



# Using longitudinal data to explore young people's social media use

Dr Amy Orben MRC Cognition and Brain Sciences Unit 16 April 2020





# Outline

- Phase I: What got me interested
- Phase II: Data that is 'too big to fail'
- Phase III: Moving longitudinal
- Phase IV: Highlighting individual differences





## Phase I

MRC Cognition and Brain Sciences Unit





Psychologically, however, they are more vulnerable than Millennials were: Rates of teen depression and suicide have skyrocketed since 2011. It's not an exaggeration to describe iGen as being on the brink of the worst mental-health crisis in decades. Much of this deterioration can be traced to their phones.

News Jeremy Hunt: Social media poses as great a threat to children as obesity



Science must begin with myths, and with the criticism of myths.

Karl Popper

Smartphones The Observer

#### Interview Are smartphones really making our children sad?

By Interview by Ian Tucker

US psychologist Jean Twenge, who has claimed that social media is having a malign affect on the young, answers critics who accuse her of crying wolf



www.parliament.uk

Accessibil

Home Partiamentary business MPs, Lords & offices About Parliament Get involved VI House of Commons: House of Lords, What's on Bits & legislation Committees Publications & ou are here Parliament home page - Parliamentary business - Commitse A 24 committees A 24 commitse Sete Common) - Ness - Nignact of social mode and screems-are ony pagefolt health inquiry businets



Commons Select Committee

on young people's health inquiry launched



#### Giving your child a smartphone is like giving them a gram of cocaine, says top addiction expert

Harley Street clinic director Mandy Saligari says many of her patients are 13-year-old girls who see sexting as 'normal'

Rachael Pells Education Correspondent | @rachaelpells | Wednesday 7 June 2017 16:29 BST | D 3 comments

#### f 💟 🖾 149K

(f stiare) (y) (

Like Click to follow



#### Have Smartphones Destroyed a Generation?

More comfortable online than out partying, post-Millennials are safer, physically, than adolescents have ever been But they're on the brink of a mental-health crisis.





Psychologically, however, they are more vulnerable than Millennials were: Rates of teen depression and suicide have skyrocketed since 2011. It's not an exaggeration to describe iGen as being on the brink of the worst mental-health crisis in decades. Much of this deterioration can be traced to their phones.







Empirical Article

Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time PSYCHOLOGICAL SCIENCE

Clinical Psychological Science 2018, Vol. 6(1) 3–17 © The Author(s) 2017 Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/2167702617723376 www.psychologicalscience.org/CPS



Jean M. Twenge<sup>1</sup>, Thomas E. Joiner<sup>2</sup>, Megan L. Rogers<sup>2</sup>, and Gabrielle N. Martin<sup>1</sup>

San Diego State University and <sup>4</sup>Florida State University







Data from Twenge et al. (2017), Orben (2017)

Big Data – Small Effects





	I take a positive attiinde toward myself	I feel I am a person of worth, on an equal plane with others	I am able to do things as well as most other people	On the whole, I ann satisfied with myself	I feel I do not have much to be proud of	Sometim- es I think that I am no good at all	I feel that I can't do anything right	l feel that my life is not very useful	Life often seems meaning- less	l enjoy life as much as anyone	The future often seems hopeless	It feels good to be alive	How happy are you these days
Newcomb, Huba and Bentler (1986)													
Maslowsky, Schulenberg and Zucker (2014)													
Twenge, Joiner, Rogers and Martin (2017)													
Midgely and Lo (2013)**													
Denham (2009)													
Merline, Jager and Schulenberg (2008)													
Twenge, Martin and Campbell (2018)													
Twenge and Campbell (2008)*													
Trzesniewski and Donnellan (2010)													
Rosenberg (1965)													
O'Malley and Bachman (1983)													
Adams (2010)													

The Garden of Forking Paths





# Data that is "Too Big To Fail"

- Large numbers of participants ensure that even extremely modest covariations (e.g. r's < 0.05) between self-report items will result in alpha levels typically interpreted as compelling evidence for rejecting the null hypothesis by psychological scientists (i.e. p's < 0.05)</li>
- Large batteries of ill-defined questions lead to an explosion of possible analytical pathways (researcher degrees of freedom)





# Phase II

MRC Cognition and Brain Sciences Unit







Statistically Significant Result





Check for updates

**General Article** 

# 

#### False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Psychological Science. 22(11) 1359–1366 © The Author(s) 2011 Reprints and permission; sagepub.com/journalsPermissions.nav DOI: 10.1177/0956797611417632 http://pss.sagepub.com

#### Joseph P. Simmons<sup>1</sup>, Leif D. Nelson<sup>2</sup>, and Uri Simonsohn<sup>1</sup>

<sup>1</sup>The Wharton School, University of Pennsylvania, and <sup>2</sup>Haas School of Business, University of California, Berkeley

#### Abstract

In this article, we accomplish two things. First, we show that despite empirical psychologists' nominal endorsement of a low rate of false-positive findings ( $\leq$  .05), flexibility in data collection, analysis, and reporting dramatically increases actual false-positive rates. In many cases, a researcher is more likely to falsely find evidence that an effect exists than to correctly find evidence that it does not. We present computer simulations and a pair of actual experiments that demonstrate how unacceptably easy it is to accumulate (and report) statistically significant evidence for a false hypothesis. Second, we suggest a simple, low-cost, and straightforwardly effective disclosure-based solution to this problem. The solution involves six concrete requirements for authors and four guidelines for reviewers, all of which impose a minimal burden on the publication process.

#### Keywords

methodology, motivated reasoning, publication, disclosure





### Solution #1

Decide on one analytical pathway beforehand using pre-registration or registered report methodologies (Chambers, 2013; Munafò et al., 2017; van 't Veer, 2016; Lakens, 2014)

Pro: Simple way to decrease researcher degrees of freedom



http://blogs.discovermagazine.com/neuroskeptic/ 2013/10/16/the-f-problem/







# Preregistration

Study preregistrations on the Open Science Framework (OSF) are doubling every year; more than 120 journals have introduced registered reports.



J. YOU/SCIENCE





## Solution #1

Decide on one analytical pathway beforehand using pre-registration or registered report methodologies (Chambers, 2013; Munafo et al., 2017; van 't Veer, 2016; Lakens, 2014)

Pro: Simple way to decrease researcher degrees of freedom

Con: Researcher needs to prove that they have not previously seen or engaged with the data







### Solution #2

Examine all possible analytical pathways using Specification Curve Analysis (SCA; Simonsohn, Simmons, & Nelson, 2015)

Pro: Works around researcher degrees of freedom even when data has been previously accessed



Simmonsohn, Simmons, Nelson (2015)







Simmonsohn, Simmons, Nelson (2015)





# Where it all began...









International Edition +

# Extreme Reather

# Female hurricanes are deadlier than male hurricanes, study says



By Holly Yan, CNN

() Updated 2127 GMT (0527 HKT) September 1, 2016





A NC





## Female hurricanes are deadlier than male hurricanes

Kiju Jung<sup>a,1</sup>, Sharon Shavitt<sup>a,b,1</sup>, Madhu Viswanathan<sup>a,c</sup>, and Joseph M. Hilbe<sup>d</sup>

<sup>a</sup>Department of Business Administration and <sup>b</sup>Department of Psychology, Institute of Communications Research, and Survey Research Laboratory, and <sup>c</sup>Women and Gender in Global Perspectives, University of Illinois at Urbana–Champaign, Champaign, IL 61820; and <sup>d</sup>Department of Statistics, T. Denny Sanford School of Social and Family Dynamics, Arizona State University, Tempe, AZ 85287-3701

Edited\* by Susan T. Fiske, Princeton University, Princeton, NJ, and approved May 14, 2014 (received for review February 13, 2014)

Do people judge hurricane risks in the context of gender-based expectations? We use more than six decades of death rates from US hurricanes to show that feminine-named hurricanes cause significantly more deaths than do masculine-named hurricanes. Laboratory experiments indicate that this is because hurricane names lead to gender-based expectations about severity and this, in turn, guides respondents' preparedness to take protective action. This finding indicates an unfortunate and unintended consequence of the gendered naming of hurricanes, with important implications for policymakers, media practitioners, and the general public concerning hurricane communication and preparedness.

gender stereotypes | implicit bias | risk perception | natural hazard communication | bounded rationality

E stimates suggest that hurricanes kill more than 200 people in the United States annually, and severe hurricanes can cause fatalities in the thousands (1). As the global climate changes, the frequency and severity of such storms is expected to increase (2). violence and destruction (23, 24). We extend these findings to hypothesize that the anticipated severity of a hurricane with a masculine name (Victor) will be greater than that of a hurricane with a feminine name (Victoria). This expectation, in turn, will affect the protective actions that people take. As a result, a hurricane with a feminine vs. masculine name will lead to less protective action and more fatalities.

#### **Archival Study**

To test this hypothesis, we used archival data on actual fatalities caused by hurricanes in the United States (1950–2012). Ninetyfour Atlantic hurricanes made landfall in the United States during this period (25). Nine independent coders who were blind to the hypothesis rated the masculinity vs. femininity of historical hurricane names on two items (1 = very masculine, 11 = very feminine, and 1 = very man-like, 11 = very woman-like), which were averaged to compute a masculinity-femininity index (MFI). A series of negative binomial regression analyses (26, 27) were performed to investigate effects of perceived masculinity-femi-





#### LETTER

7

# Are female hurricanes really deadlier than male hurricanes?

Jung et al. (1) claim to show that "femininenamed hurricanes cause significantly more deaths than do masculine-named hurricanes" (p. 1). This conclusion is mainly obtained by analyzing data on fatalities caused by hurricanes in the United States (1950–2012). By reanalyzing the same data, we show that the conclusion is based on biased presentation and invalid statistics.

The reasoning in ref. 1 is fundamentally based on the regression models reported in their table S2, in particular, model 4. However, due to the interaction terms combined with extreme values and weak significance, the analysis is based on a very fragile model; e.g., the model predicts almost 20,000 deaths for hurricane Sandy, which actually caused 159 fatalities. Their figure 1 and the dis-

Now, we explain our claim that the results are presented in a biased way. By holding the minimum pressure at its mean in prediction of counts of deaths, the authors only report the influence of MFI and normalized damage (figure 1 in ref. 1). This ignores the influence of the second interaction term MFI minimum pressure, which shows an opposite influence (see the estimated parameters on p. 5, first paragraph). By considering the counts of deaths under constant normalized damage, the results are contrary: male-named hurricanes with a low minimum pressure (strong hurricanes) are associated with more deaths than female ones (Fig. 1).

In the light of an alternating male-female naming system started in 1979, the authors differences between male- or femalenamed hurricanes for deaths, minimum pressure, category, and damages. IEI

To conclude, the analyses given in ref. 1 are examples of the fact that prediction models using interaction terms have to be handled and interpreted carefully; in particular, using insignificant variables is not expedient and may lead to statistical artifacts.

To summarize, the data do not contain evidence that feminine-named hurricanes cause more deaths than masculine-named hurricanes.

#### Björn Christensen<sup>a</sup> and Sören Christensen<sup>b,1</sup>

"Institute for Statistics and Operations





# Population matters when modeling hurricane fatalities

Jung et al. (1) find gender bias leads to misperception of hurricane risk in both experimental and historical evidence. We affirm the rich literature on gender bias. However, we argue that Jung et al.'s empirical analysis suffers from endogeneity. Once addressed, we find the previous results to be of questionable robustness. Although gender bias may exist in limited-information experiments, historical evidence does not indicate gender bias in hurricane fatalities.

Damages do not determine deaths, but rather both are simultaneously determined by multiple factors, including hurricane characteristics and the (omitted) underlying population and vulnerability (2), which lead to endogeneity, or correlation between damages and the error term that can bias estimated

point of landfall. (No annual county-level data exist back to 1950. We calculated the 2000 county-to-country population density ratio and, assuming it is constant across time, we scaled the ratio by annual US population density. We recommend future work refine this assumption.) Third, we normalized deaths by (i) dividing deaths by real damages (nominal damages taken from the International Catastrophe Insurance Managers; deflator information taken from the Bureau of Economic Analysis; population taken from the US Census) and (ii) dividing deaths by total US population in the year of landfall (we normalized deaths by the five-county population and found similar results). After a log-transformation,

The experiments in Jung et al.'s study (1) are interesting but the motivational facts are of questionable robustness. We establish this finding by controlling for population and correcting for endogeneity. Further research on the subject of hurricane naming is therefore warranted and encouraged.

Laura A. Bakkensen<sup>a,1</sup> and William Larson<sup>b</sup> <sup>a</sup>School of Government and Public Policy, University of Arizona, Tucson, AZ 85721; and <sup>b</sup>Office of the Chief Statistician, Bureau of Economic Analysis, US Department of Commerce, Washington, DC 20230





### ETTER

ick for update

# Statistics show no evidence of gender bias in the public's hurricane preparedness

Jung et al. (1) make the bold claim that a storm's assigned gender (traditionally masculine vs. feminine name) predicts its destructive potential, such that a hypothetical Hurricane Eloise would have three times the expected death toll compared with a hypothetical Hurricane Charley, by 41 deaths to 15 deaths. They say that feminine names and pronouns are perceived by the public to be less threatening, resulting in lax preparations and fewer evacuations. Their conclusion was widely reported in the popular press.

During 30+ years on the Gulf Coast, my hurricane evacuations show no gender bias (Andrew, Lili, Rita, and Gustav). I am skeptical that sexism explains the noted effect.

Ninety-four hurricanes made landfall in

deaths raises the average death toll for all female storms by 85%, to 23.5 (n = 63 deaths; maximum, 256 deaths). Hurricanes Diane (1955), Camille (1969), Agnes (1972), and Sandy (2012) collectively accounted for 732, or 38.5%, of 1,900 total deaths in the study. Virtually all of the statistical difference between the deadliness of female vs. male storms is explained by the inclusion of these four events.

Surprisingly, of those 732 deaths, at least 327 (45%) occurred well inland, mostly due to flooding and landslides in the mountain valleys of the Appalachians as the largely spent storms dropped torrential rains. Diane's deluge hit a region already rain soaked from Hurricane Connie a mere 5 d before. An es- Steve Maley 

Setting aside the issue of outliers, what are the odds that the six deadliest hurricanes of the study period would all bear feminine names?

If all of the names had been assigned randomly or alternately, the correct answer would be 1 in  $2^6$ , or about 1.6%.

However, four of the six storms occurred during a period (1953-1978) when only female names were assigned. Only Katrina and Sandy had a chance to be "male." The chance that two randomly chosen hurricanes would both have female names is 1 in 2<sup>2</sup>, or 25%: that's not bias, it's a coincidence, and not a strong one at that.

MRC Cognition and Brain Sciences Unit





### LETTER

# Female hurricanes are not deadlier than male hurricanes

Jung et al. (1) assert that hurricanes that made landfall in the United States killed more people when they had female names rather than male names. The article has stirred much controversy. Criticisms range from the inclusion of hurricanes from the era before they were given male names (2), over the selective interpretation and the overstatement of their results from the archival study in favor of their hypothesis (3), to the external validity of their six behavioral experiments for at-risk populations in at-risk situations (4).

The criticism of this letter is a different one: the results of their archival study are a function of the selective inclusion of regrestolls had smaller death tolls when the hurricanes were strong (lower pressure), but higher death tolls when the hurricanes were weak (higher pressure). The latter result is driven by the pre-1978 sample (model 5). In the post-1978 sample, the interaction effect becomes insignificant and the damage toll has a positive and significant relationship with the death toll (models 6 and 7).

Like the death toll, the damage toll is a simultaneous outcome of the storm and hence not a good explanatory variable. It merely reflects other underlying characteristics that could range from the size of the hurricane or its area of effect over the assets at risk and the safety infrastructures and more reflective of other characteristics for weaker storms and after 1978. Even though a lower importance of safety infrastructures during weaker storms and an overall improvement in or a convergence of safety infrastructures after 1978 might explain these results, the ambiguity of the death toll variable disallows any strong or definitive interpretation.

#### Daniel Malter<sup>1</sup>

Strategy Unit, Harvard Business School, Harvard University, Boston, MA 02163





# **Specification Curve Analysis**

**Table 1.** Original and alternative reasonable specifications used to test whether hurricanes with more feminine names were associated with more deaths.

Decision	Original Specifications	Alternative Specifications			
1.Which storms to analyze	Excluded two outliers with the most deaths	Dropping fewer outliers (zero or one); dropping storms with extreme values on a predictor variable (e.g., hurricanes causing extreme damages)			
2.Operationalizing hurricane names' femininity	Ratings of femininity by coders (1-11 scale)	Categorizing hurricanes names as male or female			
3.Which covariates to include	Property damages in dollars interacted with femininity; minimum hurricane pressure interacted with femininity	Log of dollar damages; year; year interacted with damages			
4. Type of regression model	Negative binomial regression	OLS with log(deaths+1) as the dependent variable			
5.Functional form for femininity	Assessed whether the interaction of femininity with damages was greater than zero	Main effect of femininity; interacting femininity with other hurricane characteristics (e.g., wind or category) instead of damages			

Simmonsohn, Simmons, Nelson (2015)







Simmonsohn, Simmons, Nelson (2015)

specifications.





# Does digital technology use influence adolescent well-being?

3 Datasets collected between 2007 and 2016

- Youth Risk and Behavior Survey (YRBS; Kann et al., 2016)
- Monitoring the Future (MTF; Johnston, Bachman, O'Malley, Schulenberg, & Miech)
- Millennium Cohort Study (MCS, University of London. Institute for Education. Centre for Longitudinal Studies., 2017)

Orben and Przybylski (Nature Human Behaviour, 2019)







Orben and Przybylski (Nature Human Behaviour, 2019)

#### 2 Implementing Specifications







Orben and Przybylski (Nature Human Behaviour, 2019)

#### 2 Implementing Specifications







Orben and Przybylski (Nature Human Behaviour, 2019)

#### 2 Implementing Specifications







Orben and Przybylski (Nature Human Behaviour, 2019)

#### 2 Implementing Specifications





#### ARTICLES

#### NATURE HUMAN BEHAVIOUR



**Fig. 5 | Comparison specifications for MCS.** Visualization of the comparison specifications hypothesized to have little or no influence on well-being: bicycle use, height, handedness and wearing glasses. This graph shows SCA for both the variable of interest (mean technology use) and the comparison variables; it highlights the range of possible results of a simple cross-sectional regression of the variables of interest on adolescent well-being. Wearing glasses has the most negative association with adolescent well-being (black, median  $\beta = -0.061$ , median n = 7,963, partial  $\eta^2 = 0.005$ , median standard error = 0.010); and more negative than the association of technology use with well-being (purple, median  $\beta = -0.042$ , median n = 7,964, partial  $\eta^2 = 0.002$ , median standard error = 0.010). Handedness (red/purple, median  $\beta = -0.004$ , median n = 7,972, partial  $\eta^2 < 0.001$ , median standard error = 0.010), height of the adolescent (red, median  $\beta = 0.065$ , median n = 7,910, partial  $\eta^2 = 0.005$ , median standard error = 0.010) and whether the adolescent often rides a bicycle (yellow, median  $\beta = 0.080$ , median n = 7,974, partial  $\eta^2 = 0.007$ , median standard error = 0.010) have more positive associations with adolescent well-being than does technology use. **a**, How different analytical decisions (specifications, shown on the x axis) lead to different statistical outcomes (standardized regression coefficient, shown on the y axis). Each line represents a different variable of interest while the error bars represent the standard error. **b**, The resulting median standardized regression coefficients for those SCAs linking the variables of interest with adolescent well-being.





#### Digital Technologies Effects on Adolescent Well-Being = Negative

- = Significant
- = Minimal
- = Worthy of intense public debate?



# SCA + Preregistration

GUI/Ireland, PSID/USA, MCS/UK; 17,314 adolescents Digital Screen Use (before bed) correlated with Well-Being and Self-Esteem 44 - 66 Specifications






SCA + Preregistration

Orben and Przybylski (2019, Psychological Science)





**Table 3.** Results of the Specification-Curve-Analysis BootstrappingTests for Confirmatory Tests

Technology measure	Median point estimate			Share of significant results in predominant direction	
	β	Partial $r^2$	p	Number	p
Self-report	-0.08 [-0.10, -0.07]	.008 [.006, .011]	.00	4	.00
Time spent	-0.02 [-0.04, -0.01]	.001 [.000, .002]	.00	5	.00
Less than 30 min on weekday	0.03 [0.01, 0.05]	.001 [.000, .003]	.00	3	.00
Less than 1 hr on weekday	0.02 [0.01, 0.04]	.001 [.000, .003]	.02	1	.27

Note: Values in brackets are 95% confidence intervals.







#### Digital Technologies Effects on Adolescent Well-Being = Negative

- = Significant
- = Minimal
- = Driven by self-report





## Phase III

MRC Cognition and Brain Sciences Unit





# Phase III



12,672 adolescents (10-15 years old)Median 1,699 per specification2,268 Specifications





#### Random Intercepts Crosslagged Panel Model

Dissociating between- and withinperson effects (Kievit et al. 2013)



9









1





Orben, Dienlin and Przybylski (PNAS, 2019)





## Phase IV

MRC Cognition and Brain Sciences Unit

3



Preferred Model: Mean lifesatisfaction, social media use, all covariates, WLSMV, imputed data, 4 waves













#### 2. Social Media Use -> Life Satisfaction Gender: Female









### 3. Life Satisfaction -> Social Media Use Gender: Female

Orben, Dienlin and Przybylski (PNAS, 2019)



7



Digital Technologies Effects on Adolescent Well-Being = Negative

- = Significant
- = Bi-directional
- = Gender Specific





# Phase IV





MRC Cognition and Brain Sciences Unit







#### Looking forward ...

Changing the question & Pushing for transparency & Getting better data