0.034	0.235
45	English word spelling (WRAT)	-0.263	0.217	0.194	0.078	-0.064	0.062	-0.207	0.295	0.080	-0.005	0.285
44	English word reading correct (OMT)	-0.376	-0.243	0.150	0.069	-0.146	-0.028	-0.056	0.460	0.061	0.263	0.203
43	Text reading time (GL&SCHR)	0.586	0.059	0.002	0.037	0.066	0.055	0.077	-0.071	0.068	-0.470	0.269
42	Spoonerisms time (GL&SCHR)	0.382	-0.041	0.066	-0.428	0.101	-0.011	0.027	-0.107	-0.058	0.032	0.329
41	Reversals (IDAA)	-0.263	0.014	0.026	0.650	-0.041	0.034	0.245	-0.131	-0.104	-0.042	0.290
40	Reversals time (GL&SCHR)	0.456	-0.007	0.109	-0.218	0.184	-0.150	0.119	0.025	0.070	0.061	0.502
39	Mental calculation mix (TTR)	0.047	0.067	0.050	0.064	0.065	0.879	0.038	0.081	-0.119	0.003	0.137
38	Silent Reading (Tickbite)	0.457	-0.029	-0.036	0.190	0.067	0.004	-0.061	-0.183	0.024	-0.346	0.443
37	Proofreading (GL&SCHR)	-0.173	0.534	0.065	-0.123	-0.061	0.104	0.027	-0.081	0.081	0.103	0.471
36	Letter naming (GL&SCHR)	-0.040	-0.052	-0.005	-0.054	0.891	0.040	0.003	0.057	0.061	-0.012	0.219
35	Digit naming (GL&SCHR)	-0.003	0.043	-0.030	-0.079	0.950	0.036	0.021	0.074	-0.034	0.001	0.135
34	Mental calculation division (TTR)	0.064	0.208	-0.027	-0.082	0.062	0.892	-0.033	0.060	0.073	-0.016	0.184
33	Reversals accuracy (GL&SCHR)	-0.035	-0.013	-0.019	0.771	-0.102	-0.047	0.085	-0.073	-0.096	-0.047	0.464
32	Text reading substantive errors (GL&SCHR)	0.834	0.042	-0.031	-0.062	-0.008	0.157	-0.082	0.076	-0.061	0.012	0.401
31	Mental calculation addition (TTR)	-0.092	-0.088	0.057	-0.056	-0.047	0.800	0.110	-0.004	-0.009	0.044	0.220
30	Morphology and syntax (GL&SCHR)	-0.064	0.659	0.033	-0.046	-0.067	0.028	0.019	0.067	-0.013	-0.163	0.494
29	Mental calculation multiplication (TTR)	0.050	0.200	-0.172	0.029	0.032	0.909	-0.157	0.024	0.033	-0.009	0.194
28	Word reading percentage error (EMT)	0.738	-0.131	-0.081	0.090	-0.116	0.088	-0.026	0.109	-0.104	0.057	0.515
27	Pseudoword reading percentage error (KLEPEL)	0.616	-0.027	-0.045	-0.077	-0.174	-0.090	0.044	0.082	-0.078	-0.010	0.550
26	Color naming (GL&SCHR)	0.004	-0.130	-0.003	0.115	0.731	-0.049	0.066	0.080	-0.091	-0.041	0.390
25	English word reading percentage error (OMT)	0.000	-0.044	-0.186	-0.264	0.224	0.096	0.046	-0.454	-0.044	0.362	0.503
24	Text reading time consuming errors (GL&SCHR)	0.622	-0.001	0.017	0.161	0.027	-0.006	0.019	0.017	0.091	0.015	0.687
23	Spoonerisms accuracy (GL&SCHR)	-0.087	0.021	0.020	0.561	0.256	-0.036	-0.057	-0.061	0.167	0.207	0.535
22	Mental calculation subtraction (TTR)	-0.054	-0.193	0.076	-0.018	-0.043	0.837	0.222	0.018	-0.039	-0.103	0.237
21	Vocabulary (GL&SCHR)	0.007	0.165	0.031	-0.120	0.126	0.058	0.007	0.857	-0.093	-0.033	0.272
19	Speed of processing correct (CDT)	0.029	-0.186	0.168	0.022	-0.172	0.364	0.100	-0.089	0.068	0.165	0.641
18	Definitions (KAIT)	-0.094	0.227	-0.094	0.147	0.105	0.097	0.010	0.324	0.020	0.040	0.591
16	Working memory (GL&SCHR)	-0.168	0.041	-0.049	0.200	-0.106	0.124	0.321	-0.123	0.024	0.042	0.660
15	Text comprehension (GL&SCHR)	-0.005	0.076	0.117	-0.078	0.023	-0.137	0.286	0.470	-0.019	0.099	0.595
14	Double meanings (KAIT)	-0.063	0.197	-0.153	0.102	0.043	-0.029	0.215	0.390	0.001	-0.045	0.585
13	Delayed auditory comprehension (KAIT)	-0.103	-0.093	-0.099	0.035	0.059	-0.047	0.203	0.586	0.013	0.032	0.559
12	Personalities (KAIT)	0.149	-0.135	-0.042	0.033	-0.005	0.090	-0.142	0.815	0.030	0.050	0.441
10	Verbal STM (GL&SCHR)	0.148	0.193	-0.063	0.275	-0.064	-0.125	0.243	-0.079	-0.069	0.211	0.807
7	Object naming (GL&SCHR)	-0.121	-0.134	0.133	0.142	0.437	-0.012	-0.032	-0.036	-0.010	-0.148	0.732
6	Block patterns (KAIT)	0.144	-0.118	0.129	0.068	0.099	0.086	0.610	-0.017	0.113	-0.097	0.532
5	Logical reasoning (KAIT)	0.043	0.081	0.024	0.036	0.000	-0.091	0.594	0.043	0.078	-0.013	0.580
4	Secret codes (KAIT)	0.006	-0.077	-0.111	0.088	-0.014	0.185	0.516	0.056	0.032	-0.077	0.613
3	Auditory comprehension (KAIT)	0.108	-0.038	-0.064	-0.145	0.040	0.066	0.078	0.747	-0.021	-0.058	0.556
2	Symbol learning (KAIT)	0.036	0.011	0.007	-0.039	-0.034	0.051	0.224	-0.024	0.765	-0.033	0.273
1	Delayed symbol learning (KAIT)	-0.089	0.002	-0.053	-0.059	-0.012	-0.08	0.178	-0.057	0.950	0.024	0.063
	SS loadings	4,519	1,955	1,481	2,133	2,816	4,153	1,325	3,487	1,688	1,775	
	Proportion Variance	0,094	0,041	0,031	0,044	0,059	0,087	0,028	0,073	0,035	0,037	

Note: Loadings above the cut-off of .32 or below -.32 are marked in bold. Below the cut-off they are printed in gray. These were taken into account for interpretation of the factors.