The Wolves in Sheep's Clothing: How Russia's Internet Research Agency Tweets Appeared in U.S. News as Vox Populi

Appendix A: Selection of top 100 IRA Accounts

We collected from the 1% Twitter archive all tweets from 2015-2017 that either originated from the 2,752 IRA Twitter handles or contained their handles in the tweet text. This gives us 460,143 tweets, the majority (351,501) of which are "retweets" starting with RT or RT@ and only a few (108,642) are non-retweets. Based on the retweet relationship, 239,616 unique edges were created (e.g., if @A retweets @B, then A -- B) containing 168,649 unique nodes/handles. We then sorted the handles by in degree (the times being retweeted by any handle) and selected the top 100 handles. These handles were best-connected in the sense that they received the most retweeting and had their messages spread presumably most far and wide in the Twitter sphere.

Table 1 below displays the 30 handles with the highest in-degree (the times being retweeted by any handle), along with other degree information such as out-degree and total-degree (calculated the same way). For a full list of 100 accounts, please visit this link: https://osf.io/vrbep/?view_only=a5a7b1541f364c23ba56b258bf5ccd84

handle	total_degree	in_degree	out_degree
ten_gop	59816	59806	10
crystal1johnson	35909	35905	4
pamela_moore13	22533	22531	2
coldwar20_ru	21801	21605	196
lavrovmuesli	19137	18991	146
comradzampolit	12516	12395	121
kadirovrussia	11897	11892	5
margosavazh	8486	8455	31
gloed_up	8552	8313	239
riafanru	7665	7665	0
novostidamask	6673	6673	0
jenn_abrams	6379	6371	8
zubovnik	5949	5912	37
trayneshacole	5954	5863	91
southlonestar	5622	5621	1
blacktolive	4689	4676	13
blacknewsoutlet	4643	4643	0
thefoundingson	4571	4557	14
maxdementiev	4571	4529	42
bydrbre	4487	4485	2
russianallies	2940	2938	2
bleepthepolice	2927	2827	100

usa_gunslinger	2661	2657	4
todayinsyria	1587	1587	0
tpartynews	1422	1419	3
pamelakealer13	1219	1219	0
jeblary2016	1171	1171	0
10_gop	1088	1088	0
nevnov_ru	1085	1085	0
warfareww	1036	1036	0
boeing_is_back	932	932	0
bydrbre_	928	928	0
pigeontoday	820	820	0

Appendix B: Number of Articles with IRA Tweets, by Outlet

	Outlet	Media	# of articles		Outlet	Media	# of articles
1	The Inquisitr News	Digital First	29	40	Quartz	Digital First	2
2	The Daily Caller	Digital First	20	41	Refinery29	Digital First	2
3	Breitbart	Digital First	13	42	al.com	Digital First	1
4	Right Scoop	Digital First	13	43	Boston Herald	Print	1
5	Daily Kos	Digital First	12	44	colorlines.com	Digital First	1
6	The Blaze	Digital First	12	45	Daily Beast	Digital First	1
7	IJR	Digital First	10	46	Daily Telegraph	Print	1
8	The Washington Post	Print	10	47	Detroit Free Press	Digital First	1
9	Elite Daily	Digital First	9	48	DNA Info	Digital First	1
10	Yahoo News	Digital First	9	49	Forbes	Print	1
11	Huffington Post	Digital First	8	50	Grist	Digital First	1
12	USA Today	Print	8	51	Mic	Digital First	1
13	Fox News	Broadcast	7	52	mintpressnews.com	Digital First	1
14	mlive	Digital First	7	53	Mother Jones	Digital First	1
15	Russia Today	Digital First	7	54	New York Times	Print	1
16	The Epoch Times	Digital First	7	55	Newsone	Digital First	1
17	BuzzFeed	Digital First	6	56	Newsweek	Print	1
18	Independent	Digital First	6	57	NPR	Radio	1
19	Mashable	Digital First	6	58	NY Daily News	Print	1
20	NBC	Broadcast	6	59	NY Post	Print	1
21	nj.com	Digital First	6	60	PBS NewsHour	Broadcast	1
22	uproxx.com	Digital First	6	61	Slate	Digital First	1
23	IBTimes	Digital First	5	62	The Guardian	Print	1
24	The Hill	Print	5	63	The Record (NJ)	Print	1
25	PJ Media	Digital First	4	64	The Verge	Digital First	1
26	Raw Story	Digital First	4	65	Time	Print	1
27	Salon	Digital First	4	66	Upworthy	Digital First	1
28	SFGate	Digital First	4	67	Vice	Digital First	1
29	US News & World Report	Print	4	68	Vox	Digital First	1
30	Business Insider	Digital First	3	69	Wall Street Journal	Print	1
31	CBS	Broadcast	3	70	Washington Examiner	Print	1
32	CNN	Broadcast	3	71	Washington Times	Print	1
33	Conservative Treehouse	Digital First	3	72	Alternet	Digital First	0
34	Market Watch	Digital First	3	73	Americal Journal Review	Digital First	0
35	ABC	Broadcast	2	74	American News	Digital First	0
36	Daily Mail	Print	2	75	bb4sp.com	Digital First	0
37	Gothamist	Digital First	2	76	BBC	Digital First	0
38	LA Times	Print	2	77	Clash Daily	Digital First	0
76	BBC	Digital First	0				More on next page.

78	common dreams	Digital First	0
79	Conservative Daily Post	Digital First	0
80	Conservative Fighters	Digital First	0
81	Daily Wire	Digital First	0
82	DC Clothes Line	Digital First	0
83	DC Whispers	Digital First	0
84	Dead State	Digital First	0
85	Free Beacon	Digital First	0
86	Gateway Pundit	Digital First	0
87	HeatStreet	Digital First	0
88	Hot Air	Digital First	0
89	Information Liberation	Digital First	0
90	InfoWars	Digital First	0
91	joeforamerica.com	Digital First	0
92	mediaite.com	Digital First	0
93	Nation of Change	Digital First	0
94	National Review	Digital First	0
95	newsbreakshere.com	Digital First	0
96	newsfrompolitics.com	Digital First	0
97	People	Print	0
98	Politico	Digital First	0
99	Prison Planet	Digital First	0
100	Real Clear Politics	Digital First	0
101	Red State	Digital First	0
102	Reverb Press	Digital First	0
103	rinf.com	Digital First	0
104	Rolling Stone	Print	0
105	Talking Points Memo	Digital First	0
106	The Political Insider	Digital First	0
107	Think Progress	Digital First	0
108	townhall.com	Digital First	0
109	truthfeed.com	Digital First	0
110	twitchy	Digital First	0
111	US Defense Watch	Digital First	0
112	US Herald	Digital First	0
113	Viral Patriot	Digital First	0
114	Web Daily	Digital First	0
115	worldnetdaily	Digital First	0
116	worldpoliticus.com	Digital First	0
117	Zero Hedge	Digital First	0

Appendix C: Codebook

Basic Information

(A)	Is the article about sockpuppets? [RELEVANT, FK: 0.83] *(A) **Cases should be excluded** when the search term picks up an irrelevant story (not related to the handle), or when the handle is discussed as "being discovered."* *(B) Cases which include Facebook pages/groups with the same name as the handle should be kept, but **notated** (you may not be able to answer all the questions)*	1 = Relevant; 0 = Not Relevant	***If it is not relevant, please stop coding here.***

Content of Article

	I. Topic of the News Story		
1	Is the story about the election? [T_elect, FK: 0.88]	1 = Yes; 0 = No	e.g., polls, campaign stops, etc.
2	Is the article talking about a specific candidate? [T_can, FK: 0.82]	1 = Trump 2 = Clinton; 3 = Other candidates; 99 = N/A	
3	Is the story talking about a political issue? (includes int'l and domestic issues) [T_issue, FK: 0.71]	1 = Yes; 0 = No	e.g., LGBTQ issues, taxes, guns, police brutality, election
3a	If the story is talking about a political issue, is the story about "policy" or "issues" political issues? [T_issue_type, FK: 0.88] (1) If the article references campaign events, a politician, party, government agency, or specific policy, code as "policy" (2) If the article is about a protest, a hashtag-call-to-action boycott, racial injustice (none of which relate to a policy) code as "issues" (3) Remarks about politicians in non-government settings should be coded as "soft"	1 = policy 0 = issues 99 = not political	

4	In three words, what political issue is discussed? [T_issue2]	(words)	
5	Is the story about a Twitter topic? [T_trend, FK: 0.87] (A) If people are discussing a topic on Twitter, this constitutes a Twitter topic. (B) If the topic is a news story (e.g., shooting at concert), and there happens to be Twitter responses, the story is not about a Twitter topic	1 = Yes; 0 = No	e.g., the article describes a "discussion" or "debate" on Twitter (includes stories that happened offline, that Twitter responded)
6	Does the story use string-of-tweets reporting? [T_motw; FK: 0.82] (A) If a story is a list of tweets (e.g. "these 12 tweets will…"), code as "yes" (often identifiable in headline) (B) If a story has more than 3 tweets in a row, code "yes" (C) Hypothetically, if you remove the tweets, how much information is left? If there is not enough information to understand the story, code as "0"	1 = Yes; 0 = No	There should be no journalistic content between the "string of tweets"
7	If not a political topic, please describe the topic of the article in one or two words [T_misc]	(words)	
8	Does the article show disagreement (i.e., competing perspectives on an issue/topic)? [T_sides, FK: 0.72] (A) Differing perspectives are not limited to tweets (B) Code 1 (yes) even if only one paragraph or tweet shows a different viewpoint	1 = Yes; 0 = No	
9	Do all the tweets in this article have congruent political ideology? [T_congruent, FK: 0.75] (A) Make inferences based on the content of the tweet, it is okay to check externally (but avoid this if possible)	1 = all lib 2 = all con 0 = mixed tweets 99 = apolitical	e.g., "The hashtag also prompted a vigorous backlash, with people critical of the refugee ban co-opting the hashtag..."
	II. Tweet Use		
10	How many tweets are in the article? [TW_n, FK: 0.92]	(number)	
11	Are other people (**not tweets**) interviewed, quoted, paraphrased, or sourced in the news story? [TW_other, FK: 1.0]	1 = Yes 0 = No	
	III. Post-Hoc Sockpuppet Use		

#	Question	Coding	Notes
12	Does the article include a post-hoc correction about the sockpuppet? [Disclosure, FK: 0.83]	1 = Yes 0 = No	
13	Has the sockpuppet been removed from the story? [Removal, FK: 0.90] (A) If you had to use the wayback machine to find the article, code this as 1 (yes)	1 = Yes 0 = No	

Article's Use of Specific Tweet

#	Question	Coding	Notes
14	What number tweet is it in the story? [ATW_n, FK: 1.0]	(number)	Note if the sockpuppet is listed as the "first" of a set of opinions.
	I. How Is the Tweet Being Described?		
15	Is the tweet handle mentioned aside from the use of the tweet? [ATW_handle, FK: 0.71] (A) If a tweet is "dropped" into a news article without mention of it before or after, code this as a 0 (No)	1 = yes; 0 = no	*If a tweet is dropped in, 17, 18, and 19 should be "0"*
16	Is the tweet being described as originating from authoritative source? [ATW_aut, FK: 0.75] (A) An authoritative source is popular or representative of some group (e.g., a "prominent" handle)	1 = yes 0 = no	
17	Is the tweet being described as a "man on the street"? [ATW_reg, FK: 0.71]	1 = yes 0 = no	If the article broadly describes a group of tweets as "coming from the people" without specific mention of the handle still code as "yes"
	II. What Use Is the Outlet Making of the Tweet		
18	Is the sockpuppet tweet a source of information? [ATW_info, FK: 0.71]	1 = yes 0 = no	
19	Is the sockpuppet tweet the primary or sole information of a claim? [ATW_info_prim, FK: 0.74]	1 = yes 0 = no	
20	Is the sockpuppet tweet a source of opinion (not providing information, but providing commentary?) [ATW_op, FK: 0.73]	1 = yes 0 = no	
21	Is the journalist describing the sockpuppet positively or negatively? [ATW_tone, FK: 0.80] (1) If the article broadly describes "comments from Twitter," with a positive/negative tone (example: "Here are some of the best	1 = positive 0 = nothing -1 negative	Reputable sources are likely to be "neutral"

	questions/comments we received from Breitbart reader")		

Content of Embedded Sockpuppet Tweet

22	Is the tweet in support of a liberal, or conservative ideology, or is it mixed? [STW_pol, FK: 0.78]	1 = liberal 2 = conservative 0 = mixed	
23	Does the tweet include deceptive (incorrect) information? [STW_false, FK: 0.89] (A) This will require you to look elsewhere (outside of the article); do not spend significant time if unverifiable (B) This specifically refers to **factual** information	1 = yes 0 = no or unknown	*If unable to verify, write "N/A" and include a note*
24	Does the tweet try to "troll" people? [STW_troll, FK: 0.86] (A) To troll is to intentionally antagonize someone online	1 = yes 0 = no	
25	Does this tweet have a call to action? [STW_action, FK: 0.85]	1 = yes 0 = no	e.g., "The hashtag #BoycottStarbucks was trending Monday on Twitter"
26	Does this tweet use a hashtag? [STW_#]	1 = yes 0 = no	
27	If so, what is the hashtag? [STW_#2]	(hashtags)	
28	Is the tweet sarcastic? [STW_sarc, FK: 0.82]	1 = yes 0 = no	
N/A	Are there any other notes you would like to include? [Notes]		

Appendix D: Stratified random sample tweets that were embedded in news stories

	IRA handle (*Ideology*)	Tweet	Publication(s)
1	@Crystal1Johnson (L)	The Wade quadruplets, Nigel, Zach, Aaron and Nick, all 18, have all been accepted to both Harvard and Yale	Elite Daily
2	@trayneshacole (L)	African American Woman who took a oath to serve & protect | I support this movement #VeteransForKaepernick pic.twitter.com/7SYSt1DXLn	CBS
3	@blacktolive (L)	They didn't associate themselves with Black Lives Matter mvmnt. Haters are just trying to clutch at 'evil black people' straw#BLMKidnapping	The Verge
4	@BlackNewsOutlet (L)	Two students were attacked Saturday night at the USM vs. Rice University football game for refusing to stand for National Anthem	Raw Story
5	@gloed_up (L)	#AllLivesDidntMatter when they didn't include the rest of American history in books to only teach us about white American contributions.	Huffington Post
6	@TEN_GOP (C)	RT if you also think Lester Holt did a bad job moderating tonight! He didn't mention: her lies, Benghazi, Emails! #debatenight #Debates	Yahoo News
7	@pamela_moore13 (C)	Hiring 10k refugees makes liberals feel warm BUT we have homeless vets that need those jobs.#BoycottStarbucks #ReasonsToProtest #MuslimBan	Fox News
8	@southlonestar (C)	Muslim woman pays no mind to the terror attack, casually walks by a dying man while checking phone#PrayForLondon #Westminster #BanIslam	*Washington Post*
9	@TEN_GOP (C)	Massive crowd waiting outside for the Trump rally in Phoenix!	Daily Beast; Inquisitr News; *USA Today*
10	@usa_gunslinger (C)	Libs, do you find this picture offensive? #HeterosexualPrideDay	The Guardian

Ideology: L = Liberal, C = Conservative [Five Liberal tweets and five Conservative tweets were randomly sampled]