Correction for: When forgetting fosters learning: A saliency map for TP computations

Ansgar D. Endress

Scott P. Johnson

Department of Psychology, City, University of London, UK

Department of Psychology, UCLA

Abstract

We discovered a programming error that changes some aspects of the results reported in the main text of Endress and Johnson (2021). The amended results are very similar to those originally reported, and our central conclusions are unaffected. When corrected, a preference for *ABC* units over *BC:D* units emerges only for a forgetting rate of at least 0.6 (rather than 0.4 in the previous report). The results reported in the Supplementary Information as well as the overall conclusions are unaffected. This note provides updated results. ADE accepts responsibility for this mistake.

Correction for: When forgetting fosters learning: A saliency map for TP computations

In Endress and Johnson (2021), we reported simulations with a neural network of a number of Statistical Learning tasks. We evaluated the network performance by comparing its “familiarity” with different types of test items. We calculated the familiarity of the network with a test item by recording the total network activation, either in the entire network (in the main text) or just in the neurons coding for the test items (in Supplementary Information D). We compared the network’s familiarity with the test items in two ways. For each comparison of test items, we calculated normalized difference scores:

$$d = \frac{\text{Item}_1 - \text{Item}_2}{\text{Item}_1 + \text{Item}_2}$$

We then (1) compared the difference scores to the chance level of zero using a signed rank test (across simulated participants) and, in analogy to analyses in developmental populations, (2) compared the proportion of positive difference scores to the chance level of 50% using a binomial test; with 100 simulations per parameter set, the chance level is exceeded when at least 61% of the simulations show positive difference scores. (Below, we call difference scores “significant” if they different from the chance level of zero in a signed rank test. We call the proportion of positive difference scores significant when the proportion of positive (or negative) difference scores differences from the chance level of 50% in a binomial test.)

We found that the network reproduced many Statistical Learning results for intermediate forgetting rates, but not for very low forgetting rates or very high forgetting rates.

We discovered a programming error that affects the results reported in the main text, while the results reported in Supplementary Information D are correct as reported. We now reran the simulations with the ameneded coded as well as with the old code (but using current versions of the R libraries required for our simulations).

The amended results are very similar to those originally reported, and our central conclusions are unaffected. The main difference to the original results concerns the forgetting rate at which a preference for *ABC* units over *BC:D* part-units emerges; these units correspond to words and part-words in linguistic Statistical Learning studies. In the amended simulations, this preference emergees only at forgetting rates of at least 0.6 (rather than 0.4 in the original report). Further, for a forgetting rate of 0.4, a preference for *BC:D* part-units emerges. As in our original report, *BC:D* part-units are thus harder to reject, but can be rejected with suitable forgetting rates, though the rates need to be slightly higher than in the original report.[1]

Except for some numerical differences for forgetting rates where learning was unreliable in the original report (i.e., where our evaluation measures above disagreed), the main results as well as the conclusions remain unaffected.

In the Supplementary Material, we provide a detailed comparison between the amended and the old results. Specifically, we provide updated versions of Figures 3, 4, and 5 as well as updated Tables C1 and C2. We also list all changes in the significance pattern. As mentioned above, except for the preference for *ABC* units over *BC:D* part-units, these changes occurred exclusively in cases where learning was unreliable in both the original and the amended simulations, in general in cases where the significance pattern was inconsistent between the continuous (signed-rank) and the count-based tests in both the original and the amended simulations. As a result, these changes do not affect our central conclusion that a Hebbian learning model can account for a variety of Statistical Learning results at intermediate forgetting rates. ADE accepts responsibility for this mistake.

[1] This is most likely because *BC:D* part-units activate an extra syllable not contained in the test item (i.e., an A item); if forgetting is strong enough, this item will no longer be activated as strongly by the C item, and units thus become preferred over part-units.

References

Endress, A. D., & Johnson, S. P. (2021). When forgetting fosters learning: A neural network model for statistical learning. *Cognition*, *104621*. doi: 10.1016/j.cognition.2021.104621

Supporting Material

Below, we report an exhaustive list of differences between the old and the amended simulations. We report these differences for forward units, backward units and phantom-units in Supplementary Materials A, B and C, respectively.

Supplementary Material A

Forward units, part-units, rule-units and class-units

As mentioned above, and as shown in Figures A1 and A2, the main difference between the old and the new simulations is that the previously reported preference for *ABC* units over *BC:D* part-units for a forgetting rate of 0.4 turned into a significant preference (by both measures) for *BC:D* part-units; this is consistent with our original finding that these part-units are relatively harder to reject than *C:DE* part-units. For higher forgetting rates, the results are similar to those reported previously, suggesting that successful discrimination between *ABC* and *BC:D* items requires slightly higher forgetting rates than in our earlier simulations when the familiarity of the network is assessed using the total network activity.

For completeness, we now report all changes in the pattern of significance between the two sets of simulations; these changes occurred exclusively when the difference scores were and remain close to zero, suggesting that conditions that did not yield reliable learning in the earlier simulations did not do so in the current simulations either.

For a forgetting rate of zero, the proportion of positive difference scores for the *ABC* vs. *CD:E* comparison dropped from 68% in the old simulations to 53% in the new simulations. However, the difference scores did not differ from chance in either set of simulations, suggesting that learning was not reliable in either set of simulations (and was characterized as unreliable in the original manuscript).

For a forgetting rate of 0.2, a significant preference for *BC:D* items over *ABC* items no longer reached significance. However, the proportion of simulations showing negative difference scores did not differ from chance in either set of simulations, suggesting that it was not reliable in either set of simulations (and was characterized as unreliable in the original manuscript).

For a forgetting rate of 0.6, the proportion of simulations with a positive difference score in the *AXC* vs. *AXF* comparison increased from 58% to 70%; when evaluated as numeric variables, the difference scores exceeded chance in both sets of simulations, strengthening the conclusion that TPs among non-adjacent items can be tracked at intermediate forgetting rates.

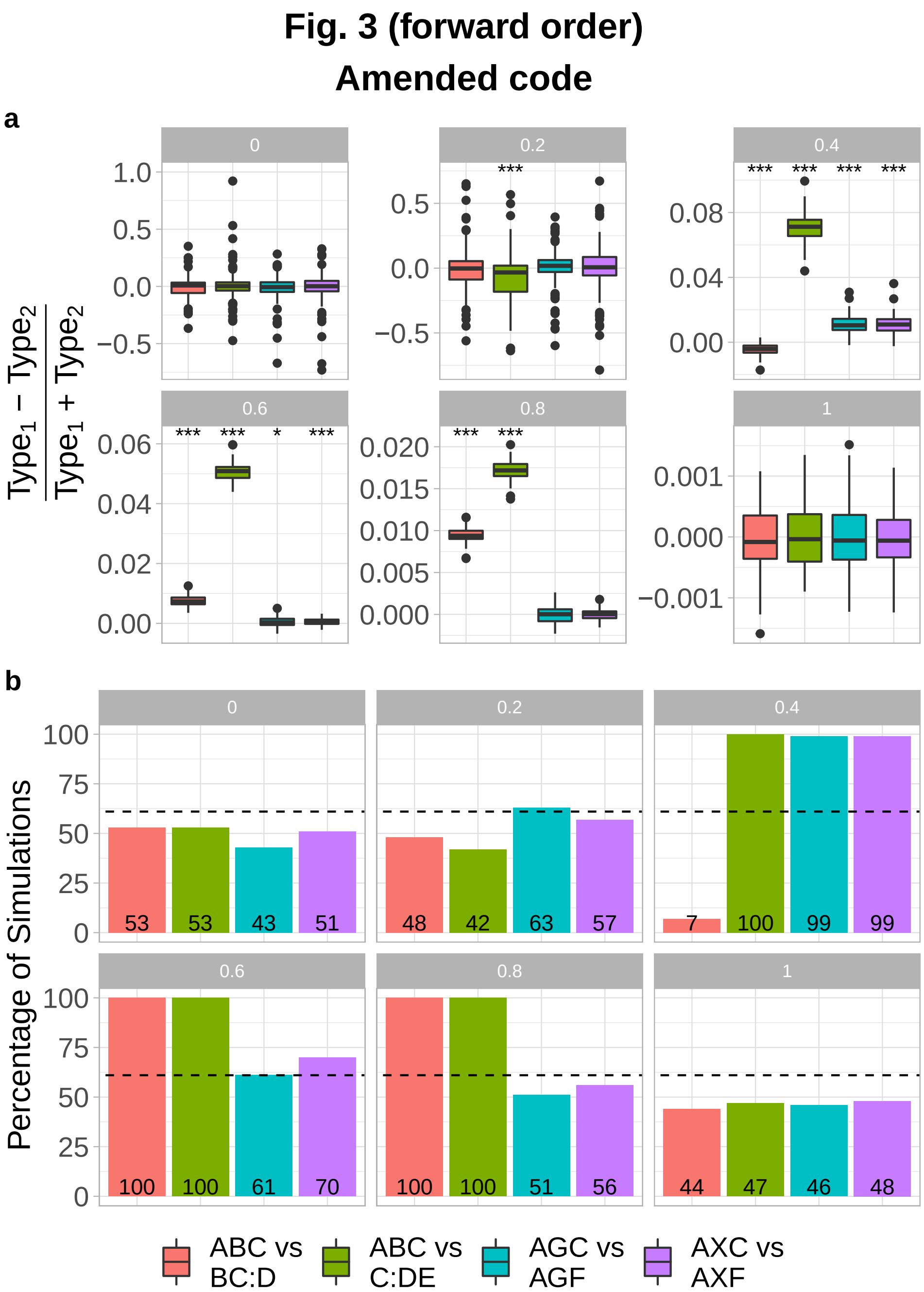

Figure A1. Results with the *amended* code for items presented in *forward order*, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Rule-Unit vs. Class-Unit: *AGC* vs. *AGF* and *AXC* vs. *AXF*). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

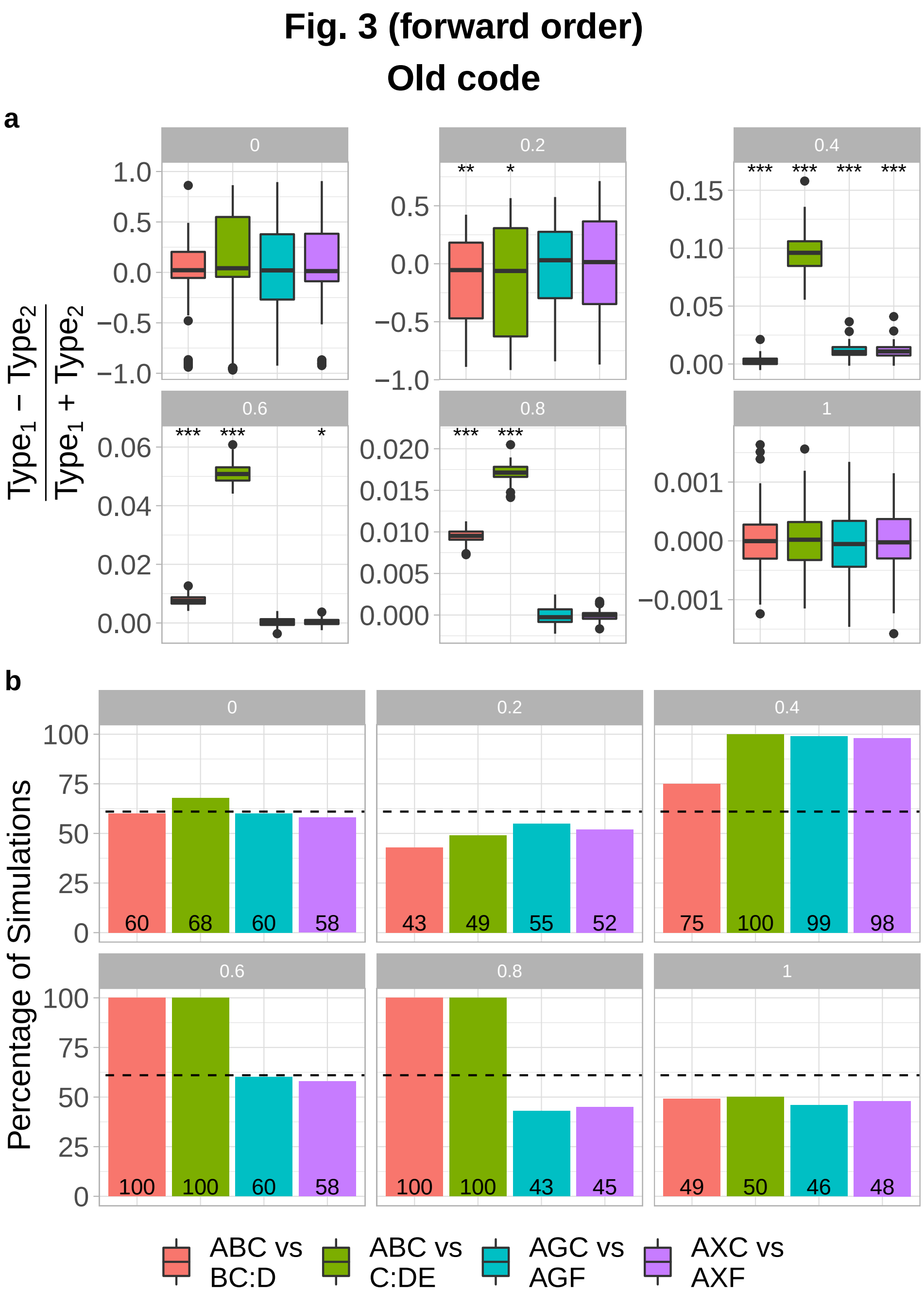

Figure A2. Results with the *original* code for items presented in *forward order*, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Rule-Unit vs. Class-Unit: *AGC* vs. *AGF* and *AXC* vs. *AXF*). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

Supplementary Material B

Backward units, part-units, rule-units and class-units

For a forgetting rate of zero, as shown in Figures B1 and B2, the preference for reverse *ABC* items over reverse *C:DE* items turned significant. However, the proportion of simulations with a positive difference score did not differ from chance in either set of simulations. Further, the proportion of simulations with a positive difference score for reverse *AXC* items over reverse *AXF* items decreased from 62% to 51%. However, the difference scores did not differ from chance in either set of simulations. These differences thus affected results that we unreliable in both sets of simulations (and that were presented as such in the original manuscript).

For a forgetting rate of 1.0, the proportion of simulations with a positive difference score for the comparison between reverse *ABC* items and reverse *BC:D* items increased from 54% to 61%, and the corresponding difference score turned to be statistically significant.

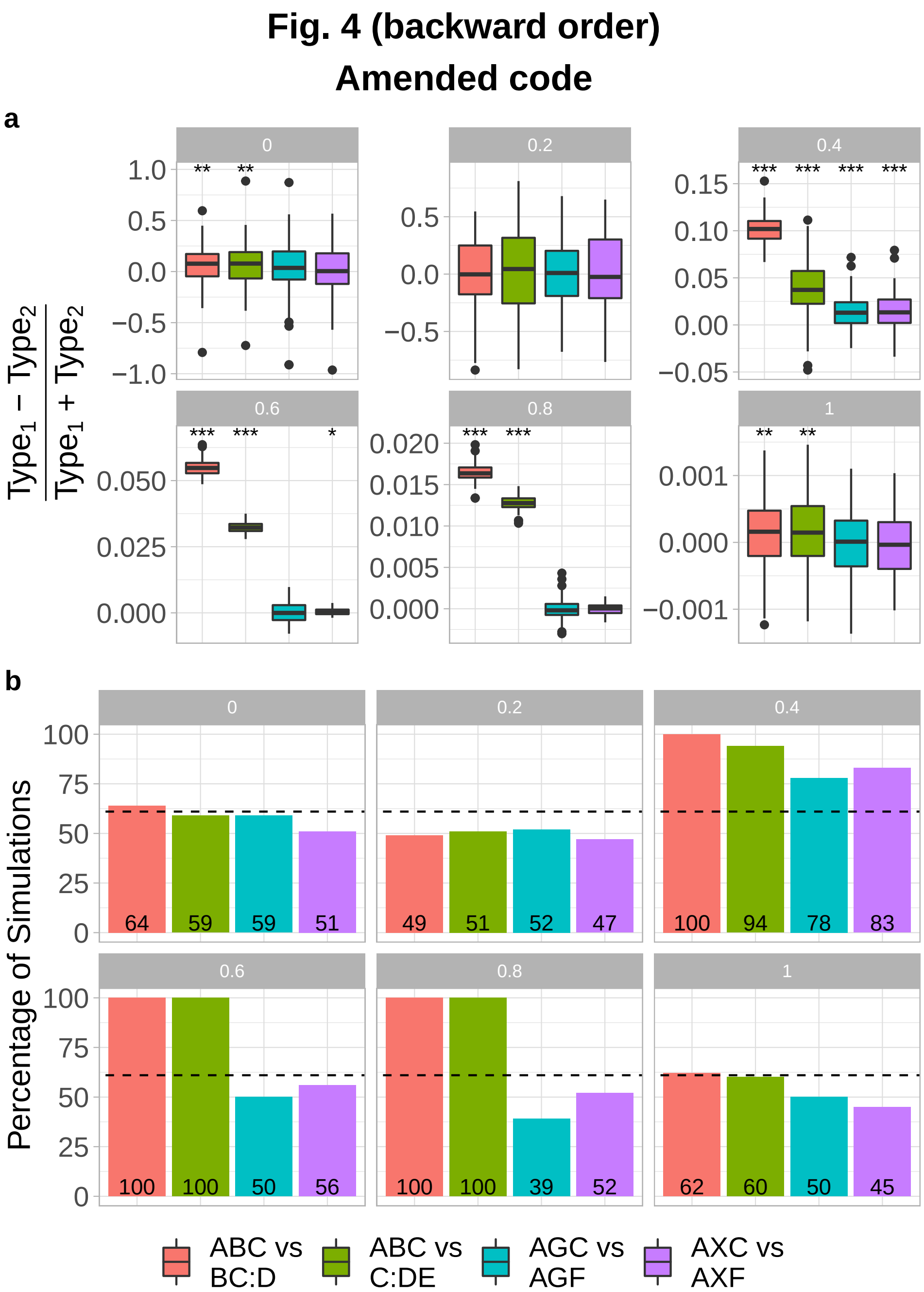

Figure B1. Results with the *amended* code for items presented in backward order, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Backward Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Backward Rule-Unit vs. Class-Unit: *AGC* vs. *AGF* and *AXC* vs. *AXF*). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

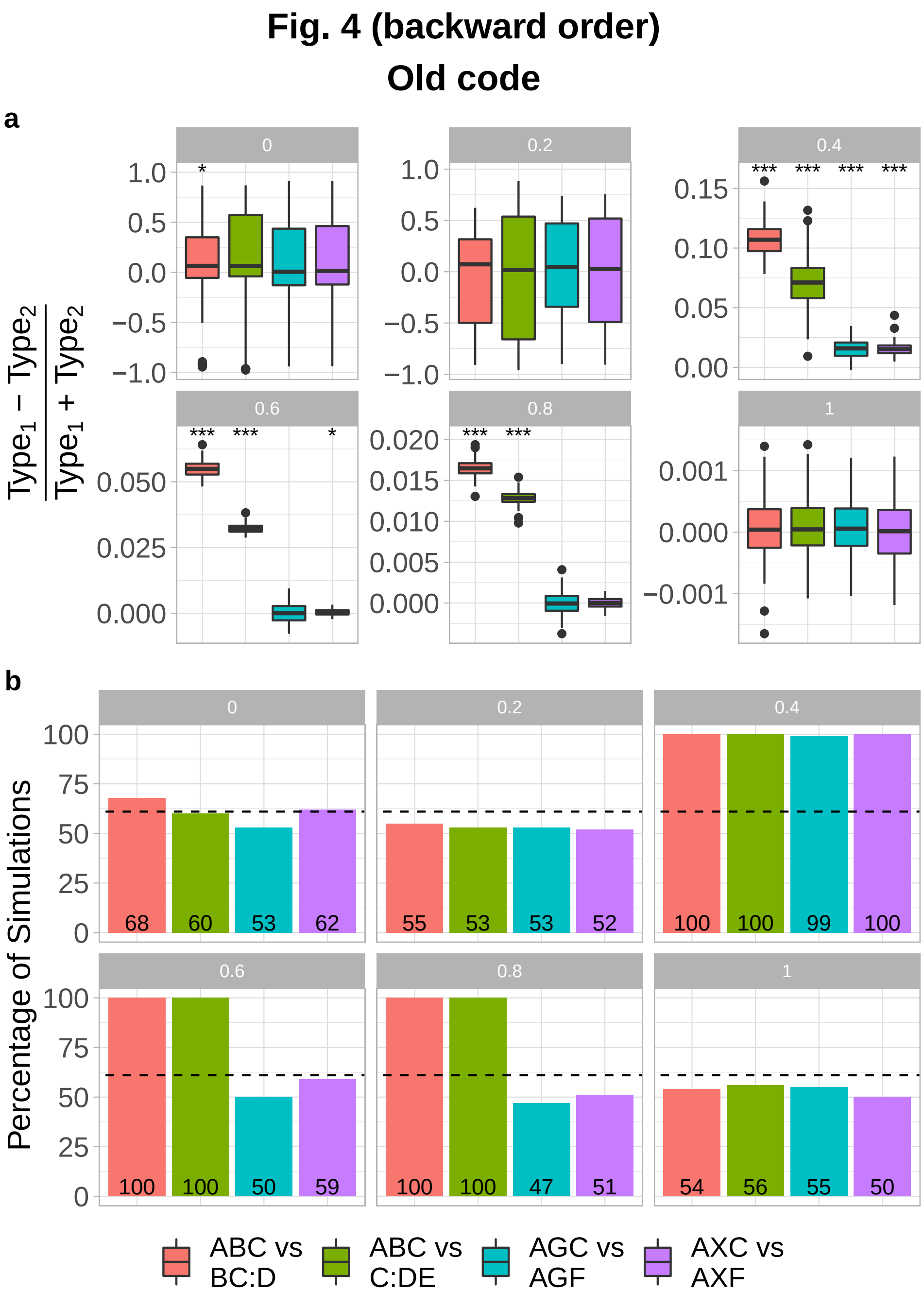

Figure B2. Results with the *original* code for items presented in backward order, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Backward Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Backward Rule-Unit vs. Class-Unit: *AGC* vs. *AGF* and *AXC* vs. *AXF*). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

Supplementary Material C

Simulations with phantom-units

For a forgetting rate of zero, as shown in Figures C1 and C2, a small but significant preference for phantom-units over units (difference score: 0.062) turned non-significant. However, the proportion of simulations with positive difference scores did not differ from chance in either set of simulations, suggesting that this effect was unreliable in both sets of simulations.

For a forgetting rate of 0.2, the proportion of simulations with a positive difference score for the *Unit* and *phantom-unit* vs. *BC:D* items turned from significantly above chance to significantly below chance, perhaps reflecting the relative difficulty of rejecting *BC:D* items. However, the difference scores did not differ from chance in either set of simulations, again suggesting that learning was unreliable in both sets of simulations. The proportion of simulations with a positive difference score for the *Unit* vs. *C:DE* comparison dropped from 62% to 54%.

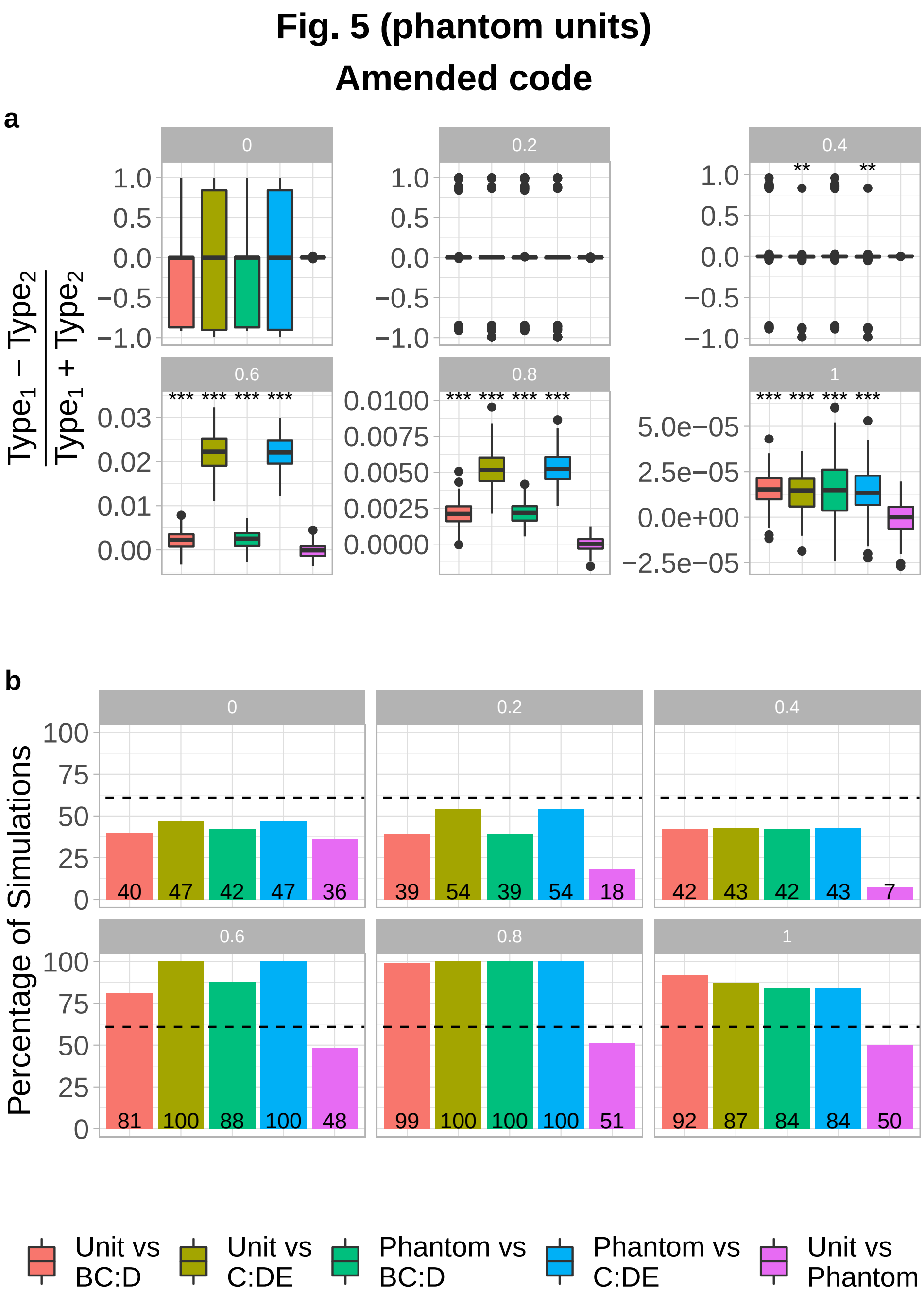

Figure C1. Results with the *amended* code of the simulations comprising phantom-units, for items presented in forward order, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Phantom-Unit vs. Part-Unit: Phantom-Unit vs. *BC:D* and Phantom-Unit vs. *C:DE*; Unit vs. Phantom-Unit). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

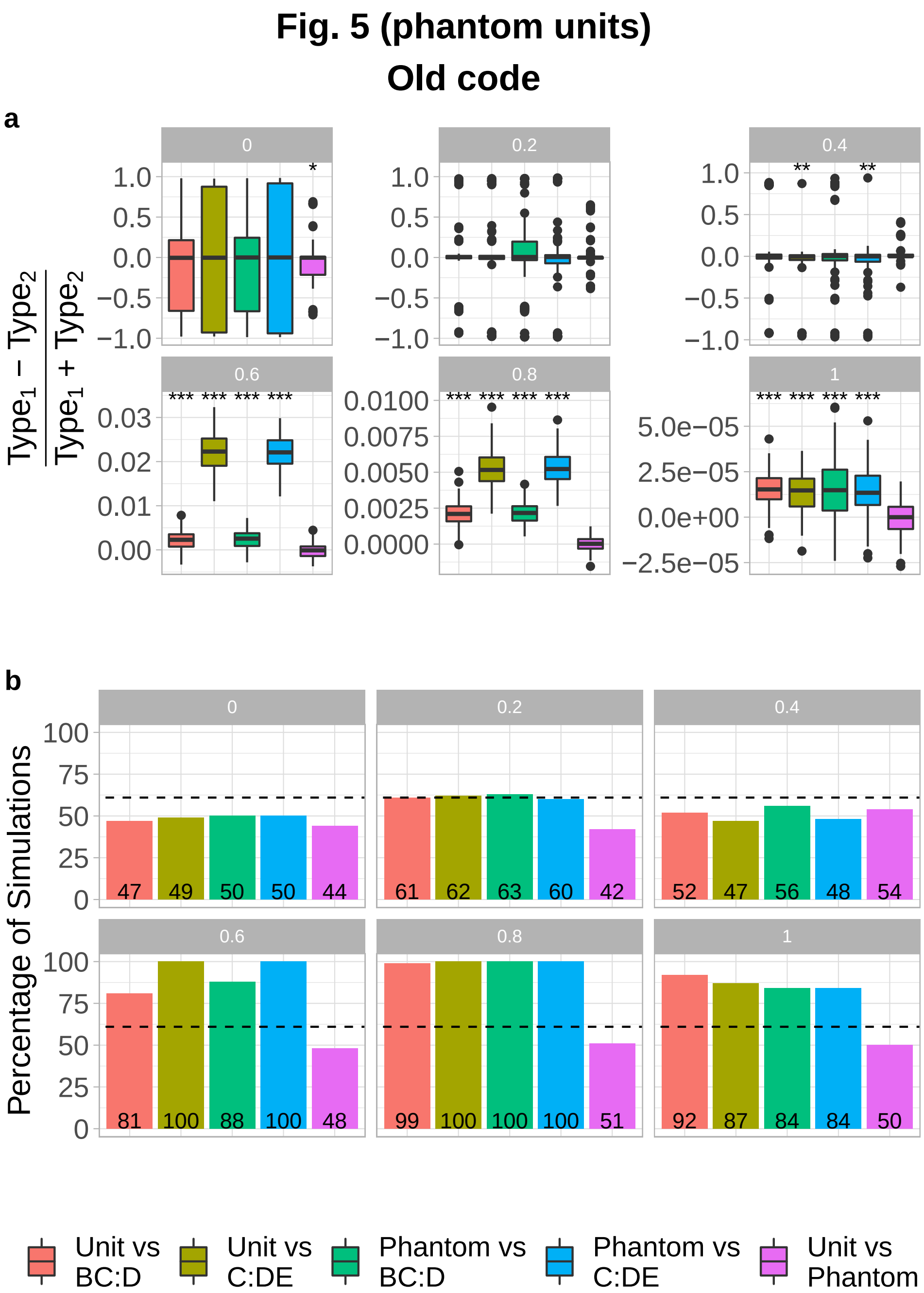

Figure C2. Results with the *old* code of the simulations comprising phantom-units, for items presented in forward order, different forgetting rates (0, 0.2, 0.4, 0.6, 0.8 and 1), and for the different comparisons (Unit vs. Part-Unit: *ABC* vs. *BC:D* and *ABC* vs. *C:DE*; Phantom-Unit vs. Part-Unit: Phantom-Unit vs. *BC:D* and Phantom-Unit vs. *C:DE*; Unit vs. Phantom-Unit). The scores are calculated based the global activation as a measure of the network's familiarity with the items. (a) Difference scores. Significance is assessed based on Wilcoxon tests against the chance level of zero. (b) Percentage of simulations with a preference for the target items. The dashed line shows the minimum percentage of simulations that is significant based on a binomial test.

Supplementary Material D

Updated tables

Table C1

Detailed results for the different forgetting rates and comparisons (Unit vs. Part-Unit: ABC *vs.* BC:D *and* ABC *vs.* C:DE*; Rule-Unit vs. Class-Unit:* AGC *vs.* AGF *and* AXC *vs.* AXF*), for items presented in forward and backward order, and using the global activation as a measure of the network's familiarity with the items.* $p_{Wilcoxon}$ *represents the p value of a Wilcoxon test on the difference scores against the chance level of zero.* $P_{Simulations}$ *represents the proportion of simulations showing positive difference scores.*

λ_a	Statistic	ABC vs. BC:D	ABC vs. C:DE	AGC vs. AGF	AXC vs. AXF
Forward					
0	*M*	-2.54×10^{-3}	6.66×10^{-3}	-16.2×10^{-3}	-11.4×10^{-3}
0	*SE*	-255×10^{-6}	670×10^{-6}	-1.62×10^{-3}	-1.15×10^{-3}
0	$p_{Wilcoxon}$	777×10^{-3}	949×10^{-3}	401×10^{-3}	777×10^{-3}
0	$P_{Simulations}$	530×10^{-3}	530×10^{-3}	430×10^{-3}	510×10^{-3}
200×10^{-3}	*M*	3.04×10^{-3}	-63.2×10^{-3}	152×10^{-6}	4.13×10^{-3}
200×10^{-3}	*SE*	306×10^{-6}	-6.35×10^{-3}	15.2×10^{-6}	415×10^{-6}
200×10^{-3}	$p_{Wilcoxon}$	753×10^{-3}	947×10^{-6}	123×10^{-3}	385×10^{-3}
200×10^{-3}	$P_{Simulations}$	480×10^{-3}	420×10^{-3}	630×10^{-3}	570×10^{-3}
400×10^{-3}	*M*	-4.64×10^{-3}	70.7×10^{-3}	11.3×10^{-3}	11.2×10^{-3}
400×10^{-3}	*SE*	-467×10^{-6}	7.11×10^{-3}	1.13×10^{-3}	1.13×10^{-3}
400×10^{-3}	$p_{Wilcoxon}$	49.5×10^{-18}	3.96×10^{-18}	4.20×10^{-18}	4.33×10^{-18}
400×10^{-3}	$P_{Simulations}$	70.0×10^{-3}	1.00	990×10^{-3}	990×10^{-3}
600×10^{-3}	*M*	7.55×10^{-3}	50.7×10^{-3}	400×10^{-6}	534×10^{-6}
600×10^{-3}	*SE*	759×10^{-6}	5.10×10^{-3}	40.2×10^{-6}	53.7×10^{-6}
600×10^{-3}	$p_{Wilcoxon}$	3.96×10^{-18}	3.96×10^{-18}	13.4×10^{-3}	28.4×10^{-6}
600×10^{-3}	$P_{Simulations}$	1.00	1.00	610×10^{-3}	700×10^{-3}
800×10^{-3}	*M*	9.44×10^{-3}	17.2×10^{-3}	525×10^{-9}	-2.66×10^{-6}

Table C1

Detailed results for the different forgetting rates and comparisons (Unit vs. Part-Unit: ABC *vs.* BC:D *and* ABC *vs.* C:DE*; Rule-Unit vs. Class-Unit:* AGC *vs.* AGF *and* AXC *vs.* AXF*), for items presented in forward and backward order, and using the global activation as a measure of the network's familiarity with the items.* $p_{Wilcoxon}$ *represents the p value of a Wilcoxon test on the difference scores against the chance level of zero.* $P_{Simulations}$ *represents the proportion of simulations showing positive difference scores. (continued)*

λ_a	Statistic	ABC vs. BC:D	ABC vs. C:DE	AGC vs. AGF	AXC vs. AXF
800×10^{-3}	*SE*	949×10^{-6}	1.73×10^{-3}	52.8×10^{-9}	-267×10^{-9}
800×10^{-3}	$p_{Wilcoxon}$	3.96×10^{-18}	3.96×10^{-18}	830×10^{-3}	851×10^{-3}
800×10^{-3}	$P_{Simulations}$	1.00	1.00	510×10^{-3}	560×10^{-3}
1.00	*M*	-43.6×10^{-6}	751×10^{-9}	-12.4×10^{-6}	-33.1×10^{-6}
1.00	*SE*	-4.38×10^{-6}	75.5×10^{-9}	-1.25×10^{-6}	-3.33×10^{-6}
1.00	$p_{Wilcoxon}$	491×10^{-3}	873×10^{-3}	619×10^{-3}	528×10^{-3}
1.00	$P_{Simulations}$	440×10^{-3}	470×10^{-3}	460×10^{-3}	480×10^{-3}
Backward					
0	*M*	50.5×10^{-3}	54.7×10^{-3}	30.2×10^{-3}	11.5×10^{-3}
0	*SE*	5.08×10^{-3}	5.50×10^{-3}	3.04×10^{-3}	1.16×10^{-3}
0	$p_{Wilcoxon}$	2.25×10^{-3}	7.51×10^{-3}	57.9×10^{-3}	369×10^{-3}
0	$P_{Simulations}$	640×10^{-3}	590×10^{-3}	590×10^{-3}	510×10^{-3}
200×10^{-3}	*M*	-11.0×10^{-3}	14.9×10^{-3}	33.4×10^{-3}	18.1×10^{-3}
200×10^{-3}	*SE*	-1.10×10^{-3}	1.49×10^{-3}	3.36×10^{-3}	1.82×10^{-3}
200×10^{-3}	$p_{Wilcoxon}$	971×10^{-3}	491×10^{-3}	320×10^{-3}	549×10^{-3}
200×10^{-3}	$P_{Simulations}$	490×10^{-3}	510×10^{-3}	520×10^{-3}	470×10^{-3}
400×10^{-3}	*M*	101×10^{-3}	38.2×10^{-3}	13.9×10^{-3}	14.4×10^{-3}
400×10^{-3}	*SE*	10.2×10^{-3}	3.84×10^{-3}	1.40×10^{-3}	1.45×10^{-3}
400×10^{-3}	$p_{Wilcoxon}$	3.96×10^{-18}	1.32×10^{-15}	222×10^{-12}	53.1×10^{-12}

Table C1

Detailed results for the different forgetting rates and comparisons (Unit vs. Part-Unit: ABC *vs.* BC:D *and* ABC *vs.* C:DE*; Rule-Unit vs. Class-Unit:* AGC *vs.* AGF *and* AXC *vs.* AXF*), for items presented in forward and backward order, and using the global activation as a measure of the network's familiarity with the items.* $p_{Wilcoxon}$ *represents the p value of a Wilcoxon test on the difference scores against the chance level of zero.* $P_{Simulations}$ *represents the proportion of simulations showing positive difference scores. (continued)*

λ_a	Statistic	ABC vs. BC:D	ABC vs. C:DE	AGC vs. AGF	AXC vs. AXF
400×10^{-3}	$P_{Simulations}$	1.00	940×10^{-3}	780×10^{-3}	830×10^{-3}
600×10^{-3}	M	55.0×10^{-3}	32.4×10^{-3}	224×10^{-6}	362×10^{-6}
600×10^{-3}	SE	5.52×10^{-3}	3.26×10^{-3}	22.6×10^{-6}	36.4×10^{-6}
600×10^{-3}	$p_{Wilcoxon}$	3.96×10^{-18}	3.96×10^{-18}	684×10^{-3}	21.1×10^{-3}
600×10^{-3}	$P_{Simulations}$	1.00	1.00	500×10^{-3}	560×10^{-3}
800×10^{-3}	M	16.5×10^{-3}	12.7×10^{-3}	-85.0×10^{-6}	-24.1×10^{-6}
800×10^{-3}	SE	1.66×10^{-3}	1.28×10^{-3}	-8.54×10^{-6}	-2.42×10^{-6}
800×10^{-3}	$p_{Wilcoxon}$	3.96×10^{-18}	3.96×10^{-18}	260×10^{-3}	661×10^{-3}
800×10^{-3}	$P_{Simulations}$	1.00	1.00	390×10^{-3}	520×10^{-3}
1.00	M	136×10^{-6}	164×10^{-6}	-14.2×10^{-6}	-43.8×10^{-6}
1.00	SE	13.7×10^{-6}	16.5×10^{-6}	-1.43×10^{-6}	-4.40×10^{-6}
1.00	$p_{Wilcoxon}$	9.02×10^{-3}	3.77×10^{-3}	819×10^{-3}	340×10^{-3}
1.00	$P_{Simulations}$	620×10^{-3}	600×10^{-3}	500×10^{-3}	450×10^{-3}

Table C2

Detailed results for the different forgetting rates and comparisons, using the global activation as a measure of the network's familiarity with the items. $p_{Wilcoxon}$ *represents the p value of a Wilcoxon test on the difference scores against the chance level of zero.* $P_{Simulations}$ *represents the proportion of simulations showing positive difference scores.*

λ_a	Statistic	Unit vs. BC:D	Unit vs. C:DE	Phantom vs. BC:D	Phantom vs. C:DE	Unit vs. Phantom
0	M	-146×10^{-3}	-107×10^{-3}	-146×10^{-3}	-107×10^{-3}	1.08×10^{-3}
0	SE	-14.6×10^{-3}	-10.7×10^{-3}	-14.6×10^{-3}	-10.7×10^{-3}	109×10^{-6}
0	$p_{Wilcoxon}$	58.8×10^{-3}	67.6×10^{-3}	115×10^{-3}	52.3×10^{-3}	206×10^{-3}
0	$P_{Simulations}$	400×10^{-3}	470×10^{-3}	420×10^{-3}	470×10^{-3}	360×10^{-3}
200×10^{-3}	M	-46.8×10^{-3}	-78.1×10^{-3}	-46.6×10^{-3}	-78.1×10^{-3}	-241×10^{-6}
200×10^{-3}	SE	-4.71×10^{-3}	-7.85×10^{-3}	-4.68×10^{-3}	-7.84×10^{-3}	-24.2×10^{-6}
200×10^{-3}	$p_{Wilcoxon}$	85.3×10^{-3}	488×10^{-3}	138×10^{-3}	482×10^{-3}	682×10^{-3}
200×10^{-3}	$P_{Simulations}$	390×10^{-3}	540×10^{-3}	390×10^{-3}	540×10^{-3}	180×10^{-3}
400×10^{-3}	M	50.3×10^{-3}	-50.7×10^{-3}	50.3×10^{-3}	-50.8×10^{-3}	42.9×10^{-6}
400×10^{-3}	SE	5.06×10^{-3}	-5.10×10^{-3}	5.05×10^{-3}	-5.10×10^{-3}	4.31×10^{-6}
400×10^{-3}	$p_{Wilcoxon}$	530×10^{-3}	1.33×10^{-3}	528×10^{-3}	1.33×10^{-3}	80.1×10^{-3}
400×10^{-3}	$P_{Simulations}$	420×10^{-3}	430×10^{-3}	420×10^{-3}	430×10^{-3}	70.0×10^{-3}
600×10^{-3}	M	2.14×10^{-3}	22.0×10^{-3}	2.32×10^{-3}	22.1×10^{-3}	-178×10^{-6}
600×10^{-3}	SE	215×10^{-6}	2.21×10^{-3}	233×10^{-6}	2.23×10^{-3}	-17.8×10^{-6}
600×10^{-3}	$p_{Wilcoxon}$	451×10^{-15}	3.96×10^{-18}	24.5×10^{-15}	3.96×10^{-18}	242×10^{-3}
600×10^{-3}	$P_{Simulations}$	810×10^{-3}	1.00	880×10^{-3}	1.00	480×10^{-3}
800×10^{-3}	M	2.13×10^{-3}	5.23×10^{-3}	2.16×10^{-3}	5.25×10^{-3}	-23.9×10^{-6}
800×10^{-3}	SE	214×10^{-6}	525×10^{-6}	217×10^{-6}	528×10^{-6}	-2.40×10^{-6}
800×10^{-3}	$p_{Wilcoxon}$	4.08×10^{-18}	3.96×10^{-18}	3.96×10^{-18}	3.96×10^{-18}	900×10^{-3}
800×10^{-3}	$P_{Simulations}$	990×10^{-3}	1.00	1.00	1.00	510×10^{-3}
1.00	M	14.7×10^{-6}	13.6×10^{-6}	15.6×10^{-6}	14.5×10^{-6}	-890×10^{-9}
1.00	SE	1.48×10^{-6}	1.36×10^{-6}	1.57×10^{-6}	1.45×10^{-6}	-89.4×10^{-9}
1.00	$p_{Wilcoxon}$	112×10^{-18}	10.1×10^{-15}	271×10^{-15}	1.20×10^{-12}	634×10^{-3}
1.00	$P_{Simulations}$	920×10^{-3}	870×10^{-3}	840×10^{-3}	840×10^{-3}	500×10^{-3}