Can Tweets Become the New Folktales

Carolina Arriaga, Alice Hua, Manpreet Khural, Randy Moran

W241 Experiment and Causal Inference
Section 4
August 4, 2021

Agenda

1. Research Question
2. Experiment Design
3. Assumptions and Generalizability
4. Randomization/Covariate Check
5. Attrition Analysis
6. Power Analysis
7. Regression Model
8. Outcome

01 Research Question

Motivations

Hypothesis

How do a tweet's truthfulness and sentiment affect a reader's ability to remember the content?

Motivation

- Folktales are persistent narratives of a society.
- Social Media content like Tweets may be the new folktale.
- What kind of content is being conveyed?

Goal

Identify whether memory retention is affected by the sentiment and truthfulness of a tweet.

How do a tweet's truthfulness and sentiment affect a reader's ability to remember the content?

Hypotheses to Test

- **H1:** There's no difference in memory retention of tweets caused by truthfulness
- **H2:** There's no difference in memory retention of tweets caused by sentiment

Expected Outcome Per Research

Expectation 1: More truthful tweets will be remembered more than fake ones

➔ False information in makes less false memory; recall unclear[1,2]

Expectation 2: Negative tweets will be remembered more than positive ones

➔ People remember negative information more often and more vividly[3]

1. Fenn, K, et al. The effect of Twitter Exposure on false memory formation.
2. Robinson, B. A New Study Shows Fake News May Benefit Your Memory.
3. Li, K.K. Asymmetric memory recall of positive and negative events in social interactions.

02

Experiment Design

Treatment & Survey Design

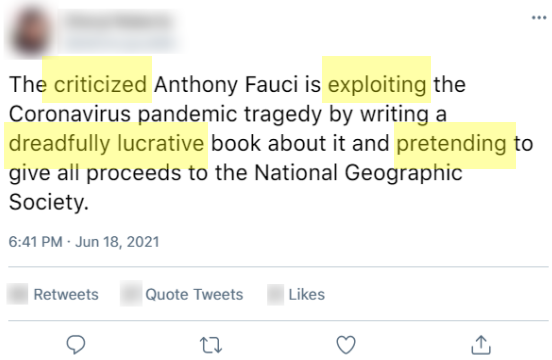
Treatment Design

Using a Factorial Design with two treatments in tweet content, which combination has the strongest effect?

1. **False**: Pulled actual “False”-rated tweets from Politifact for variety of topics
2. **True**: Pulled associated Politifact-corrected tweet statements
3. **Positive/Negative Sentiment:** Added terms that accentuated sentiment to Fake/Fact tweets w/o meaning change. (+/-) 0.8 sentiment scores.
4. **Generated/Blurred tweet info**: fake images, names, dates and reactions
5. **Presentation:** scroll format, similar to real experience.

Example Tweet - True/Positive

Example Tweet - False/Negative



Experimental Design

The experiment helps us test the possibility of controlling the memory retention of respondents considering the following relationship:

$$y = \beta_0 + \beta_1 T + \beta_2 S + \beta_3 T*S \text{ (Eq. 1)}$$

Where:

y: number of tweets remembered
T: truthfulness in the tweet (0,1)
S: sentiment in the tweet (0,1)

Factor	Level (0)	Level (1)
Truthfulness (T)	False	True
Sentiment (S)	Positive	Negative

Baseline treatment: False-Positive (FP)

Treatments = Surveys

True Positive Survey

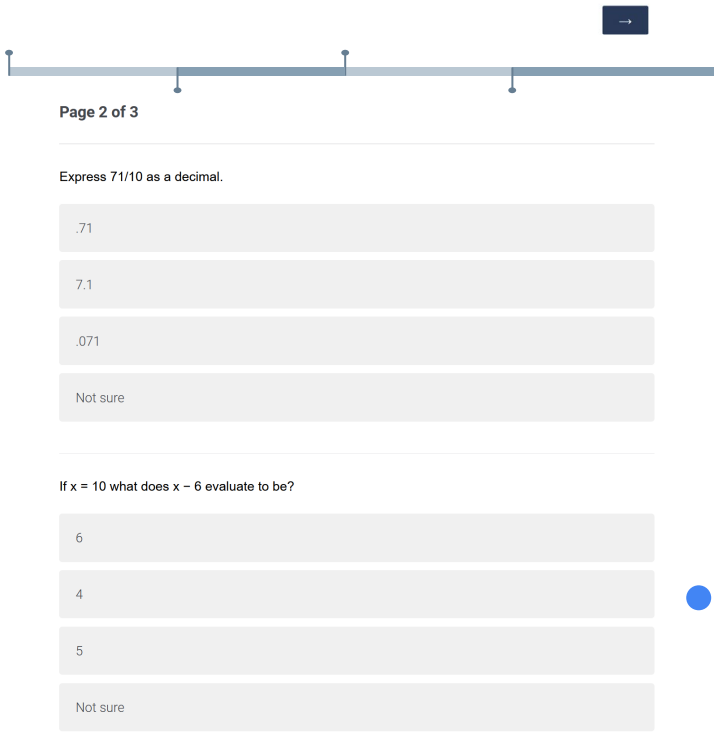
True Negative Survey

False Positive Survey

False Negative Survey

Warning & Distraction Questions

Present unbiased views
&
simulate a time-lapse



Implementation

- **Participant Selection**
  - Chose to use *our social network* to solicit participants (over MTurk, *Qualtrics* Service)
    - Cost effective option
    - Provided better ability to control the randomization for the factorial design to ensure near equal distribution
    - Incentivised Participants with a potential *raffle prize*
- **Survey Design**
  - Four sections: *demographic data, treatment tweets, distraction questions, treatment questions.*
  - *Randomized the treatment question* display so it did not align with the tweet sequence
  - Recorded timing, pre-treatment abandonment and post-treatment abandonment, attention check
- **Measures**
  - Treatment Outcome: Total *Correct Responses* count (Max =7) for the survey
  - Treatment Conditions: *Truthfulness* (False/True) and *Sentiment* (Positive/Negative)
  - Demographic Data: *Age, Education, and Gender*

03

Assumptions & Generalizability

Notes for readers

Assumptions about Potential Outcomes

1. **Excludability** — We assume that the outcome is a response only to the treatment

1. **Non-interference** — No strategic interaction among units

1. **Randomization** — The probability of being assigned to any treatment group is the same for all units.

   Therefore, treatment status is statistically independent of unit's potential outcome and attributes.

Generalizability

- Our study only tests the effect of memory retention with *designed Twitter tweets*. We also conducted the experiment with a given *demographic that agreed to take the survey* + composed by *asymmetric distributions of education level and age*.

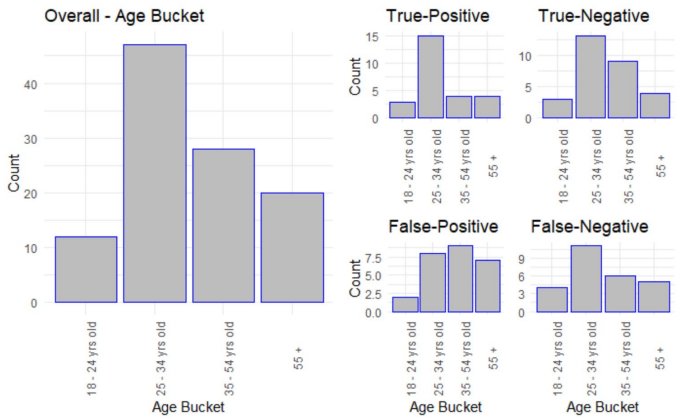
- There is *no certainty* that our results can *generalize* to *other social media mediums* or to a *wider demographic*.

04

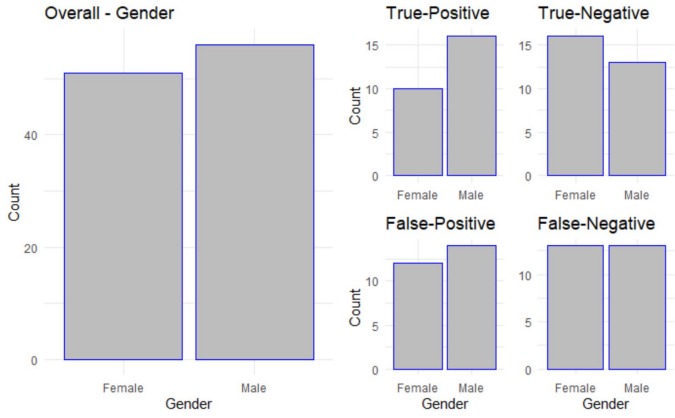
Randomization Check

Between 4 treatment groups

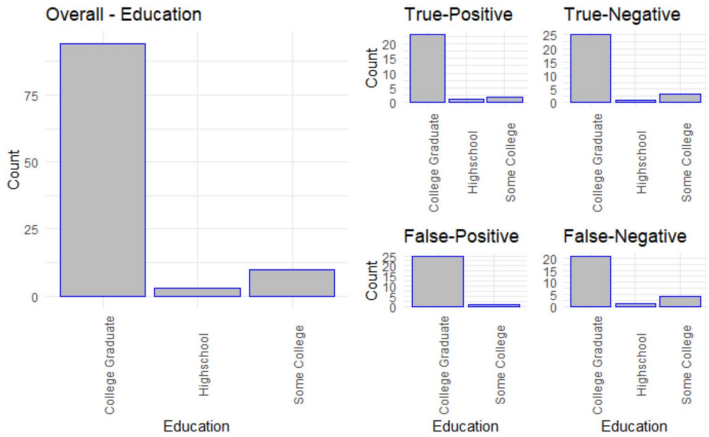
Age



Gender



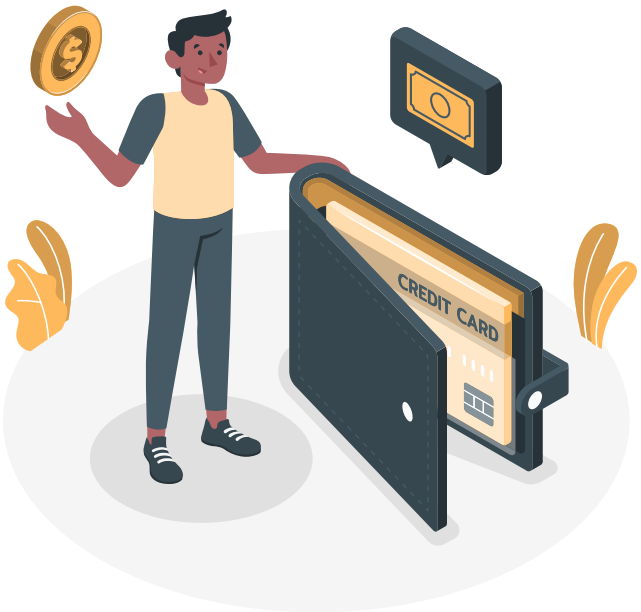
Education



Randomization Check

	Randomization check			
	Four	different	treatment	groups
	FN	TN	TP	FP
	(1)	(2)	(3)	(4)
Male	-0.098	-0.009	0.091	0.017
	(0.090)	(0.087)	(0.086)	(0.085)
Age 25 -34	0.052	-0.068	0.048	-0.032
	(0.150)	(0.145)	(0.143)	(0.142)
Age 35 - 54	0.098	-0.087	-0.139	0.128
	(0.161)	(0.156)	(0.154)	(0.152)
Age 55+	-0.061	-0.080	-0.065	0.206
	(0.169)	(0.163)	(0.161)	(0.159)
Educ Highschool	0.097	0.122	0.171	-0.389
	(0.274)	(0.265)	(0.262)	(0.258)
Educ College	0.055	0.160	-0.039	-0.175
	(0.156)	(0.151)	(0.149)	(0.147)
Constant	0.277*	0.297**	0.222	0.204
	(0.142)	(0.138)	(0.136)	(0.134)
Observations	107	107	107	107
R2	0.023	0.019	0.046	0.070
Adjusted R2	-0.036	-0.040	-0.012	0.014
Residual Std. Error (df = 100)	0.454	0.439	0.433	0.428
F Statistic (df = 6; 100)	0.393	0.325	0.795	1.250
Note:				*p<0.1; **p<0.05; ***p<0.01

05 Attrition Analysis

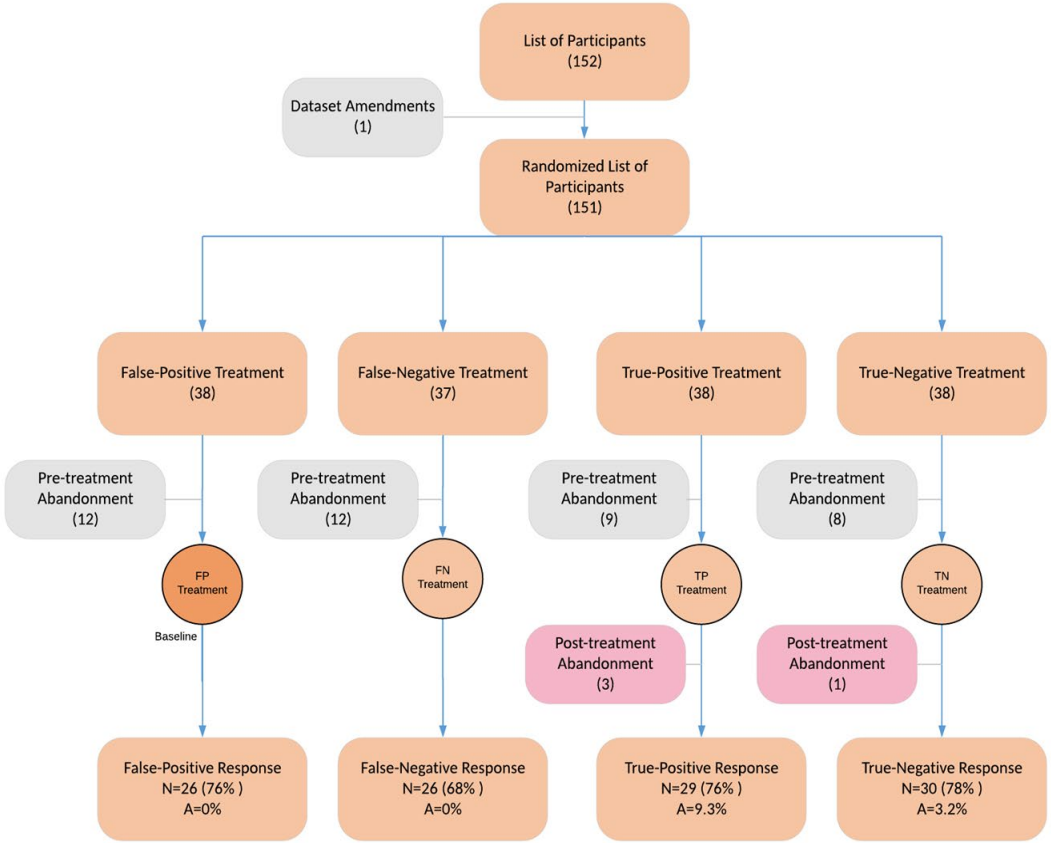


Attrition Analysis

Dataset Amendments
- Duplicate Entry

Pre-Treatment Abandonment
- Did not respond to survey
- Opened survey but exited prior to seeing tweets

Post-Treatment Abandonment
- Did not respond to the survey
- Saw the tweets



Attrition Analysis

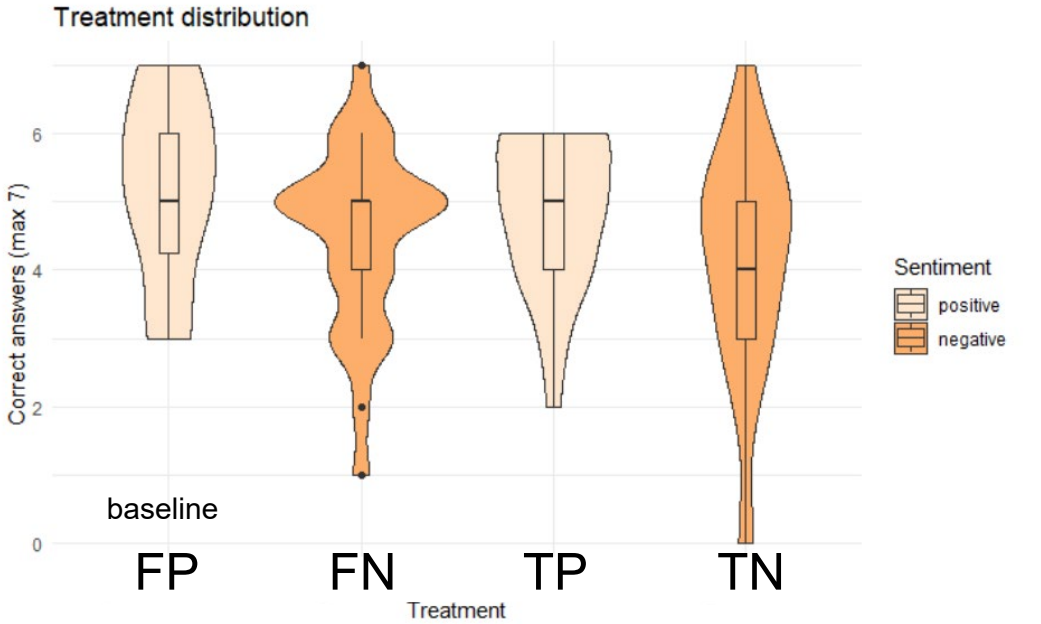
The TP treatment had the highest attrition of 9%, followed by TN with 3%.

To overcome this, we compared the different distributions and checked which bound (lower/upper) in the distribution would decrease the effect size.

Extreme value bounds

TP = higher bound = 7

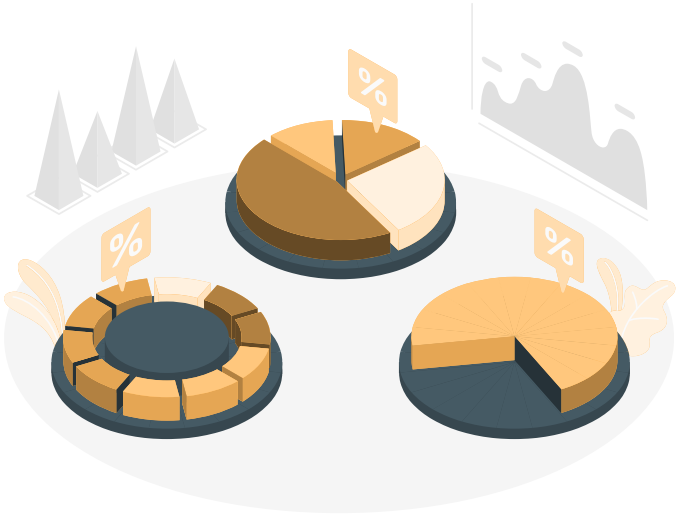
TN = lower bound = 0



06

Power Analysis

You can enter a subtitle here if you need it



Power Analysis

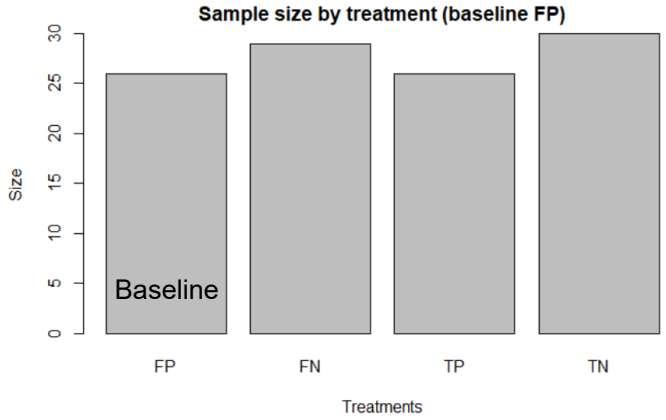
We compared all the sample sizes by treatment, they were very similar (All 26, except for TN which had 29)

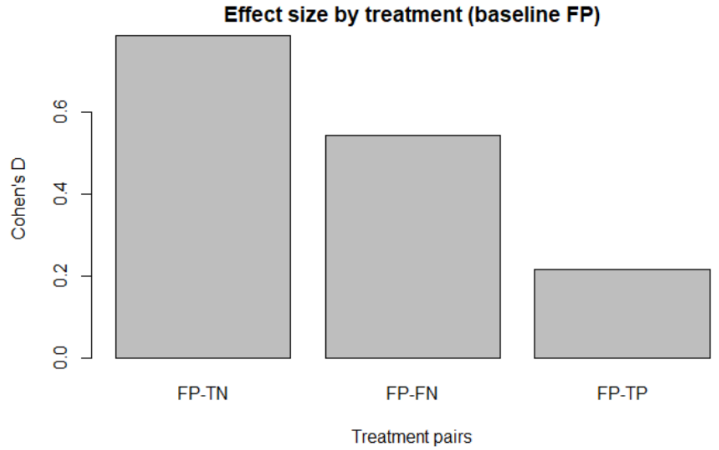
We calculated the power by comparing pairs of treatments to the baseline (FP) to get the effect size. The smallest effect was FP-TP.

We achieved a 96% Power, with a 95 significance level.

```
t test power calculation

        n1 = 26
        n2 = 29
         d = 0.2164672
 sig.level = 0.95
     power = 0.9637201
alternative = two.sided
```





07 Regression Model

Three models

• Modeling

Reduced model (one covariate):

$$y = \beta_0 + \beta_1 T + \epsilon_1 \qquad (1)$$

Extended model:

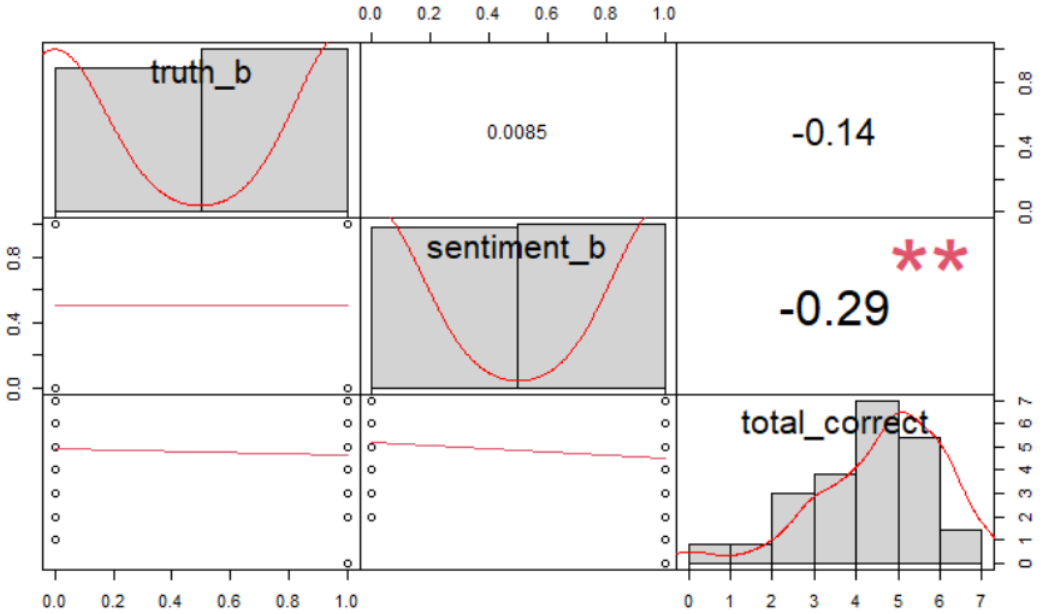
$$y = \beta_0 + \beta_1 T + \beta_2 S + \epsilon_2 \qquad (2)$$

Full model:

$$y = \beta_0 + \beta_1 T + \beta_2 S + \beta_3 T \cdot S + \epsilon_3 \qquad (3)$$

• Correlation check

We tested for independence in our covariates by measuring the correlation between our covariates.



• Regression results (Overall)

Model 4 - proposed by G&G:

H1: Fail to reject

H2: Fail to reject

Added Benjamini-Hochberg p-value for reference.

We believe there's evidence that sentiment had an impact in short-term memory retention (model 2 and 3).

Regression results

	Dependent variable:			
	Correct responses			
	T (1)	S (2)	T+S (3)	T+S+T*S (4)
Truthfulness	-0.408 (0.284)		-0.400 (0.341)	-0.265 (0.388)
Sentiment		-0.877*** (0.274)	-0.873** (0.382)	-0.731* (0.399)
Truth:Sentiment				-0.268 (0.547)
Constant	4.865*** (0.194)	5.091*** (0.232)	5.302*** (0.267)	5.231*** (0.282)
Benjamini-Hochberg pval	0.15	0.05	0.05	0.08
Observations	111	111	111	111
R2	0.019	0.086	0.104	0.106
Adjusted R2	0.010	0.078	0.088	0.081
Residual Std. Error	1.492 (df = 109)	1.440 (df = 109)	1.432 (df = 108)	1.437 (df = 107)
F Statistic	2.064 (df = 1; 109)	10.285*** (df = 1; 109)	6.277*** (df = 2; 108)	4.235*** (df = 3; 107)
Note:				*p<0.1; **p<0.05; ***p<0.01

Factor	Level (0)	Level (1)
Truthfulness (T)	False	True
Sentiment (S)	Positive	Negative

• Regression results (Question level)

Sports tweet
True information had a negative impact in memory.

Pandemic tweet
Negative sentiment had a negative impact in memory, and a strong interaction negatively correlated to sentiment alone.

Regression results

	Dependent variable:					
	Georgians (1)	Energy (2)	Soccer (3)	Pollution (4)	Fauci (5)	Election (6)
False information	-0.194	0.074	-0.294**	-0.149	-0.137	-0.084
	(0.137)	(0.133)	(0.121)	(0.127)	(0.137)	(0.118)
Positive sentiment	-0.000	0.038	0.038	-0.154	-0.500***	-0.077
	(0.144)	(0.139)	(0.098)	(0.131)	(0.122)	(0.121)
False:Positive	-0.211	-0.061	-0.157	-0.0002	0.716***	0.086
	(0.182)	(0.187)	(0.166)	(0.187)	(0.176)	(0.170)
Constant	0.538***	0.615***	0.846***	0.769***	0.654***	0.808***
	(0.102)	(0.099)	(0.074)	(0.086)	(0.097)	(0.080)
Observations	111	111	111	111	111	111
R2	0.121	0.003	0.166	0.049	0.192	0.006
Adjusted R2	0.097	-0.025	0.142	0.022	0.169	-0.022
Residual Std. Error (df = 107)	0.463	0.483	0.439	0.484	0.457	0.441
F Statistic (df = 3; 107)	4.923***	0.113	7.088***	1.820	8.474***	0.211

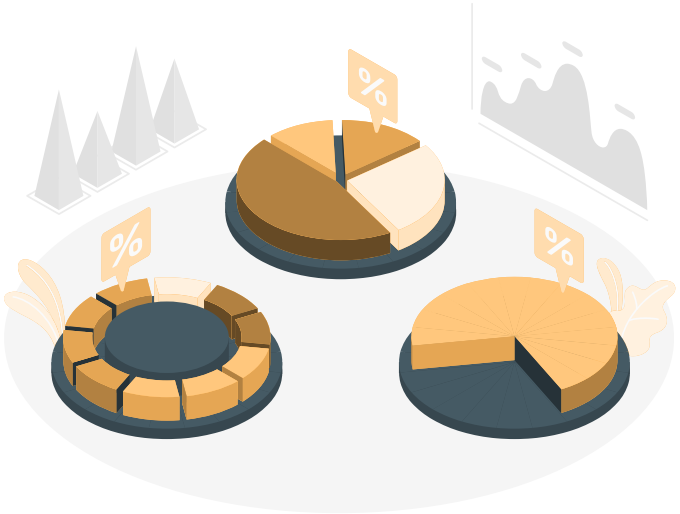
Note: *p<0.1; **p<0.05; ***p<0.01

Factor	Level (0)	Level (1)
Truthfulness (T)	False	True
Sentiment (S)	Positive	Negative

08

Outcome

Hypothesis Validation



Conclusions

- **H1:** There's no difference in memory retention of tweets caused by truthfulness

We didn’t find evidence suggesting that short-term memory was caused by truthfulness in the tweet.

- **H2:** There's no difference in memory retention of tweets caused by sentiment

We found marginal evidence that negative sentiment cause a reduction in short-memory retention.

This is an interesting finding because the literature suggest the opposite (Li, K.K.). A larger study would help us understand the effect.

Threats to validity

- Hard to replicate due to being dependent on text. We used third party verification to facilitate the process: Polifact and VADER sentiment package.
- Selection of subjects can be questionable because it came from a relatively close network -a mix of colleagues, mids students, facebook groups (general) and close friends.
- Attrition of 9% in one of the treatment can point out that something in the tweet affected the outcome.
- Spill-over effect was mitigated by telling people to not share information after responding the survey. Depends on a system of honor.

Further discussion

- Extreme sentiments seem to have somewhat an effect (marginal p-value) which could be further investigated.
- Possibility to try different platforms and see if there's consistency in the results from the tweets.
- We tried as best to have a single standard test. All the questions were the same, but we had two answer sheets. A follow up experiment could seek to combine all into a single standard test and response sheet.
- Both Sports and Pandemic tweets are related in terms of Covid-19 and mortality. This could create a special attention to the content.

Thanks!

Feedback

1. Did we achieve the treatment we meant in our tweet design?
2. We did not use all our budget, could increasing the prize amount or number have helped increase our participants?
3. Did our warning on the tweets increase our abandonment rate?
4. Did our choice to use the personal network limit the results?
5. Threats to validity:
6. Any additional analyses we could do to improve strength of conclusions?

Appendix

1. Fenn, K, et al. The effect of Twitter Exposure on false memory formation, Psychonomic Bulletin & Review, https://www.researchgate.net/publication/262301675_The_effect_of_Twitter_exposure_on_false_memory_formation

1. Robinson, B, A New Study Shows Fake News May Benefit Your Memory, Forbes, Oct 17, 2020 https://www.forbes.com/sites/bryanrobinson/2020/10/17/a-new-study-shows-fake-news-may--benefit-your-memory/?sh=bb9b55b2687b.

1. McGrath D, People remember fake news more if it aligns with their personal beliefs, TheJournal.ie Aug 22, 2019, https://www.thejournal.ie/fake-news-repeal-abortion-eighth-amendment-ucc-cork-4777533-Aug2019/