

Appendix Supplemental Statistical Materials

Table W1: Topics predictions for click-through rate, for each of the models

	Full Model	LASSO	Holm-Bonferroni
physical health	-.0001221***(.0000443)	-.0001263***(.0000443)	-.0001153***(.0000439)
mental health	-.0000535(.0000335)	-.0000493(.0000335)	-.0000479(.0000335)
drugs	3.800e-06(.0000327)	2.000e-07(.0000328)	3.000e-06(.0000328)
financial	-.0002679***(.0000533)	-.0002794***(.0000533)	-.0003208***(.0000514)
sex	.0001407***(.0000436)	.0001431***(.0000436)	.0001352***(.0000434)
religion	9.900e-06(.0000409)	.0000164(.0000409)	-3.100e-06(.0000409)
famous	.0000551*(.0000316)	.0000527*(.0000317)	.000055*(.0000317)
government & crime	-.0000914***(.0000352)	-.0001032***(.000035)	-.0001249***(.000034)
business	.0000284(.0000352)	.0000303(.0000351)	.0000291(.000035)
education	.0000863***(.0000334)	.0000857***(.0000335)	.0001043***(.0000334)
entertainment	.0001109***(.0000355)	.0001121***(.0000355)	.0001169***(.0000354)
news	.0000362(.0000313)	.0000346(.0000313)	.000042(.0000313)
identity	.0001877***(.0000338)	.0001768***(.0000338)	.0001836***(.0000337)
hobby	-.0000416(.000034)	-.0000398(.0000339)	-.000046(.0000337)
technology	-.0001386***(.0000364)	-.0001427***(.0000364)	-.0001366***(.0000364)
impressions	2.000e-07(4.000e-07)	2.000e-07(4.000e-07)	2.000e-07(4.000e-07)

Standard errors are in parentheses: *** p<.01, ** p<.05, * p<.1

Table W2: Lasso-selected Construct by Construct interactions

	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
impressions	2.000e-07	4.000e-07	0.53	.591655	-5.000e-07	9.000e-07	
Reading Ease	-.0004014	.0000431	-9.30	.0000000	-.0004859	-.0003168	***
Women	.0004436	.0000383	11.60	.0000000	.0003686	.0005186	***
Common Words	.0002843	.0000447	6.36	.0000000	.0001967	.0003719	***
Social Plural Family	-.0001994	.0000423	-4.72	2.400e-06	-.0002823	-.0001165	***
Visual Language	.0002885	.0000372	7.75	.0000000	.0002155	.0003614	***
Intensity of Emotion	.0001907	.0000561	3.40	.0006816	.0000807	.0003008	***
Verb Categories	-.0001622	.0000513	-3.16	.0015527	-.0002627	-.0000618	***
Concrete	-.0000141	.0000439	-0.32	.747569	-.0001001	.0000719	
Compare Contrast	-.000217	.0000482	-4.50	6.900e-06	-.0003116	-.0001224	***
Secret	.0001374	.0000282	4.88	1.100e-06	.0000822	.0001926	***
Men	.0003031	.0000354	8.56	.0000000	.0002337	.0003725	***
Positive Tone	-.0001422	.0000358	-3.98	.0000701	-.0002122	-.0000721	***
Conflict	.0000774	.0000452	1.71	.0868781	-.0000112	.000166	*
Past	.0003377	.0000317	10.65	.0000000	.0002755	.0003998	***
Harm	.0000741	.000051	1.45	.146399	-.0000259	.0001742	
Positive Emotion	-.0000507	.0000482	-1.05	.2926948	-.0001453	.0000438	
Location	-.000323	.0000387	-8.35	.0000000	-.0003988	-.0002472	***
Present	7.900e-06	.0000327	0.24	.8097339	-.0000562	.0000719	
Questions	-.0002958	.00004	-7.39	.0000000	-.0003742	-.0002173	***
Time Related	.0001296	.0000277	4.69	2.800e-06	.0000754	.0001838	***
Aesthetics	.0000886	.0000275	3.23	.0012475	.0000348	.0001425	***
Safekeeping	-.0001381	.0000369	-3.74	.0001831	-.0002104	-.0000657	***
Negative Emotion	.0000735	.0000446	1.65	.0995142	-.0000014	.0001609	*
Disgusting	.0001488	.0000399	3.73	.0001954	.0000705	.0002271	***
Angry	.0001117	.0000471	2.37	.0177333	.0000194	.000204	**
First Person	.0000425	.000049	0.87	.3853437	-.0000535	.0001386	
Second Person	.0000522	.0000335	1.56	.1184474	-.0000133	.0001178	
Fairness	-.0000734	.0000412	-1.78	.0752695	-.0001542	7.500e-06	*
Morality	-.0000489	.0000379	-1.29	.197032	-.0001231	.0000254	
Threat	-.0000573	.0000416	-1.38	.168904	-.0001389	.0000243	
Authority	-.0002163	.0000629	-3.44	.0005908	-.0003397	-.0000929	***
Cuteness	.0000478	.0000293	1.63	.1022911	-9.500e-06	.0001052	
Negative Tone	.0000546	.0000416	1.31	.1897833	-.000027	.0001362	
Reading Ease X Reading Ease	-.0001244	.0000385	-3.23	.0012247	-.0001998	-.000049	***
Reading Ease X Women	-.0000587	.0000405	-1.45	.146575	-.0001381	.0000206	
Common Words X Social Pl. Family	.0000273	.000043	0.64	.520378	-.0000569	.0001115	
Common Words X Visual Language	.0000633	.0000412	1.54	.1244381	-.0000174	.000144	
Common Words X Intensity of Emotion	-.0000871	.0000444	-1.96	.0497503	-.0001742	-1.000e-07	**
Common Words X Verb Categories	-.00004	.0000454	-0.88	.3788044	-.000129	.000049	
Concrete X Conflict	.0000823	.0000503	1.64	.1017704	-.0000163	.000181	
Concrete X Angry	.000189	.0000557	3.39	.0006992	.0000797	.0002982	***
Concrete X Verb Categories	-.0001062	.0000464	-2.29	.0220287	-.0001972	-.0000153	**
Compare Contrast X Men	-.0000977	.0000469	-2.08	.0372414	-.0001896	-5.800e-06	**
Conflict X Past	-.0000256	.000041	-0.62	.5326936	-.000106	.0000548	
Conflict X Harm	-.0000934	.0000326	-2.86	.0041898	-.0001573	-.0000295	***
Conflict X Positive Emotion	.000104	.0000684	1.52	.1287381	-.0000302	.0002381	
Location X Past	-.00002	.0000323	-0.62	.5358865	-.0000833	.0000433	
Present X Questions	.0000812	.0000327	2.49	.0129123	.0000172	.0001452	**
Questions X Women	.0000266	.0000406	0.65	.5126889	-.000053	.0001061	
Secret X Safekeeping	7.300e-06	.0000362	0.20	.8392326	-.0000636	.0000783	
Secret X Disgusting	.0000509	.000028	1.82	.0689168	-3.900e-06	.0001057	*
Secret X Angry	-.0000771	.000043	-1.79	.0732172	-.0001614	7.200e-06	*
First Person X Second Person	.000052	.0000617	0.84	.3989404	-.0000689	.0001729	
First Person X Fairness	.0000988	.0000606	1.63	.1026462	-.0000199	.0002175	
Women X Men	.0001567	.0000345	4.54	5.700e-06	.000089	.0002244	***
Women X Verb Categories	-.0000582	.0000384	-1.52	.129751	-.0001334	.0000171	
Men X Verb Categories	-.000137	.000037	-3.70	.0002163	-.0002095	-.0000644	***
Authority X Cuteness	-.0000541	.0000538	-1.01	.3148083	-.0001595	.0000513	
Threat X Aesthetics	.0000809	.0000408	1.98	.0473074	1.000e-06	.0001609	**
Harm X Negative Tone	-.0001316	.0000287	-4.59	4.400e-06	-.0001879	-.0000754	***
Cuteness X Negative Tone	.0000692	.000027	2.56	.0104904	.0000162	.0001222	**
Mean dependent var	0.01107			SD dependent var	0.00849		
R-squared	0.05021			Number of obs	33677		
F-test	17.71398			Prob > F	0.00000		
Akaike crit. (AIC)	-288804.25684			Bayesian crit. (BIC)	-288147.14035		

*** p<.01, ** p<.05, * p<.1

Table W3: Lasso-selected Construct by Topic interactions

	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval	Sig
impressions	2.000e-07	4.000e-07	0.55	.5808555	-5.000e-07	9.000e-07	
Goals	-.000197	.000036	-5.48	0	-.0002674	-.0001265	***
Financial	-.0003352	.0000597	-5.61	0	-.0004523	-.0002181	***
Government & Crime	-.0001728	.0000357	-4.84	1.300e-06	-.0002429	-.0001028	***
Visual Language	.0001763	.0000349	5.05	4.000e-07	.0001079	.0002447	***
Business	.0000921	.0000341	2.70	.0068824	.0000253	.0001589	***
News	.0000263	.0000331	0.80	.4264933	-.0000385	.0000911	
Auditory Language	.000083	.0000327	2.54	.011492	.0000189	.0001472	**
Religion	-.0000107	.0000433	-0.25	.804797	-.0000956	.0000742	
Identity	.0001915	.0000349	5.50	0	.0001232	.0002598	***
Cuteness	.0000574	.0000296	1.94	.0523999	-6.000e-07	.0001153	*
Sex	.0001201	.0000444	2.71	.0068327	.0000331	.0002071	***
Famous	.0000354	.0000337	1.05	.29348	-.0000307	.0001016	
Positive Tone	-.0001729	.0000345	-5.01	6.000e-07	-.0002406	-.0001052	***
Intensity of Emotion	.0002143	.0000488	4.39	.0000115	.0001186	.00031	***
Drugs	6.800e-06	.0000355	0.19	.8471417	-.0000627	.0000763	
Positive Emotion	-.0000572	.0000485	-1.18	.2378614	-.0001522	.0000378	
Negative Emotion	.000046	.0000418	1.10	.2706174	-.0000358	.0001279	
Entertainment	.0000212	.0000325	0.65	.5138199	-.0000425	.000085	
Disgusting	.0001942	.0000427	4.55	5.400e-06	.0001105	.0002778	***
Physical Health	-.0002004	.0000449	-4.46	8.100e-06	-.0002885	-.0001124	***
Angry	.0001753	.000047	3.73	.0001911	.0000832	.0002673	***
Education	.0001396	.0000364	3.84	.0001257	.0000683	.0002109	***
Verb Categories	.0000176	.0000382	0.46	.6458357	-.0000573	.0000925	
Conjunction	.0000893	.0000259	3.45	.0005617	.0000386	.00014	***
Goals X Financial	.0000669	.0000303	2.21	.027337	7.500e-06	.0001262	**
Goals X Government & Crime	.0000972	.0000318	3.06	.0022159	.000035	.0001595	***
Visual Language X Business	7.500e-06	.0000353	0.21	.8323787	-.0000616	.0000766	
Visual Language X News	-.0000819	.0000354	-2.31	.0209042	-.0001513	-.0000124	**
Auditory Language X Religion	-.0000183	.0000447	-0.41	.681551	-.000106	.0000693	
Auditory Language X Identity	.0000765	.00003	2.55	.0108385	.0000177	.0001354	**
Cuteness X Sex	.0000833	.0000363	2.29	.0217793	.0000121	.0001545	**
Cuteness X Famous	-.0000375	.0000441	-0.85	.3947729	-.0001238	.0000489	
Cuteness X News	-.0000346	.0000406	-0.85	.3939527	-.0001142	.000045	
Positive Tone X Religion	-3.900e-06	.0000353	-0.11	.91152	-.0000732	.0000653	
Positive Tone X News	-.0000227	.0000342	-0.67	.505424	-.0000897	.0000442	
Intensity of Emotion X Drugs	.0000295	.0000429	0.69	.4915314	-.0000546	.0001136	
Positive Emotion X Drugs	-.0000753	.0000389	-1.94	.0529201	-.0001516	9.000e-07	*
Positive Emotion X Sex	-.0000261	.0000588	-0.44	.65675	-.0001413	.0000891	
Positive Emotion X Religion	.0001041	.0000568	1.83	.06673	-7.200e-06	.0002155	*
Positive Emotion X Identity	.0001397	.0000448	3.12	.001799	.000052	.0002275	***
Negative Emotion X Entertainment	-6.600e-06	.0000417	-0.16	.8741928	-.0000884	.0000752	
Disgusting X Physical Health	-.0000143	.0000215	-0.66	.5064533	-.0000565	.0000279	
Disgusting X Drugs	-.000028	.0000295	-0.95	.3429713	-.0000858	.0000298	
Angry X Financial	-2.500e-06	.000079	-0.03	.9743155	-.0001575	.0001524	
Angry X Sex	-.00001	.000068	-0.15	.8833318	-.0001433	.0001233	
Angry X Education	.0002009	.000061	3.29	.0009923	.0000813	.0003205	***
Angry X News	.0000818	.0000575	1.42	.1550203	-.0000309	.0001945	
Verb Categories X Identity	7.800e-06	.0000376	0.21	.8360679	-.0000659	.0000815	
Conjunction X Religion	-.000018	.000033	-0.54	.5860857	-.0000827	.0000467	
Constant	.0103166	.0013311	7.75	0	.0077076	.0129256	***
Mean dependent var		0.01107		SD dependent var		0.00849	
R-squared		0.01479		Number of obs		33677	
F-test		7.75285		Prob > F		0.00000	
Akaike crit. (AIC)		-287625.15920		Bayesian crit. (BIC)		-287195.50611	

*** p<.01, ** p<.05, * p<.1

Robustness Check Details

Robustness Check	Construct Changes
Including Excluded Constructs	Inclusion of these constructs did not significantly change any of the results for the other constructs, excepting for “concrete”, “deliberation”, and “negative emotion”, all of which now had a significant (at p<.05 level, with only “concrete” passing the Holm-Bonferroni correction) and positive impact on click-through rate, and “verb categories” was now not significant.
Using all the original features	This analysis revealed eight weakly significant variables (clickbait-related words, tutorial words, insight-related words, negating words, positive emotion words, usage of ‘they’, usage of ‘I’, and usage of articles) that were not part of any of the constructs that were significant in at least one of the models tested (i.e., the full model, Holm-Bonferroni model or the LASSO-selected model). However, only one of these variables (usage of articles) was significant as an individual variable after Holm-Bonferroni correction to account for multiple comparisons.
LN(1+X) model	Using either approach, the Holm-Bonferroni corrected model using the alternative constructs revealed nearly the same results as in our main Holm-Bonferroni-corrected model, except that “intensity of emotion”, “other senses”, and “action words”, were no longer significant after these corrections.
Logit Transformed Model	Using Holm-Bonferroni correction, we see largely similar results as in the original model except that five constructs (“social plural family”, “desires”, “goals”, “cuteness”, and “positive tone”) become significant in the logit-based model.