Supplementary Materials

1. Material and Methods:

Our stimuli set comprised 24 exemplars for each of the five visual categories presented in the study: faces, houses, tools, strings and false-fonts. Examples of each category are showed in figure S1.

All stimuli were black on a white background. Faces, houses and tools were highly contrasted gray-level photographs matched for size and overall luminance. Faces were frontal or slightly lateral views of non-famous people. Tools were common hand-held household objects (e.g. knife, hairdryer) presented in normal orientation. Houses comprised exterior pictures of common houses, apartments and buildings. Letter strings were pseudo-words, four letters long, made of the lower-case letters ‘bdmnlpqiou’, and were orthographically legal in Portuguese. Other letters could not be used, because when presented in mirror image they would not have formed a possible pseudoword. Each of the 24 pseudowords comprised one and only one asymmetric letter (b, d, p or q) per stimuli, the other three being symmetric. A false font was designed in which the pseudo-letters were matched one-to-one in stroke complexity with the letters used in the pseudo-word stimuli. The 24 false font strings were generated by one-to-one replacement of the letters in the pseudo-word strings with the false-font letters. Strings and false-fonts were ~130 pixel wide x 40 pixel high or 4 x 1.2 degrees. The pseudo-words used in strings category were all mirror-reversible pseudo-words such as “obli/ildo”, using a slightly modified font where letters were strictly reversible, so that even in mirror image the stimuli still looked like possible pseudo-words.

Stimuli were presented sequentially as pairs which could be physically identical, left-right mirrored, or different. Subjects judged if the pairs were same or different, pressing the right or left button respectively.

In such a same-different task, the choice of “different” trials can have a considerable impact on strategy and performance. Here, to facilitate the task, we tried to maximize the difference between the exemplars presented on “different” trials.

In the face category, which comprised 13 male and 11 female faces, the “different” trials were formed by pairing each face with another very different one, thus forming 24 easily discriminable “different” pairs (13 of which had the same sex). 10 of these pairs were randomly chosen to be presented to a given subject as the “different” trials. The rationale for using only a subset of 10 pairs was to have the same overall stimulus set as the one used in our fMRI and ERPs studies, while restricting the duration of the present behavioral experiment for practical reasons (all studies were performed in the same day). The same information applies to the tools and house categories.

In the string category, each of the 24 pseudowords was paired with another one to create the “different” pairs. We tried to maximize letter dissimilarity between the two members of each pair, but it was impossible to fully exclude pairs of pseudowords sharing the same letters. In the end, the pairs exhibited quite a bit of variability in similarity on “different” trials, ranging from pairs such as “inpo / mubi” to others such as “obli / iqno” or “idum / oubi”. Quantitatively, the “different” pairs shared, on average, only 0.21 letters at the same location (out of 4 letters). Even after mirroring one of the two stimuli (as needed for the task, which requires responding “same” to mirror pairs such oubi/iduo), the “different” pairs still only shared 1.08 letter at the same location. Finally, without considering location, the “different” pairs shared 1.83 letters on average. We never used as “different” stimuli two pseudowords that shared the same exact letter content but different only in letter order (e.g. oubi/ibou). Thus, it was always possible to respond “different” by spotting that at least one of the letters differed between the two pseudowords. 75% of the pairs also differed in the number of ascender (bdl), descender (pq) versus midline letters (mniou), thus creating a noticeable difference in overall contour of the two pseudowords.

Finally, the false-font stimuli were matched one-to-one with the pseudowords and hence were controlled in exactly the same manner.

Importantly, the stimulus set was the same for all three groups, and therefore in principle the difficulty induced by the choice of “different” trials was the same in all groups.

2. Supplementary Results:

2.1 Error rates (ERs)

An ANOVA on error rates revealed main effects of group ($F(2,56) = 34.5$; $p < 0.0001$), category ($F(2,112) = 176.4$; $p < 0.0001$) and condition ($F(2,112) = 25.3$; $p < 0.0001$). Significant condition-by-group ($F(4,112) = 5.4$; $p < 0.001$) and category-by-group ($F(4,112) = 3.6$; $p < 0.01$) interactions were also noticed (see figure S2). Consistent with RT data, illiterates did not show more errors, irrespective of stimulus category, in mirror relative to identical trials (strings: $F(1,9) = 1.6$, $p > .2$; false-fonts: $F(1,9) = 1.6$, $p > .2$; pictures: $F(1,9) < 1$). In contrast, ex-illiterates made more errors in mirror relative to identical trials, for strings ($F(1,20) = 22.6$, $p < .0005$) and false-fonts ($F(1,20) = 22.7$, $p < .0005$) but not for pictures ($F(1,20) = 1.46$, $p > .2$). The same was true in the literate group for strings ($F(1,27) = 54.3$, $p < .0001$), false-fonts ($F(1,27) = 58.6$, $p < .0001$) and pictures ($F(1,27) = 10.8$, $p < .003$), even though error rates were very low in this category (3.1% for mirror and 0.8% for identical). We then computed the mirror cost index for error rates. For strings, literates presented a higher mirror cost than illiterates ($F(1,36) = 11.0$, $p = .002$) but ex-illiterates did not differ from illiterates ($F(1,29) = 1.3$, $p > .2$). Concerning false-fonts, literates again presented higher mirror cost than illiterates ($F(1,36) = 5.5$, $p < .03$), and ex-illiterates vs. illiterates were marginally distinct ($F(1,29) = 4.3$, $p = .05$). Finally, for pictures, neither ex-illiterates nor literates differed from illiterates ($F(1,29) < 1$; $F(1,36) = 3.0$, $p > .09$; respectively).

2.2 Signal Detection Theory (SDT) analysis

SDT matrix definitions and d' results are reported in the main text. See also table 2 for hit rates and false-alarms separately for each of the 5 original categories.

Bias

We calculated individual bias values [-0.5 * (Zscore hits + Zscore FAs)] for the critical comparison (mirror vs different trials; SDT matrix defined earlier), with positive bias reflecting an overall tendency to answer “same” and negative bias a tendency to respond “different”. A 3 x 3 ANOVA revealed main effects of category ($p<0.0001$; mean bias: pictures=0.15; strings= -0.32; false-fonts=0.05) and group ($p=0.005$; illiterates=0.34; ex-illiterates=0.12; literates= - 0.13) but no group x category interaction ($F=1$) (see figure S3).

When restricting the analysis to strings, we observed a significant group effect ($p<0.03$): illiterates were non-significantly biased to answer “same” (bias=0.22, t=0.73, p=0.49) whereas literates showed a significant bias to answer “different” (bias= -0.63, t= -5.3, p<0.0001). Finally, ex-illiterates presented an intermediate pattern, i.e. a non-significant disposition to answer “different” (bias= -0.18, t= -0.74, p=0.47). This important observation suggests that literacy, especially when acquired in childhood, enhances the tendency to judge mirror-symmetrical strings such as “ildo” and “obli” as entirely different images. Comparing groups within the string category, literates were marginally more biased to respond “different” than ex-illiterates (p<0.08) but clearly more than illiterates (p<0.005), whereas illiterates and ex-illiterates did not differ (p>0.25). For false-fonts, no significant group difference was found (p>0.30; illiterates: 0.37; ex-illiterates: 0.11; literates: -0.09). For pictures a significant group effect was observed (p<0.002): whereas literates were unbiased (bias=0.01, t =0.7, p=0.5), illiterates presented a bias to answer “same” (bias =0.38, t =2.9, p <0.01), and ex-illiterates too (bias=0.22, t =4.1, p < 0.001). When investigating category effects in each group, while literates presented no important bias for all categories except strings (p<0.0001), illiterates showed a general tendency to answer “same” that did not vary with category (p>0.8) and ex-illiterates exhibited an intermediary pattern, with the same propensity as illiterates to answer “same” for pictures but a literate bias pattern for strings (p<0.03).

Finally, we tested the correlation of bias with the reading scores of the participants, separately for each visual category. For pictures, we found a negative correlation (r = -0.31; p<0.0001), mainly due to the tendency of illiterates to answer “same”, as explained before. For strings, we also found a negative correlation (r = -0.38; p < 0.003), due to the bias of literates to answer “different” and of illiterates to answer “same”. This effect was not significant for the unfamiliar false-fonts stimuli (r = -0.21; p>0.10).

2.3 Speed-Accuracy trade-off analysis

No speed-accuracy trade-off was found (i.e., no negative correlation between RTs and Error Rates; see figure S4). Indeed there was even a positive correlation (r = 0.42; p < 0.001) meaning that the subjects making less errors were also the fastest (essentially literates).

3. Additional analysis restricted to matched groups

We provide here additional results obtained when restricting the analysis to a subset of 31 Brazilian subjects (called ILB, EXB and LB2 in Dehaene et al., 2010b) matched for origin, age, and socio-economic status. We closely follow the order in which these analyses are reported in the main text for the entire group of subject.

3.1 Response Times (RTs)

As previously, we first focused on 'mirror' vs. 'same' trials and then analysed 'different' trials separately.

The first ANOVA on the natural log-transformed RTs showed main effects of category ($F(2,56) = 25.9$; $p < 0.0001$), condition ($F(1,28) = 34.3$; $p< 0.0001$) and a marginal group effect ($F(2,28) = 2.7$; $p = 0.087$). More importantly, these main effects were qualified by the only significant interaction: group x condition ($F(2,28) = 6.9$; $p < 0.005$) (see figure S5A). To understand the crucial group x condition interaction, we restricted the analysis to each group searching for condition effects (i.e., significant additional cost of assigning the "same" response to mirror pairs relative to identical pairs). The literate group exhibited a significant delay for responding "same" to mirror relative to identical trials ($F(1,10) = 28.1$; $p < 0.005$), with a 0.19 log-RT difference for pictures, 0.32 for false-fonts and also 0.32 for strings ($p < 0.01$ for each category) (see figure S5A). Ex-illiterates also presented an additional cost ($F(1,9) = 33.0$; $p < 0.001$) for all categories, i.e., 0.19 for pictures, 0.26 for false-fonts and 0.31 for strings stimuli ($p < 0.02$ for each). In clear contrast, illiterates did not exhibit any additional cost ($F(1,9) = 1.2$; $p = 0.3$) for strings, false-fonts or pictures ($p > 0.3$ for each).

We then tested directly for group differences in mirror cost for each category, using the normalized index: ((logRTmirror - logRTsame) / (logRTmirror + logRTsame)). We observed main effects of group ($F(2,28) = 6.3$, $p < 0.005$) but no effect of category ($F(2,56) < 1$) nor category-by-group interaction ($F(4,56) < 1$). Nevertheless, the literate group showed a greater mirror cost index than illiterates in all categories (paired t-test, uncorrected for multiple comparisons; for pictures: $p = 0.04$; for false-fonts: $p = 0.01$ and for strings: $p = 0.05$). Interestingly, ex-illiterates also showed a greater mirror cost than illiterates for pictures ($p = 0.05$), false-fonts ($p = 0.02$) and marginally for strings ($p = 0.08$) . Overall, the results suggest again that learning to read, whether early or late, reduces the efficiency with which we judge two mirror images as "same".

We then analysed the log-RTs for “different” trials, by declaring category and literacy-groups as factors. The ANOVA showed no significant main effects or interactions (see Figure S5B).

3.2 Regressions with reading fluency scores

The reduced variability in reading fluencies within this restricted subgroup of participants render the regression approach rather insensitive although a similar pattern as previously observed with the full set of subjects can be noticed graphically (see Figure S6). For pictures, the mirror cost was not significantly correlated with reading fluency ($r^2 = 3.3\%$; $p = 0.33$). Such was also the case for strings ($r^2 = 4.3\%$; $p = 0.26$). Only the mirror cost for false-fonts was positively correlated with reading fluency ($r^2 = 22.1\%$; $p = 0.01$).

3.3 D-primes:

As previously done with the full set of subjects we performed an ANOVA on d-primes, by declaring literacy-group (illiterates, ex-illiterates, literates) restricted to the matched subjects as between-subject factor and visual category (pictures, strings, false-fonts) and d’ type (“same vs. different”, “mirror vs. different”) as within-subject factor. The results revealed main effects of group ($F(2,28) = 15.7$; $p < 0.0001$), category ($F(2,56) = 105.8$; $p < 0.0001$) and d’ type ($F(1,28) = 67.4$; $p < 0.0001$). These main effects were qualified by the following interactions: category by group ($F(4,56) = 5.7$; $p < 0.001$), d’ type by category ($F(2,56) = 38.9$; $p < 0.0001$), d’ type by group ($F(2,28) = 7.9$; $p = 0.002$) and the triple category x group x d’ type interaction ($F(4,56) = 4.2$, $p < 0.005$) (see figure S7).

The main effect of d’ type shows that it was easier for subjects to distinguish “same vs different” trials than “mirror vs different” (respectively: d’ = 1.8 and 0.9). Pairwise group comparisons (t-tests, Holm-Bonferroni corrected for multiple comparisons) revealed better d’ for literates (d’=1.9) than for illiterates (d’ = 0.6; $p < 0.0001$). Ex-illiterates (d’ = 1.7) were also significantly better than illiterates ($p < 0.0005$), but no difference was found between literates and ex-illiterates ($p > 0.4$), suggesting an overall improvement of visual decisions with literacy. The main effect of category demonstrated that it was easier to judge pictures (d’ = 2.2) than strings (d’ = 1.1; $p < 0.0001$) or false-fonts (d’= 0.8; $p < 0.0001$) but no difference between the two latter categories was found ($p > 0.2$).

We then turned to the crucial part of our analysis: the difference between the two d' types (indexing the level of mirror invariance) and the influence of literacy on it (i.e., d' type X group interaction). First, as previously reported in the main text, illiterates performed better than the chance level and they showed no significant difference between the two d' types, neither at a global level nor for each of the visual categories (see main text). These results indicate an invariant mirror representation in illiterates (i.e., "mirror" images were not treated significantly differently from "same" images) for all visual categories. In a clear contrast, literates showed much higher d' for "same vs different" (d' = 2.4) than for "mirror vs different" (d'= 1.3; $p < 0.0001$; see figure S7), reflecting a lower mirror invariance level in this group. This difference in literates d' was also present for each of the three categories ($p < 0.05$ for pictures; $p < 0.001$ for strings and $p < 0.0001$ for false-fonts). Additionally, ex-illiterates also presented a higher d' for "same vs different" than for "mirror vs different", at a global level (respectively d'= 2.4 and d'= 1.0; $p < 0.0001$) and separately for strings and false-fonts ($p < 0.0005$ for each), although not significantly for pictures ($p = 0.13$). The between-group comparisons showed that literates exhibited higher d' differences (i.e., reduced mirror invariance) compared to illiterates for strings and false-fonts ($p < 0.001$ each) but only marginally for pictures ($p > 0.3$). Ex-illiterates also showed reduced mirror invariance compared to illiterates for strings and false-fonts ($p < 0.001$ each) but not for pictures ($p > 0.3$). Finally, literates presented reduced mirror invariance compared to ex-illiterates for all categories ($p < 0.0001$for strings and false-fonts and $p = 0.01$ for pictures).

These results suggest again that the reduction in mirror invariance with literacy can be observed even when literacy is acquired late in life for visual objects related to reading (letter strings) or physically similar stimuli (false-fonts). However the transfer of this mirror discrimination to visual categories outside the reading domain (pictures) may depend on early literacy acquisition.

3.4 Error rates (ERs)

Consistent with RT data, illiterates did not show more errors, irrespective of stimulus category, in mirror relative to identical trials (strings: $F(1,9)=1.6, p>0.2$; false-fonts: $F(1,9)=1.6, p>0.2$; pictures: $F(1,9) <1$). In contrast, ex-illiterates made more errors in mirror relative to identical trials, for strings ($F(1,9)=38.0, p<0.0005$) and false-fonts ($F(1,9)=44.2, p<0.0001$) but not for pictures ($F(1,9)= 2.6, p>0.13$). The same was true in literate group for strings($F(1,10)=54.5, p<0.0001$), false-fonts ($F(1,10)=24.4, p<0.001$) and pictures ($F(1,27)=5.4, p<0.05$) even though error rates were very low in this category (4.2% for mirror and 1.5% for identical). We then computed the mirror cost index for error rates. For strings, literates presented a higher mirror cost than illiterates ($F(1,19)=6.2, p=0.02$) but ex-illiterates did not differ from illiterates ($F(1,18)=2.1, p>0.16$). Concerning false-fonts, literates did not present a higher mirror cost than illiterates ($F(1,19)=1.2$),$p>0.25$), but ex-illiterates were slightly worse than illiterates ($F(1,18)=4.5, p<0.05$). Finally, for pictures, neither ex-illiterates nor literates differed from illiterates ($F(1,18)=1.7, p>0.2$; $F(1,19)=2.3, p>0.14$; respectively).

In conclusion, the supplementary data shows that (1) as literacy increased there was an increasing bias to judge mirror-symmetrical images such as ‘ildo’ and ‘obli’ as different images; (2) there was no speed-accuracy trade-off and (3) our results still held when the analysis were restricted to matched groups in terms of origin, age and socio-economic status. These complementary findings reinforce the hypothesis that learning to read breaks the mirror invariance of the visual system.

Figure S1. Examples of stimuli used for each visual category.

(A) faces; (B) houses; (C) Tools; (D) strings; (E) false-fonts.

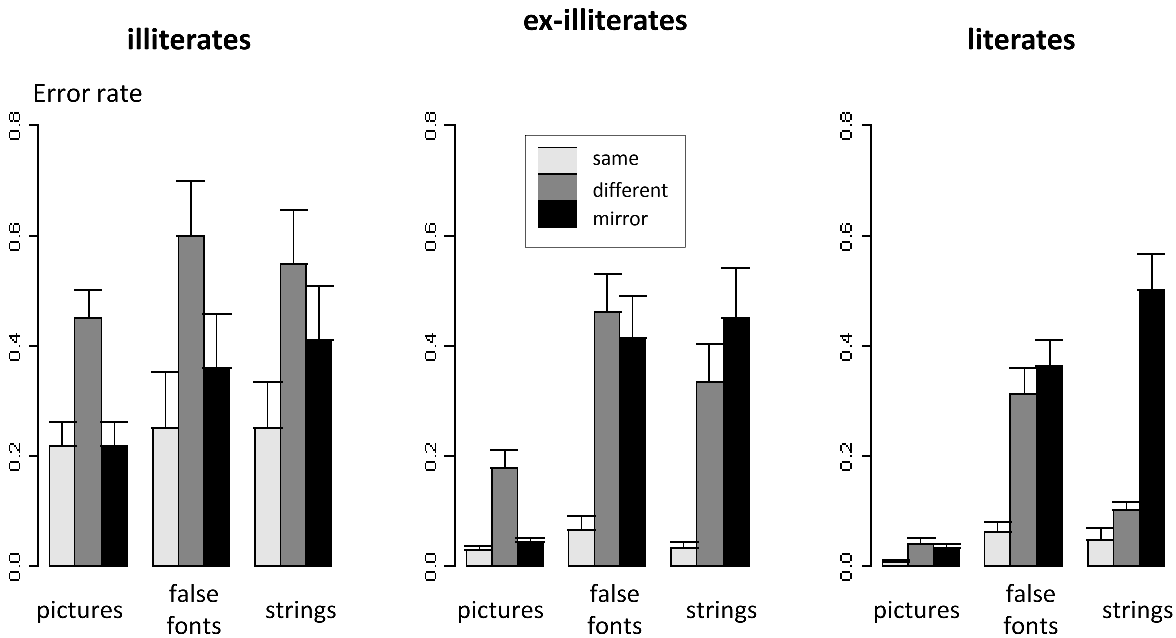

Figure S2. Error rates in the same-different task. Error bars represent one standard error of the mean.

Bias

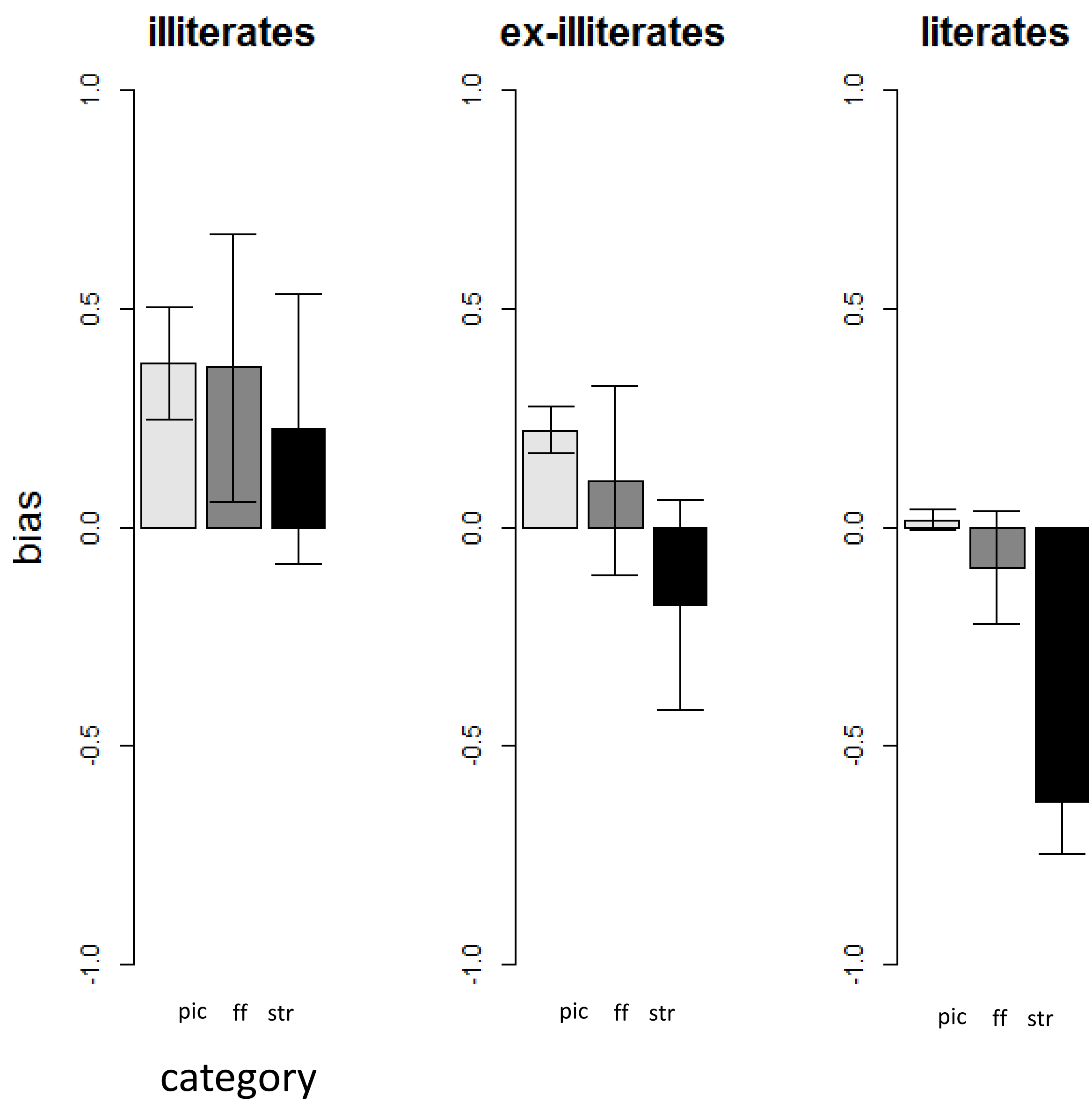

Figure S3: Bias. Bias values [-0.5 * (Zscore hits + Zscore FAs)] for « mirror vs. different » trials (hits = different trials answered “different”; false-alarms (FAs) = mirror trials answered “different”) are plotted for each literacy group and category. Positive bias reflecting an overall tendency to answer “same” and negative bias a tendency to respond “different” . pic = pictures; ff = false-fonts; str = strings. Error bars represent one standard error of the mean.

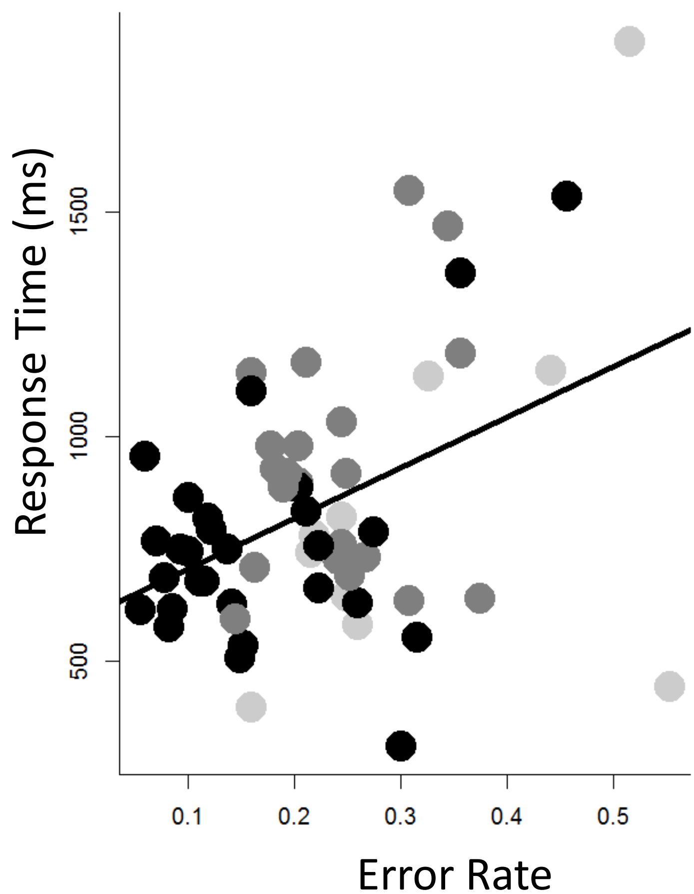

Figure S4. Relationship between Response Times (ms) and Error Rates. No speed-accuracy trade-off (i.e., negative correlation) was found. Indeed there was even a positive correlation ($r = 0.42$; $p < 0.001$) meaning that the subjects making less errors were also the fastest (essentially literates). Literate subjects are plotted in black, ex-illiterates in dark grey and illiterates in light grey.

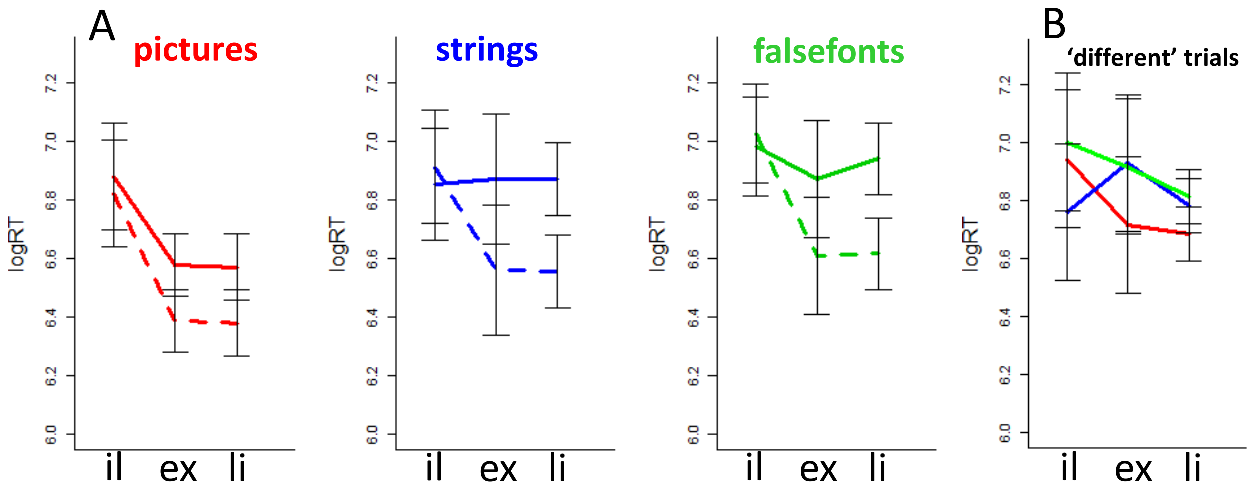

Figure S5: Log-transformed Response-Times (log-RTs) restricted to matched-groups. Response times were transformed according to a natural logarithm to correct for group differences in variance. (A) 'Same' trials (dashed lines) and 'mirror' trials (normal line) are plotted for each literacy-group and category . (B) 'Different' trials are plotted for each literacy-group and category. il = illiterates; ex = ex-illiterates; li = literates. The same color code for categories was used in (B). Error bars represent 95% within-subject Confidence Intervals (see equation #3 in Masson & Loftus, 2003).

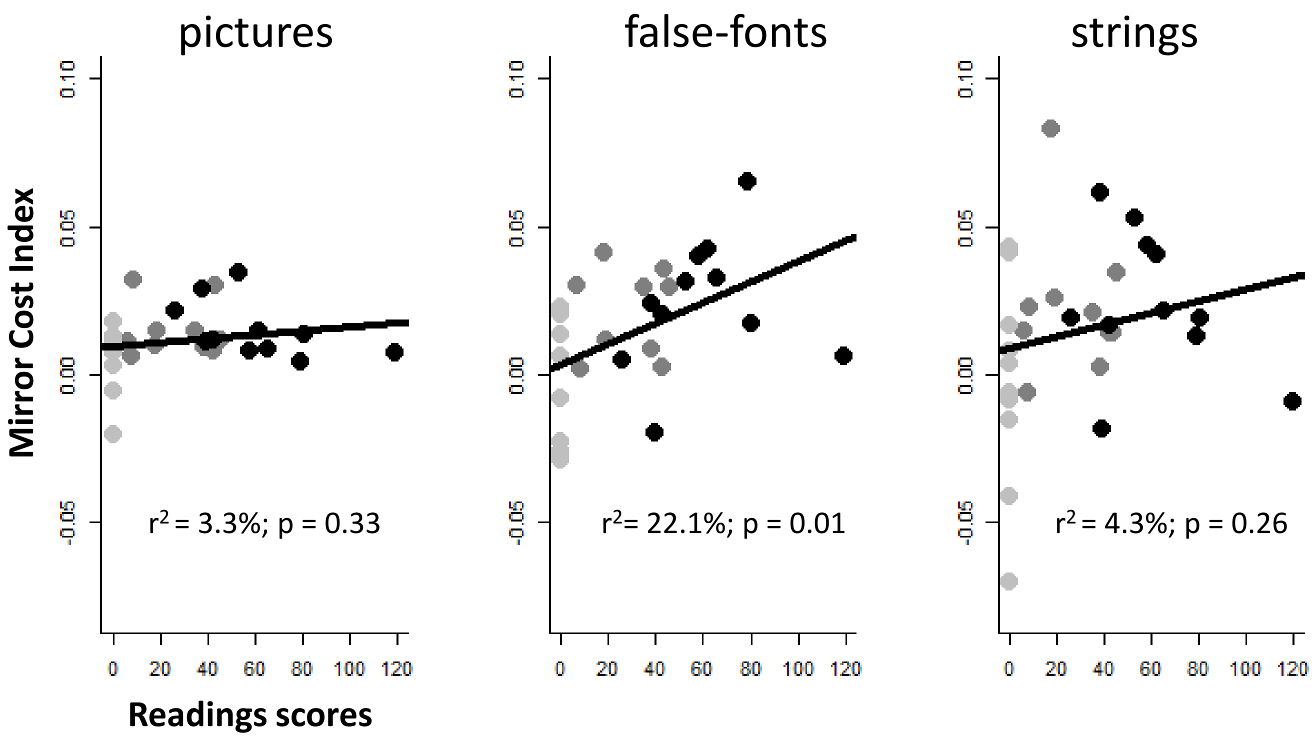

Figure S6. **Regressions restricted to matched-groups**. Correlations between the response time mirror cost index [i.e., ($logRT_{mirror}$ - $logRT_{same}$) / ($logRT_{mirror}$ + $logRT_{same}$)] and reading performance (average of words and pseudowords read per minute) restricted to subjects of matched-groups for each of the three categories of stimuli are presented. Literate subjects are plotted in black, ex-illiterates in dark grey and illiterates in light grey.

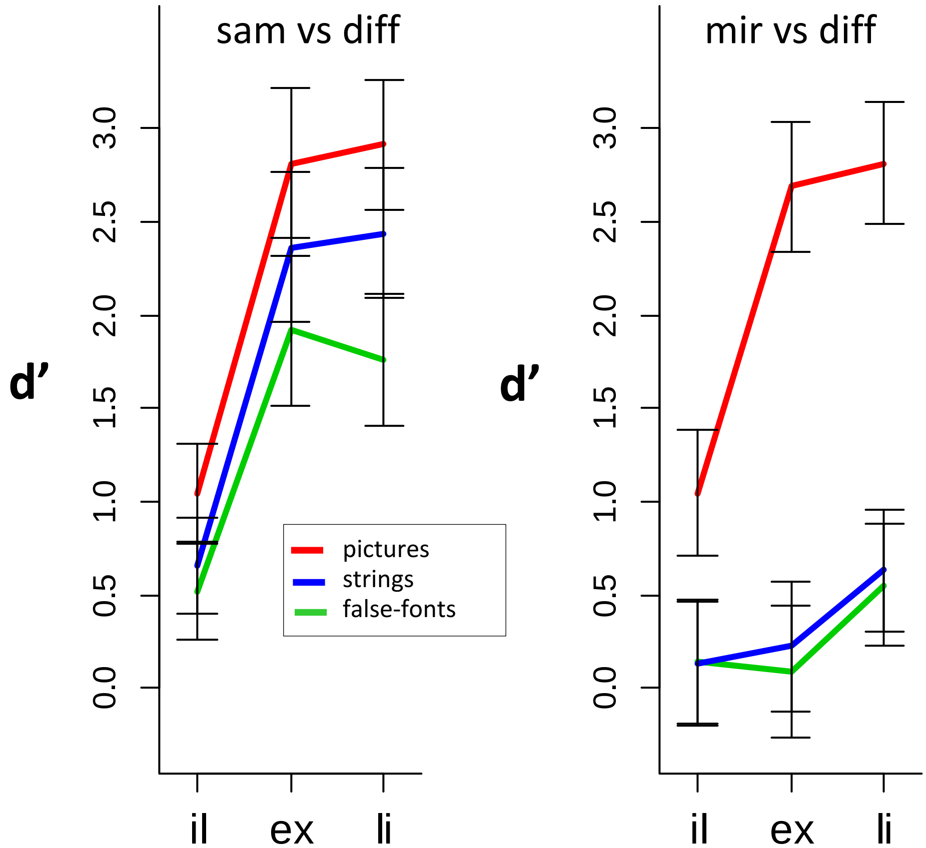

Figure S7: D-primes restricted to matched-groups. D-primes for « same vs. different » and « mirror vs. different » are plotted for each of the matched literacy-groups and category. il = illiterates; ex = ex-illiterates; li = literates. Error bars represent 95% within-subject Confidence Intervals (see equation #3 in Masson & Loftus, 2003).

Table 1. Response times (miliseconds) for each literacy-group, category and condition

illiterates

	faces	houses	tools	strings	falsefonts
same	1039 ± 652	1227 ± 832	1071 ± 496	1182 ± 646	1231 ± 519
mirror	1188 ± 772	1102 ± 549	1119 ± 600	1006 ± 391	1171 ± 486
different	1218 ± 800	1148 ± 510	1146 ± 458	958 ± 475	1247 ± 647

ex-illiterates

	faces	houses	tools	strings	falsefonts
same	540 ± 109	552 ± 113	519 ± 120	615 ± 164	639 ± 182
mirror	641 ± 145	700 ± 201	619 ± 174	856 ± 443	837 ± 313
different	714 ± 165	787 ± 166	728 ± 175	949 ± 577	953 ± 659

literates

	faces	houses	tools	strings	falsefonts
same	584 ± 204	561 ± 152	554 ± 167	641 ± 208	687 ± 233
mirror	659 ± 164	783 ± 231	643 ± 190	1068 ± 372	993 ± 273
different	737 ± 190	776 ± 207	723 ± 169	875 ± 218	895 ± 207

Table 2. Hit rates and False alarms for each literacy-group and category

hit rates

	faces	houses	tools	strings	falsefonts
illiterates	.49	.53	.63	.45	.40
ex-illiterates	.84	.74	.90	.67	.54
literates	.97	.95	.98	.90	.68

false alarms (mirr vs diff)

	faces	houses	tools	strings	falsefonts
illiterates	.16	.25	.24	.41	.35
ex-illiterates	.03	.07	.03	.47	.43
literates	.02	.05	.01	.47	.35

false alarms (sam vs diff)

	faces	houses	tools	strings	falsefonts
illiterates	.20	.19	.27	.25	.25
ex-illiterates	.03	.03	.01	.03	.07
literates	.01	.00	.01	.04	.05