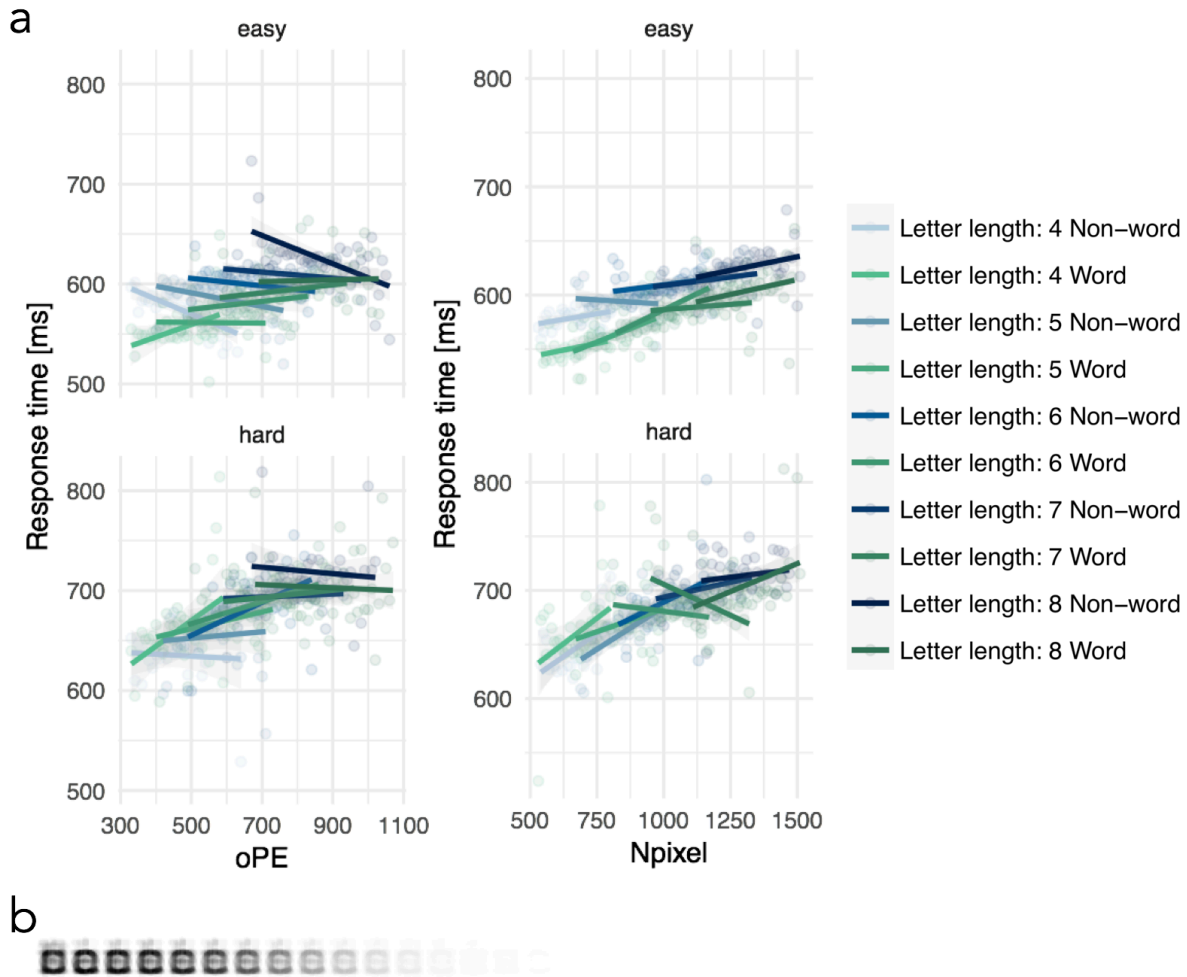


1388  
1389  
1390  
1391  
1392  
1393  
1394

## Supplementary Materials

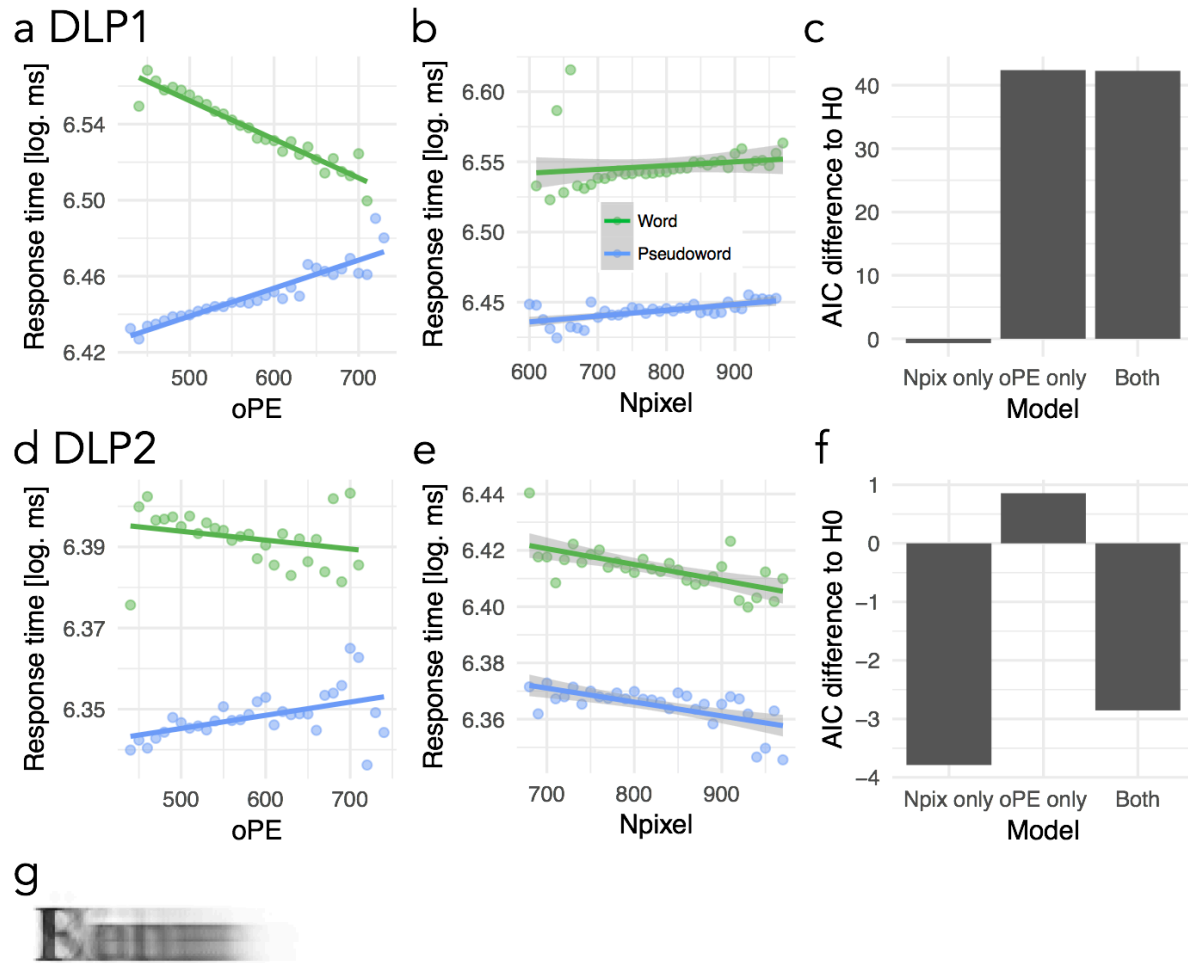
Fig. S1. Behavioral evaluation, including stimuli with 4-8 letters.  
Fig. S2. Dutch lexical decision behavior and prediction using a proportional script.  
Fig. S3. Detailed description of significant activation clusters in the EEG study  
Table S1. Results from linear mixed model regression analysis  
Table S2. Reliable activation clusters from the fMRI evaluation with respective anatomical labels



1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408

*Figure S1.* Behavioral evaluation including multiple word lengths. (a) Response times aggregated across participants from the British lexicon (BLP) project (Keuleers et al., 2012) for the word lengths 4-8. The left panel shows the word/non-word by orthographic prediction error (oPE) interaction and the right panel shows the word/non-word by number of pixels (Npixel) interaction for each word length separately. In addition, the upper panel shows letter strings that are correctly categorized in nearly all cases (accuracy > .95) and the lower panel shows the response times to the items, which were less accurately processed (i.e., accuracy < .95). The median split resulted in a subset of the BLP (i.e., the easy words) which are roughly comparable to words used in the previous experiments (e.g. see Fig. 3), as the BLP study includes a large number of very rare words (median log. word frequency per million is .3). Bluish colors represent non-words (N) and greenish colors represent words (W), while the hue of the colors reflects word length (i.e., bright to dark reflects short to long letter strings). For both effects, we first estimated linear regression models with either the oPE or the Npixel effect and allowing interactions with word/non-word

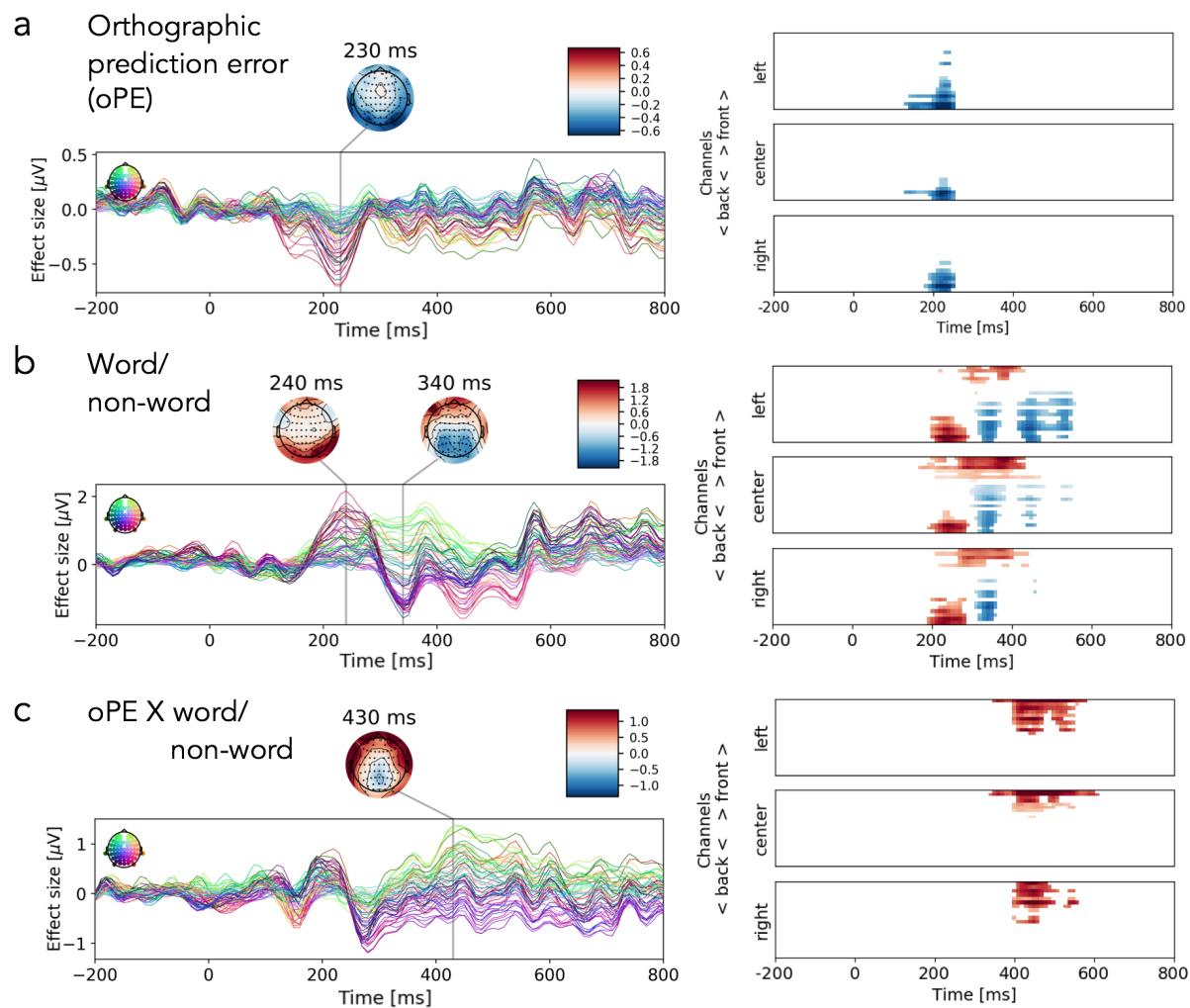
1409 status, word length, and accuracy. Note that the oPE in this first analysis was based on length-  
1410 specific predictions (i.e., for the estimation of the oPE of four-letter words, all four-letter words  
1411 of the lexicon were included in the prediction). For the oPE model, a significant four-way  
1412 interaction was found (estimate =  $-1.078e-04$ ; SE =  $4.199e-05$ ;  $t = -2.567$ ). Separating hard vs.  
1413 easy words allowed us to disentangle the four-way interaction: In easy words/non-words, we  
1414 found a consistent (i.e., across length levels) oPE by word/non-word interaction (estimate =  
1415  $1.530e-04$ ; SE =  $4.047e-05$ ;  $t = 3.780$ ) in the same direction as previously shown (positive effect  
1416 for words and a negative effect for non-words). For hard words/non-words, we found that the  
1417 oPE by word/non-word interaction was inconsistent across letter length levels, which was  
1418 indicated by a significant oPE and letter length interaction (estimate =  $-3.530e-05$ ; SE =  $8.092e-$   
1419  $06$ ;  $t = -4.363$ ). In addition, for the hard words both the oPE by word/non-word interaction  
1420 (estimate =  $-1.685e-04$ ; SE =  $6.905e-05$ ;  $t = -2.440$ ) and the main effect of oPE were reversed  
1421 (estimate =  $2.828e-04$ ; SE =  $5.802e-05$ ;  $t = 4.874$  compare to estimate =  $-1.000e-04$ ; SE =  $2.440e-$   
1422  $05$ ;  $t = -4.101$ , for easy words). For the Npixel model, no four-way interaction and no Npixel  
1423 interaction or main effect were found. In sum, in this analysis we showed that the oPE by  
1424 word/non-word interaction shown previously for word lengths of five letters (see main text) is  
1425 consistent for easy-to-process English items with word lengths from 4-8 letters. Secondly, the  
1426 word/non-word by orthographic prediction error interaction was also reliable when the prediction  
1427 included all words of all letter lengths from the English lexicon (see part b of this Figure) and the  
1428 orthographic prediction error estimation was based on this length-unspecific prediction (estimate:  
1429  $0.02$ ; SE= $0.007$ ;  $t=3.349$ ). (b) Letter-length unspecific prediction for English based on ~60,000  
1430 English words from the SUBTLEX database (Heuven et al., 2014).  
1431



1432

1433 *Figure S2.* Dutch lexical decision behavior and prediction using a proportional script. (a) Effect  
 1434 of the orthographic prediction error parameter, (b) number of pixels parameter and (c) showing  
 1435 the same model comparisons as implemented in Figure 3 for the data from the first Dutch lexicon  
 1436 project (DLP1; (Keuleers, Diependaele, & Brysbaert, 2010); 4,305 five-letter stimuli; 39  
 1437 participants) and the same effects and model comparisons for the second Dutch lexicon project  
 1438 (DLP2; (Brysbaert, Stevens, Mandera, & Keuleers, 2016); 3,145 five-letter stimuli; 81  
 1439 participants) are presented in (d,e,f). Before going into the details of the two studies one has to  
 1440 note that the patterns we have found in the data in relation to our parameters of interest do not  
 1441 replicate within these two Dutch studies and, in addition, do not replicate with the findings from  
 1442 German, English, and French shown in Figure 3. In general, this is difficult for the interpretations  
 1443 of the results. For the DLP1 pattern we found a significant interaction of the orthographic  
 1444 prediction error with word/non-words and no significant effect of number of pixels. The  
 1445 interaction pattern in contrast to the findings in other languages (Fig. 3a), however, was  
 1446 qualitatively different as it showed a negative orthographic prediction error effect for words and a  
 1447 positive effect for non-words. The pattern is exactly the inverse from all other languages. Still  
 1448 model comparisons highlighted that the orthographic prediction error was relevant for the model  
 1449 fit since the predictor increased the model fit with no further increase of fit when the number of  
 1450 pixel parameter was included. None of these findings could be replicated in the DLP2 dataset,  
 1451 showing no significant fixed effects or interactions and no substantial changes in model fit

1452 relation to the null model. (g) Prediction image from a PEMoR implementation using five-letter  
1453 words with a proportional Times New Roman script.  
1454



1455

1456 *Figure S3.* Detailed description of significant activation clusters in the EEG study for (a) the  
 1457 orthographic prediction error; (b) word/non-word effect; (c) interaction of word/non-word and the  
 1458 orthographic prediction error. On the left, the effect sizes from regression ERPs are presented as  
 1459 time courses for each sensor and time-point (color coding reflects scalp position). This part of the  
 1460 Figure reproduces Figure 5. The right column displays time courses with one line per channel,  
 1461 masked by significance using cluster statistics (see Methods for details; Maris & Oostenveld,  
 1462 2007).  
 1463

*Table S1.* Results from linear mixed model regression analysis (with the exception of the British data including multiple word lengths was estimated based on word aggregated data) for the behavioral lexical decision tasks (LDT) and handwriting analyses.

	E	SE	<i>t</i>
German LDT N°1: Orthographic prediction error based on word length specific prediction			
Intercept	6.49	0.023	288
Orthographic prediction error (oPE)	-0.03	0.004	6.5
Number of pixels (Npixel)	-0.007	0.004	1.8
Word/non-word (Lex)	0.33	0.009	33.1
Word frequency	-0.12	0.004	33.5
Error	-0.03	0.005	6.2
oPE X Lex	0.03	0.006	5.0
Npixel X Lex	0.000	0.006	0.1
German LDT N°1: Orthographic prediction error based on word length general prediction			
Intercept	6.48	0.023	288.3
Orthographic prediction error (oPE)	-0.03	0.004	6.3
Number of pixels (Npixel)	-0.01	0.004	1.7
Word/non-word (Lex)	0.33	0.010	33.2
Word frequency	-0.12	0.004	35.5
Error	-0.03	0.005	6.2
oPE X Lex	0.03	0.006	4.5
Npixel X Lex	-0.00	0.006	0.0
German LDT N°1: Orthographic prediction error based on word length specific prediction including orthographic Levenshtein distance and word frequency			
Intercept	6.66	0.023	237.1
Orthographic prediction error (oPE)	-0.02	0.004	4.3
Number of pixel (Npixel)	-0.00	0.004	0.2
Word/non-word (Lex)	0.29	0.011	27.0
Error	-0.03	0.005	6.2
Orthographic Levenshtein distance	-0.08	0.008	10.5
Word frequency	-0.12	0.004	35.5
oPE X Lex	0.03	0.006	5.2
Npixel X Lex	-0.00	0.005	0.6
German LDT N°2 including noise: 0%			
Intercept	6.32	0.024	263.9
Orthographic prediction error (oPE)	-0.02	0.016	1.4
Number of pixels (Npixel)	-0.00	0.015	0.2
Word/non-word (Lex)	0.27	0.05	5.4

Word frequency	-0.07	0.02	4.9
Error	-0.07	0.010	6.8
oPE X Lex	0.05	0.02	2.3
Npixel X Lex	-0.02	0.021	1.2

German LDT N°2 including noise: 20%

Intercept	6.45	0.026	245.4
Orthographic prediction error (oPE)	-0.06	0.017	3.3
Number of pixels (Npixel)	-0.00	0.013	0.3
Word/non-word (Lex)	0.37	0.049	7.5
Word frequency	-0.14	0.02	6.1
Error	-0.14	0.010	5.4
oPE X Lex	0.04	0.022	1.6
Npixel X Lex	0.02	0.022	0.7

German LDT N°2 including noise: 40%

Intercept	6.84	0.042	162.9
Orthographic prediction error (oPE)	-0.02	0.021	1.0
Number of pixels (Npixel)	-0.08	0.018	4.1
Word/non-word (Lex)	0.14	0.049	2.8
Word frequency	-0.11	0.06	1.9
Error	-0.00	0.010	0.1
oPE X Lex	-0.00	0.028	0.1
Npixel X Lex	0.08	0.026	2.9

British LDT

Intercept	6.39	0.013	507.1
Orthographic prediction error (oPE)	-0.007	0.001	5.3
Number of pixels (Npixel)	0.008	0.001	6.7
Word/non-word (Lex)	0.12	0.003	46.2
Word frequency	-0.067	0.001	58.0
oPE X Lex	0.008	0.002	4.2
Npixel X Lex	-0.003	0.002	1.9

British LDT 4-8 Letters: Length specific prediction

Intercept	6.26	0.157	39.7
Orthographic prediction error (oPE)	-0.001	0.000	5.0
Number of letters (Nletters)	0.062	0.027	2.3
Word/non-word (Lex)	0.155	0.162	0.3
Error	0.043	0.165	0.8
oPE X Lex	-0.001	0.000	4.5
oPE X Nletters	-0.001	0.000	3.3
oPE X Error	-0.002	0.000	5.1
Nletters X Lex	-0.006	0.028	0.8
Nletters X Error	-0.245	0.172	1.4

Lex X Error	-0.036	0.028	1.3
oPE X Lex X Nletters	0.001	0.000	2.4
oPE X Lex X Error	0.002	0.000	5.0
oPE X Nletters X Error	0.001	0.000	3.2
Nletters X Lex X Error	0.003	0.030	0.1
oPE X Lex X Nletters X Error	-0.001	0.000	2.6

British LDT 4-8 Letters: Length general prediction

Intercept	5.25	0.421	12.5
Orthographic prediction error (oPE)	0.002	0.000	3.7
Number of letters (Nletters)	0.250	0.061	4.1
Word/non-word (Lex)	1.064	0.438	2.4
Error	1.264	0.443	2.9
oPE X Lex	-0.002	0.001	3.1
oPE X Nletters	-0.000	0.000	3.6
oPE X Error	-0.002	0.001	4.0
Nletters X Lex	-0.183	0.065	2.9
Nletters X Error	-0.002	0.001	4.0
Lex X Error	-1.426	0.467	3.1
oPE X Lex X Nletters	0.001	0.000	2.9
oPE X Lex X Error	0.002	0.001	3.6
oPE X Nletters X Error	0.001	0.000	4.0
Nletters X Lex X Error	0.228	0.068	3.5
oPE X Lex X Nletters X Error	-0.001	0.000	3.3

British LDT 4-8 Letters: Number of pixel

Intercept	6.590	0.157	42.0
-----------	-------	-------	------



Number of pixel (Npixel)	0.000	0.001	0.3
Number of letters (Nletters)	0.092	0.028	3.2
Word/non-word (Lex)	-0.124	0.162	0.8
Error	-0.309	0.165	1.9
Npixel X Lex	0.000	0.001	0.2
Npixel X Nletters	0.000	0.001	1.4
Npixel X Error	0.000	0.001	0.4
Nletters X Lex	-0.059	0.029	2.0
Nletters X Error	-0.090	0.030	3.0
Lex X Error	0.035	0.171	0.2
Npixel X Lex X Nletters	0.000	0.001	0.9
Npixel X Lex X Error	0.000	0.001	0.1
Npixel X Nletters X Error	0.000	0.001	1.2
Nletters X Lex X Error	0.069	0.031	2.2
Npixel X Lex X Nletters X Error	0.000	0.001	1.2

#### French LDT

Intercept	6.63	0.005	1,333
Orthographic prediction error (oPE)	-0.002	0.001	2.0
Number of pixels (Npixel)	0.002	0.001	1.3
Word/non-word (Lex)	-0.040	0.003	11.6
Word frequency	-0.042	0.001	34.1
oPE X Lex	0.005	0.002	2.0
Npixel X Lex	-0.007	0.002	3.0

#### Dutch LDT

Intercept	6.45	0.019	348.1
Orthographic prediction error (oPE)	0.005	0.002	3.2
Number of pixels (Npixel)	0.001	0.002	0.6
Word/non-word (Lex)	0.101	0.004	23.8
Word frequency	-0.061	0.002	36.9
oPE X Lex	-0.016	0.002	6.6
Npixel X Lex	0.002	0.002	1.0

#### Dutch LDT2

Intercept	6.35	0.016	391.1 <sup>1464</sup>
Orthographic prediction error (oPE)	0.002	0.002	1.1
Number of pixels (Npixel)	-0.001	0.002	0.6
Word/non-word (Lex)	0.048	0.005	9.4
Word frequency	-0.023	0.001	26.9
oPE X Lex	-0.003	0.003	1.3
Npixel X Lex	0.003	0.003	0.5

Handwriting: Script based orthographic prediction error

Intercept	1.465	0.010	154.3
Mean prediction strength	0.052	0.007	7.4
Number of pixels with a prediction	0.015	0.008	2.1
Letter case	0.039	0.012	3.2

Handwriting: Readability ratings

Intercept	11.5	1.4	8.1
Mean prediction strength	-5.9	1.0	6.2

*Note.* E: Estimate; SE: Standard error; *t*: t-value. All *t*'s >2 are considered a significant effect.