LMM for Full Original Sample (N = 175) | LMM for Online Original Sample (N = 80) | LMM for Online Replication Sample (N = 96)

	Predictors	Estimates	Std. Error	CI	t	p	Estimates	Std. Error	CI	t	p	Estimates	Std. Error	CI	t	p
Main Effects	reading time of previous trial (log-transformed)	0.110235	0.001409	0.107472 – 0.112997	78.210829	$< 1.33 \times 10^{-322}$ *	0.076150	0.002059	0.072113 – 0.080186	36.977238	6.662×10^{-297} *	0.149053	0.001818	0.145491 – 0.1526166	81.997617	$< 1.33 \times 10^{-322}$ *
	d-prime	-0.006293	0.001767	-0.009755 – -0.002831	-3.562277	4.903×10^{-4} *	0.058155	0.003309	0.051668 – 0.064641	17.573579	2.265×10^{-68} *	-0.010803	0.002175	-0.015066 – -0.006541	-4.967139	1.043×10^{-6} *
	mean d-prime singletasks	0.084636	0.019585	0.045974 – 0.123298	4.321414	3.717×10^{-5} *	0.093503	0.027546	0.038639 – 0.148366	3.394411	1.801×10^{-3} *	0.111773	0.020097	0.071862 – 0.151684	5.561623	4.223×10^{-7} *
	mean comprehension question performance	0.002841	0.001276	0.000321 – 0.005360	2.225728	3.282×10^{-2} *	0.003365	0.001733	-0.000086 – 0.006816	1.941918	7.555×10^{-2}	0.003458	0.001659	0.000163 – 0.006753	2.084131	5.401×10^{-2}
	de-meaned comprehension question performance	0.000053	0.000038	-0.000023 – 0.000128	1.374475	1.693×10^{-1}	-0.000479	0.000059	-0.000595 – -0.000363	-8.099410	1.282×10^{-15} *	-0.000715	0.000047	-0.000807 – -0.000623	-15.273789	4.111×10^{-52} *
	word frequency	0.643570	0.362057	-0.067231 – 1.354371	1.777538	8.280×10^{-2}	0.262085	0.300729	-0.329714 – 0.853885	0.871498	4.909×10^{-1}	0.276327	0.253880	-0.223607 – 0.776261	1.088417	3.358×10^{-1}
	word length	0.007839	0.000407	0.007039 – 0.008638	19.240591	7.407×10^{-73} *	0.006343	0.000414	0.005530 – 0.007155	15.314692	2.944×10^{-48} *	0.006351	0.000362	0.005641 – 0.007061	17.540547	2.396×10^{-61} *
	word entropy	0.001410	0.000785	-0.000129 – 0.002949	1.796533	8.280×10^{-2}	-0.000629	0.000933	-0.002459 – 0.001201	-0.674362	6.054×10^{-1}	0.000358	0.000826	-0.001262 – 0.001978	0.433541	6.646×10^{-1}
	n-back reaction [reaction vs. no reaction]	0.317464	0.001799	0.313937 – 0.320990	176.444743	$< 1.33 \times 10^{-322}$ *	0.366360	0.002602	0.361259 – 0.371461	140.773620	$< 1.33 \times 10^{-322}$ *	0.335489	0.002293	0.330995 – 0.339982	146.322474	$< 1.33 \times 10^{-322}$ *
	block number	-0.006495	0.000164	-0.006816 – -0.006173	-39.637620	$< 1.33 \times 10^{-322}$ *	-0.008491	0.000247	-0.008975 – -0.008007	-34.390513	2.814×10^{-257} *	-0.008041	0.000224	-0.008479 – -0.007602	-35.940224	5.567×10^{-301} *
	trial number	-0.000456	0.000007	-0.000470 – -0.000442	-64.085235	$< 1.33 \times 10^{-322}$ *	-0.000408	0.000009	-0.000425 – -0.000390	-45.734724	$< 1.33 \times 10^{-322}$ *	-0.000386	0.000008	-0.000402 – -0.000371	-48.625434	$< 1.33 \times 10^{-322}$ *
	recording location (online vs. lab)	-0.219339	0.032284	-0.283072 – -0.155606	-6.793975	2.686×10^{-10} *										
	surprisal	0.001707	0.000151	0.001411 – 0.002002	11.320677	1.368×10^{-28} *	0.001467	0.000186	0.001103 – 0.001832	7.890451	8.560×10^{-15} *	0.001372	0.000144	0.001091 – 0.001654	9.561841	7.969×10^{-21} *
	age	0.009113	0.000991	0.007158 – 0.011068	9.199100	1.751×10^{-16} *	0.005753	0.001802	0.002164 – 0.009343	3.192287	3.150×10^{-3} *	0.008225	0.001452	0.005339 – 0.011110	5.664597	3.219×10^{-7} *
	cognitive load [1-back vs. Reading Only]	0.473800	0.013916	0.446336 – 0.501264	34.046321	8.399×10^{-79} *	0.463516	0.026092	0.411590 – 0.515443	17.764421	5.087×10^{-29} *	0.507560	0.022200	0.463492 – 0.551627	22.863082	4.013×10^{-40} *
	cognitive load [2-back vs. Reading Only]	0.791540	0.026090	0.740046 – 0.843034	30.338989	7.320×10^{-71} *	0.826086	0.048824	0.728872 – 0.923300	16.919748	2.635×10^{-27} *	0.722375	0.031926	0.658996 – 0.785754	22.626797	1.149×10^{-38} *
2-way Interactions	surprisal x age	0.000035	0.000004	0.000027 – 0.000042	9.287151	3.481×10^{-20} *	0.000029	0.000008	0.000013 – 0.000045	3.585275	5.960×10^{-4} *	0.000006	0.000006	-0.000006 – 0.000018	0.973889	3.796×10^{-1}
	surprisal x cognitive load [1-back vs. Reading Only]	-0.001093	0.000161	-0.001409 – -0.000776	-6.771521	2.043×10^{-11} *	-0.000034	0.000311	-0.000644 – 0.000576	-0.108760	9.549×10^{-1}	0.001500	0.000203	0.001102 – 0.001898	7.380984	3.301×10^{-13} *
	surprisal x cognitive load [2-back vs. Reading Only]	-0.001255	0.000163	-0.001575 – -0.000935	-7.681261	2.709×10^{-14} *	-0.000111	0.000317	-0.000732 – -0.000511	-0.349672	8.349×10^{-1}	0.001366	0.000203	0.000968 – 0.001764	6.723279	3.417×10^{-11} *
	age x cognitive load [1-back vs. Reading Only]	-0.002798	0.000776	-0.004330 – -0.001267	-3.606479	5.135×10^{-4} *	-0.000501	0.001651	-0.003788 – -0.002786	-0.303507	8.349×10^{-1}	0.000830	0.001587	-0.002320 – -0.003980	0.523125	6.595×10^{-1}
	age x cognitive load [2-back vs. Reading Only]	-0.002458	0.001454	-0.005329 – -0.000412	-1.690400	9.681×10^{-2}	0.000161	0.003086	-0.005984 – 0.006307	0.052299	9.584×10^{-1}	-0.001084	0.002285	-0.005622 – -0.003453	-0.474395	6.646×10^{-1}
3-way Interactions	surprisal x age x cognitive load [1-back vs. Reading Only]	-0.000111	0.000009	-0.000129 – -0.000094	-12.266076	3.748×10^{-34} *	-0.000074	0.000020	-0.000113 – -0.000035	-3.739961	3.529×10^{-4} *	-0.000023	0.000015	-0.000052 – 0.000005	-1.605011	1.386×10^{-1}
	surprisal x age x cognitive load [2-back vs. Reading Only]	-0.000078	0.000009	-0.000096 – -0.000060	-8.483676	4.384×10^{-17} *	-0.000044	0.000020	-0.000083 – -0.000004	-2.161418	4.408×10^{-2} *	-0.000057	0.000015	-0.000086 – -0.000028	-3.871008	1.559×10^{-4} *
Model Fit	Intra-class Correlation (ICC)	0.46					0.47					0.51				
	Marginal R² / Conditional R²	0.643 / 0.807					0.588 / 0.782					0.612 / 0.808				

Note. All continuous predictors were centered. Degrees of freedom for *p*-values, standard errors and confidence intervals (CI) were computed using Satterthwaite's approximation. All *p*-values reported here are FDR-corrected and were computed using ANOVAs with type III sum of squares. Results that are significant on an alpha-level of 0.05 are marked with a star.

Control Analysis 2-back vs. 1-back: LMM for Full Original Sample (N = 175)

	Predictors	Estimates	Std. Error	CI	t	p
Main Effects	reading time of previous trial (log-transformed)	0.062181	0.001755	0.058740 – 0.065622	35.420630	3.296×10^{-273} *
	d-prime	-0.006667	0.001898	-0.010388 – -0.002947	-3.512595	7.398×10^{-4} *
	mean d-prime singletasks	0.094749	0.020772	0.053746 – 0.135752	4.561318	1.756×10^{-5} *
	mean comprehension question performance	0.003146	0.001352	0.000478 – 0.005815	2.327238	3.018×10^{-2} *
	de-meaned comprehension question performance	0.000038	0.000047	-0.000053 – 0.00013	0.821932	4.837×10^{-1}
	word frequency	-0.128276	0.319365	-0.756126 – 0.499574	-0.401659	6.882×10^{-1}
	word length	0.005220	0.000414	0.004408 – 0.006031	12.616084	4.020×10^{-34} *
	word entropy	0.001711	0.000914	-0.000081 – 0.003502	1.871897	8.171×10^{-2}
	n-back reaction [reaction vs. no reaction]	0.316007	0.001917	0.312250 – 0.319764	164.836212	$< 1.33 \times 10^{-322}$ *
	block number	-0.011363	0.000295	-0.011941 – -0.01079	-38.535522	1.33×10^{-322} *
	trial number	-0.000450	0.000009	-0.000467 – -0.000433	-52.053811	$< 1.33 \times 10^{-322}$ *
	recording location (online vs. lab)	-0.226062	0.034131	-0.293438 – -0.158686	-6.623323	8.894×10^{-10} *
	surprisal	0.001848	0.000161	0.001532 – 0.002164	11.467327	3.901×10^{-29} *
	age	0.008762	0.001101	0.006591 – 0.010934	7.961804	3.751×10^{-13} *
	cognitive load [2-back vs. 1-back]	0.338896	0.020892	0.297669 – 0.380123	16.221045	1.838×10^{-36} *
2-way Interactions	surprisal x age	0.000003	0.000005	-0.000007 – 0.000012	0.545912	6.159×10^{-1}
	surprisal x cognitive load [2-back vs. 1-back]	-0.000148	0.000173	-0.000486 – 0.000191	-0.855146	4.837×10^{-1}
	age x cognitive load [2-back vs. 1-back]	0.000689	0.001156	-0.001593 – 0.00297	0.595977	6.133×10^{-1}
3-way Interaction	surprisal x age x cognitive load [2-back vs. 1-back]	0.000033	0.000010	0.000014 – 0.000052	3.372931	1.144×10^{-3} *
Model Fit	Intra-class Correlation (ICC)	0.44				
	Marginal R² / Conditional R²	0.442 / 0.690				

Note. All continuous predictors were centered. Degrees of freedom for *p*-values, standard errors and confidence intervals (CI) were computed using Satterthwaite's approximation. All *p*-values reported here are FDR-corrected and were computed using ANOVAs with type III sum of squares. Results that are significant on an alpha-level of 0.05 are marked with a star.

	Predictors	Estimates	Std. Error	CI	t	p
Main Effects	surprisal	0.001707	0.000151	0.001411 – 0.002002	11.320677	1.368×10^{-28} *
	age	0.009113	0.000991	0.007158 – 0.011068	9.199100	1.751×10^{-16} *
	cognitive load [1-back vs. Reading Only]	0.473800	0.013916	0.446336 – 0.501264	34.046321	8.399×10^{-79} *
	cognitive load [2-back vs. Reading Only]	0.791540	0.026090	0.740046 – 0.843034	30.338989	7.320×10^{-71} *
2-way Interactions	surprisal x age	0.000035	0.000004	0.000027 – 0.000042	9.287151	3.481×10^{-20} *
	surprisal x cognitive load [1-back vs. Reading Only]	-0.001093	0.000161	-0.001409 – -0.000776	-6.771521	2.043×10^{-11} *
	surprisal x cognitive load [2-back vs. Reading Only]	-0.001255	0.000163	-0.001575 – -0.000935	-7.681261	2.709×10^{-14} *
	age x cognitive load [1-back vs. Reading Only]	-0.002798	0.000776	-0.004330 – -0.001267	-3.606479	5.135×10^{-4} *
	age x cognitive load [2-back vs. Reading Only]	-0.002458	0.001454	-0.005329 – -0.000412	-1.690400	9.681×10^{-2}
3-way Interactions	surprisal x age x cognitive load [1-back vs. Reading Only]	-0.000111	0.000009	-0.000129 – -0.000094	-12.266076	3.748×10^{-34} *
	surprisal x age x cognitive load [2-back vs. Reading Only]	-0.000078	0.000009	-0.000096 – -0.000060	-8.483676	4.384×10^{-17} *
Model Fit	Intra-class Correlation (ICC)	0.46				
	Marginal R² / Conditional R²	0.643 / 0.807				

Note. All continuous predictors were centered. Degrees of freedom for *p*-values, standard errors and confidence intervals (CI) were computed using Satterthwaite's approximation. All *p*-values reported here are FDR-corrected and were computed using ANOVAs with type III sum of squares. Results that are significant on an alpha-level of 0.05 are marked with a star.