

Supplemental Information For:

Investigating the Generation and Spread of Numerical Misinformation:

A Combined Eye Movement and Social Transmission Approach

Study 2a

In order to obtain our target sample size of 300, we recruited a larger pool of participants on Amazon Mechanical Turk given that we expected that some participants would not meet our study criteria. In particular, we recruited a total of 432 participants who qualified for our study, but 132 of participants were removed using specific criteria to ensure quality of the data. Of the 132, 101 were removed because their combined numerical estimates exceeded the maximum threshold stated in the blog posts. For example, a participant who estimated the proportion of Americans who prefer male bosses at 75% and 50% for female bosses would be removed (since the values need to sum to 100% or less). We removed such cases because they can introduce a potential confound when the values are transmitted to other people in the chains (e.g., receivers may direct attention to the fact that the values exceed 100%). In Supplementary Table 11 below, we analyzed the data including these 101 excluded participants. Inclusion of these participants did not change our substantive results (see Supplementary Figure 1). The remaining 31 cases were instances in which individuals did not fill out all questions in our memory task or provided the incorrect format for our numerical values (e.g., provided percentage estimates for questions on the number of Mexican immigrants in the country). We also stopped data collection after we reached our 300-participant limit.

Study 2b

The primary reason we conducted Study 2b was to ensure that the observed effects in Study 2a were not due to our exclusion criteria of participants. Specifically, Study 2b was designed to prevent individuals from entering numerical values over 100% (same-sex marriage and boss issues) or 963 (police shooting issue) given that most of our participant exclusions in Study 2a were based on these errors (101 out of 132 rejections). In Study 2b, participants were asked to re-

enter numerical values if they summed above 100% (for the same-sex marriage and boss issue) or 963 (for the police shooting issue). We also used sliders instead of open-ended response boxes to ensure that individuals were aware of the numerical format (percentages, integers). Similar to Study 2a, we also recruited a larger pool of participants from the NIH's *ResearchMatch* service in order to obtain our target sample size of 300. Of the 332 participants who qualified for our study, we removed 31 because they took an unusual amount of time to complete our study (e.g., more than one hour), and one participant completed the study in less than three minutes.

Supplementary Exploratory Analyses

Supplementary Tables 4 to 7 show potential interactions between schema consistency and partisan ID in each of the four issues in Study 1. We estimated separate models for each issue given that the direction of the interaction might differ depending on the issue. Our exploratory analyses did not reveal any statistically significant schematic consistency by partisan ID interactions. One possibility for lack of a statistically significant interaction is that both Democrats and Republicans in our sample possess similar schemas of the issues. Another possibility is that our study is not sufficiently powered to detect such interactions. Disentangling these different possibilities is important for future work to explore.

Supplementary Table 8 shows the potential interaction between schema consistency and the factual accuracy of the pairwise gist relationships in Study 1. For example, in the Mexican immigration issue, the schema-inconsistent version (e.g., 12.8 million for 2007 and 11.7 for 2014) is coded as the factually-accurate numerical relationship. In contrast, the schema-consistent version (e.g., 11.7 million for 2007 and 12.8 for 2014) is coded as the factually-inaccurate numerical relationship. Our exploratory analyses revealed no statistically significant interaction between schema consistency and factual accuracy.

In Supplementary Table 11, we analyzed the Study 2a data including the 101 participants who were excluded because their responses summed above 100% (for the same-sex marriage and gender preference for bosses issues) or 963 (for the police shootings issue). Inclusion of these participants did not change our substantive results (also see Supplementary Figure 1).

Supplementary Table 1

Paragraphs Used as Foils

The number of homeless in the United States stands at 564,708 people as of 2016. This study defined homelessness as: "living on the streets, in cars, homeless shelters, or in subsidized transitional housing". Of that over half a million people, about a quarter of them are children.	An Indian company which plans to build a power plant that runs on city waste will aim for a \$200 million IPO next month, according to a recent report from Bloomberg.
According to PEW Research Center, gun ownership by households in the United States stands at 44%. This represents a 7-percentage point increase in the past two years. This is still lower than it was in the 1970s.	If you are thinking about tweeting about clouds, pork, exercise, think again. The Department of Homeland Security has been forced to release a list of keywords and phrases it uses to monitor various social networking sites.
According to preliminary reporting, the 2016 presidential election saw the lowest rate of voter turnout in 20 years. About 55.4% of eligible voters casted their ballots in the past election, which was down from about 60% in 2012	Vulture reports this morning that the Gremlins horror movie about creatures who can't be fed after midnight is perhaps on its way to a reboot. Steven Spielberg, who executive produced the original film has reportedly kept previous remake attempts from becoming reality/
The national debt currently is at \$19.9 trillion. Although many believe a massive portion of the U.S. debt is owned by China, the country is second at \$1.049 trillion to Japan, who owns \$1.108 trillion American debt.	In recent months, the Food and Drug Administration has begun examining the safety of energy drinks following reports of several deaths and numerous injuries potentially associated with the products.
A Gallup survey from 2016 revealed which countries Americans view most favorably. Canada tops the list. The next four favorites were Great Britain, France, Germany, and Japan. The bottom of the list includes Iran, Syria, and North Korea.	The discovery of horse DNA in hamburgers on sale at supermarkets in Ireland and Britain is testing the appetite of meat lovers there. The Food Safety Authority of Ireland said that 10 out of 27 hamburger products it analyzed were found to contain horse DNA.
A scene captured by a Google Street View vehicle in Kweneng, Botswana, that went viral seems to show the car hitting and possibly killing a donkey on the road. But Google insists the animal is fine.	Last year the US was ranked 10th place in the list of the World's Happiest Countries. This year the U.S.A. has slipped to 12th. This marks the first time that America is not in the top 10.
A Cheesecake Factory pasta dish with more than 3,000 calories - or more than a day and a half of the recommended caloric intake for an average adult - is one of the most unhealthy dishes at U.S. chain restaurants.	A new study published in Administrative Science Quarterly shows how employees' wages change immediately after a male chief executive officer has a child. It found that when a male CEO has a child, his employees' wages decrease.
Recently the Pew Research Center ran a report about smartphone use in the United States. 64% of American adults were smartphone owners as of Spring 2015. This is a staggering number, when you realize that in 2011 that statistic was only 35%.	Three major scientific projects set out this season to seek evidence of life in lakes deep under the Antarctic ice - evidence that could provide clues in the search for evidence of life elsewhere in the solar system such as under the surface of Enceladus, one of Saturn's moons.

Supplementary Table 2

Attention and Memory Accuracy Analyses – Controls Included

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
Schema-Consistency (inconsistent = 1, consistent = 0)	.60** (.19)	117.96* (52.93)	.33* (.14)	.16 (.23)	-1.16*** (.30)
Age	-.03 (.02)	-6.08 (5.17)	-.02 (.01)	.01 (.02)	-.02 (.02)
Female	.46+ (.26)	127.74+ (68.30)	.25 (.19)	.08 (.27)	.04 (.30)
Political Knowledge	.03 (.04)	5.48 (10.00)	.02 (.03)	.04 (.04)	.08+ (.04)
Party ID	.10 (.09)	10.17 (22.99)	.08 (.06)	.22* (.09)	.20* (.10)
First Number Fixation	-.03 (.23)	10.16 (64.55)	.05 (.17)	.17 (.28)	-.09 (.34)

Note. Mixed-effects regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Mixed-effects logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). Party ID (Strong Democrat = 6, Democrat = 5, Independents who identify as closer to Democratic Party's position = 4, Independents who identify as closer to Republican Party's positions = 3, Republican = 2, Strong Republican = 1). Political knowledge score is based on a 20-item questionnaire. First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Supplementary Table 3

Issue Type by Wave Interaction for Studies 2a and 2b with Controls

	Study 2a	Study 2b
Issue Type x Wave	.74*** (.18)	1.13*** (.17)
Wave	-.41** (.13)	-.50*** (.13)
Issue Type	-2.55*** (.25)	-2.55*** (.24)
Age	.008 (.007)	.004 (.004)
Female	-.02 (.13)	.40** (.14)
Interest in Politics	-.02 (.08)	-.08 (.07)
Party ID	-.03 (.04)	.02 (.04)

Note. Logistic regression coefficients for Study 2a and Study 2b with standard errors clustered on participants. Please note that we did not use the 20-item political knowledge questionnaire from Study 1 in Study 2. We asked participants to self-report their level of interest in politics (Not Interested = 1, Very Interested = 4) and used this as a control variable instead. In Study 1, the scores on the 20-item political knowledge questionnaire is correlated with participants' self-reported interest in politics ($B = 2.01$, $SE = .29$, $p < .001$). Finally, estimating logistic regression model (with standard errors clustered on participants) that combines both Studies 2a and 2b and exploring the possibility of Wave x Issue Type x Study Type (Study 1 = 0, Study 2 = 1) interaction yields a non-significant interaction ($B = .36$, $SE = .27$, $p = .17$). ** $p < .01$, *** $p < .001$.

Supplementary Table 4

Attention and Memory Accuracy Analyses for Schema Consistency x Party ID interaction – Gender Preference for Bosses Issue

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Schema-Consistency X Party ID	-0.22 (.28)	-123.22 (82.09)	-.17 (.20)	.08 (.07)	.87 (.92)
Schema Consistency	1.42 (1.30)	550.96 (384.31)	1.02 (.94)	-.40 (.31)	-6.51 (5.04)
Party ID	.06 (.20)	41.60 (59.24)	.07 (.15)	-.02 (.05)	-.84 (.90)
Age	-.02 (.03)	-1.26 (9.54)	-.003 (.02)	.02* (.008)	.007 (.05)
Female	.57 (.43)	163.28 (125.82)	.30 (.31)	-.04 (.10)	0.47 (.65)
Political Knowledge	-.07 (.07)	-24.54 (19.36)	-.09* (.05)	-.009 (.02)	.14 (.10)
First Number Fixation	-1.84*** (.48)	-454.27** (140.61)	-1.52*** (.35)	.06 (.11)	-.09 (.73)

Note. Regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). Party ID (Strong Democrat = 6, Democrat = 5, Independents who identify as closer to Democratic Party's position = 4, Independents who identify as closer to Republican Party's positions = 3, Republican = 2, Strong Republican = 1). First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). ⁺ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Supplementary Table 5

Attention and Memory Accuracy Analyses for Schema Consistency x Party ID interaction – Number of Immigrants from Mexico
2007/2014 Issue

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Schema-Consistency x Party ID	-.22 (.24)	-85.84 (55.86)	-.17 (.20)	-.03 (.07)	.54 (.37)
Schema Consistency	1.37 (1.14)	445.49+ (265.12)	.83 (.93)	.09 (.33)	-3.37+ (1.85)
Party ID	.22 (.17)	41.63 (38.43)	.20 (.14)	.07 (.05)	-.29 (.30)
Age	.008 (.03)	3.27 (6.40)	.0007 (.02)	-.007 (.008)	-.02 (.04)
Female	.06 (.37)	10.80 (84.94)	-.09 (.30)	-.08 (.11)	-.28 (.50)
Political Knowledge	.0007 (.05)	-8.86 (12.59)	.001 (.04)	.02 (.02)	.04 (.07)
First Number Fixation	.61 (.46)	106.74 (107.10)	.43 (.38)	-.12 (.13)	-.31 (.64)

Note. Regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). Party ID was coded as Strong Democrat = 6, Democrat = 5, Independents who identify as closer to Democratic Party's position = 4, Independents who identify as closer to Republican Party's positions = 3, Republican = 2, and Strong Republican = 1. First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). + $p < .10$.

Supplementary Table 6

Attention and Memory Accuracy Analyses for Schema Consistency x Party ID interaction – Number of White/Black Individuals killed by Police Officers in 2016

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Schema-Consistency x Party ID	.25 (.37)	91.11 (96.81)	.28 (.25)	-.11* (.05)	-.71 (.60)
Schema Consistency	-.13 (1.78)	-156.04 (464.10)	-.68 (1.18)	.49* (.23)	2.25 (2.37)
Party ID	.006 (.30)	-27.19 (77.73)	-.06 (.20)	.13** (.04)	.77 (.52)
Age	-.004 (.04)	-2.27 (11.27)	-.002 (.03)	-.001 (.006)	.01 (.08)
Female	.65 (.57)	124.90 (149.95)	.27 (.38)	.01 (.08)	.73 (.89)
Political Knowledge	.03 (.08)	7.59 (20.84)	.02 (.05)	.005 (.01)	.09 (.12)
First Number Fixation	1.17* (.56)	326.20* (146.06)	1.08** (.37)	.02 (.07)	.54 (.86)

Note. Regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). Party ID (Strong Democrat = 6, Democrat = 5, Independents who identify as closer to Democratic Party's position = 4, Independents who identify as closer to Republican Party's positions = 3, Republican = 2, Strong Republican = 1). First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). * $p < .05$, ** $p < .01$.

Supplementary Table 7

Attention and Memory Accuracy Analyses for Schema Consistency x Party ID interaction – Level of Support/Opposition to Same-Sex Marriage

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
Schema-Consistency x Party ID	.09 (.25)	47.37 (72.26)	.06 (.18)	-.08 (.06)	.15 (.45)
Schema Consistency	-.17 (1.18)	-149.31 (337.64)	-.14 (.84)	.56 ⁺ (.30)	-2.19 (1.79)
Party ID	.20 (.19)	16.00 (53.51)	.12 (.13)	.06 (.05)	.48 (.40)
Age	-.10 ^{***} (.03)	-27.27 ^{**} (8.37)	-.08 ^{***} (0.02)	.00001 (.007)	-.04 (.04)
Female	.30 (.39)	122.40 (111.34)	.28 (.28)	.15 (.10)	-.23 (.64)
Political Knowledge	.07 (.06)	25.09 (16.99)	.06 (.04)	.02 (.01)	.08 (.09)
First Number Fixation	.47 (.43)	153.30 (123.93)	.50 (.31)	.06 (.11)	-.06 (.72)

Note. Regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). Party ID (Strong Democrat = 6, Democrat = 5, Independents who identify as closer to Democratic Party's position = 4, Independents who identify as closer to Republican Party's positions = 3, Republican = 2, Strong Republican = 1). First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). ⁺ $p < .10$, ^{**} $p < .01$, ^{***} $p < .001$.

Supplementary Table 8

Attention and Memory Accuracy Analyses for Schema Consistency x Factual Accuracy Interaction

	Encoding Analyses				Memory Retrieval Analysis
	Number of Fixations During Regressions to First Number Region	Duration of Fixations During Regressions to First Number Region	Total Number of Regressions to First Number Region	Self-Report Measure of Interest in Article	Memory Accuracy
	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed Effects					
Schema Consistency x Factual Accuracy	-0.17 (.46)	-138.83 (155.37)	-.07 (.46)	.65 (1.40)	-.66 (.98)
Schema-Consistency (inconsistent = 1, consistent = 0)	.68 ⁺ (.30)	188.19 (93.63)	.37 (.27)	-.20 (.74)	-.87 (.54)
Factual Accuracy	.12 (.30)	97.36 (94.14)	-.009 (.27)	-.61 (.74)	.81 (.65)
Age	-.03 (.02)	-6.01 (5.17)	-.02 (.01)	.01 (.02)	-.02 (.02)
Female	.46 ⁺ (.26)	127.37 ⁺ (68.29)	.25 (.19)	.08 (.27)	.04 (.30)
Political Knowledge	.03 (.04)	5.60 (10.00)	.02 (.03)	.04 (.04)	.08 ⁺ (.04)
Party ID	.10 (.09)	10.69 (22.99)	.08 (.06)	.22* (.09)	.20* (.10)
First Number Fixation	-.02 (.24)	11.19 (64.65)	.05 (.17)	.17 (.28)	-.09 (.34)

Note. Mixed-effects regression coefficients are shown with standard errors in parentheses for Models 1, 2, and 3. Mixed-effects logistic regression coefficients are shown for Models 4 and 5. For the dependent variable in Model 5, accurate memory = 1 and inaccurate = 0 (pairwise gist). First Number Fixation (1 = first number was fixated first before the second number, 0 = second number was fixated first before the first number). ⁺ $p < .10$, * $p < .05$.

Supplementary Table 9

Demographic Characteristics of Participants in each Study

	% Female	Average Age (Range)	Average Self-reported Interest in Politics	% Republican	% Democrat	Source of Sample
Study 1	57%	24.82 (18-56)	1.83 (<i>SD</i> = 0.96)	22%	78%	U.S. State of Ohio (City of Columbus/University Community)
Study 2a	38%	33.08 (19-67)	1.84 (<i>SD</i> = 0.82)	35%	65%	Across Multiple States in U.S. (MTurk)
Study 2b	70%	43.96 (19-84)	2.02 (<i>SD</i> = .91)	30%	70%	U.S. State of Ohio (NIH <i>ResearchMatch</i>)

Note. Independents who identify as closer to Democratic Party's position are classified as "Democrat." Independents who identify as closer to Republican Party's position are classified as "Republican."

Supplementary Table 10

Race/Ethnicity Distribution of Participants in each Study

	White/Caucasian	Black/African American	Asian/Asian American/Pacific Islander	Latina/o/Hispanic	Multiracial	Other
Study 1	69	13	23	1	3	1
Study 2a	233	26	28	5	4	4
Study 2b	255	19	11	2	7	6

Supplementary Table 11

Issue Type by Wave Interaction for Study 2a including the Excluded Participants

	Study 2a (300 original participants + 101 excluded participants)
Issue Type x Wave	.65*** (.16)
Wave	-.37** (.11)
Issue Type	-2.33*** (.23)
Age	.004 (.006)
Female	-.02 (.11)
Interest in Politics	.02 (.06)
Party ID	-.05 (.03)
Over Numerical Limit	-.31+ (.23)

Note. Logistic regression coefficients with standard errors clustered on participants. Ns for waves 1, 2, and 3 are 119, 140, and 142, respectively. ⁺ $p < .10$, ^{**} $p < .01$, ^{***} $p < .001$. “Over numerical limit” variable indicates whether for a given issue for a particular participant, the responses sum to above 100% (same-sex marriage and gender preference for bosses) or 936 (for the police shooting issue).

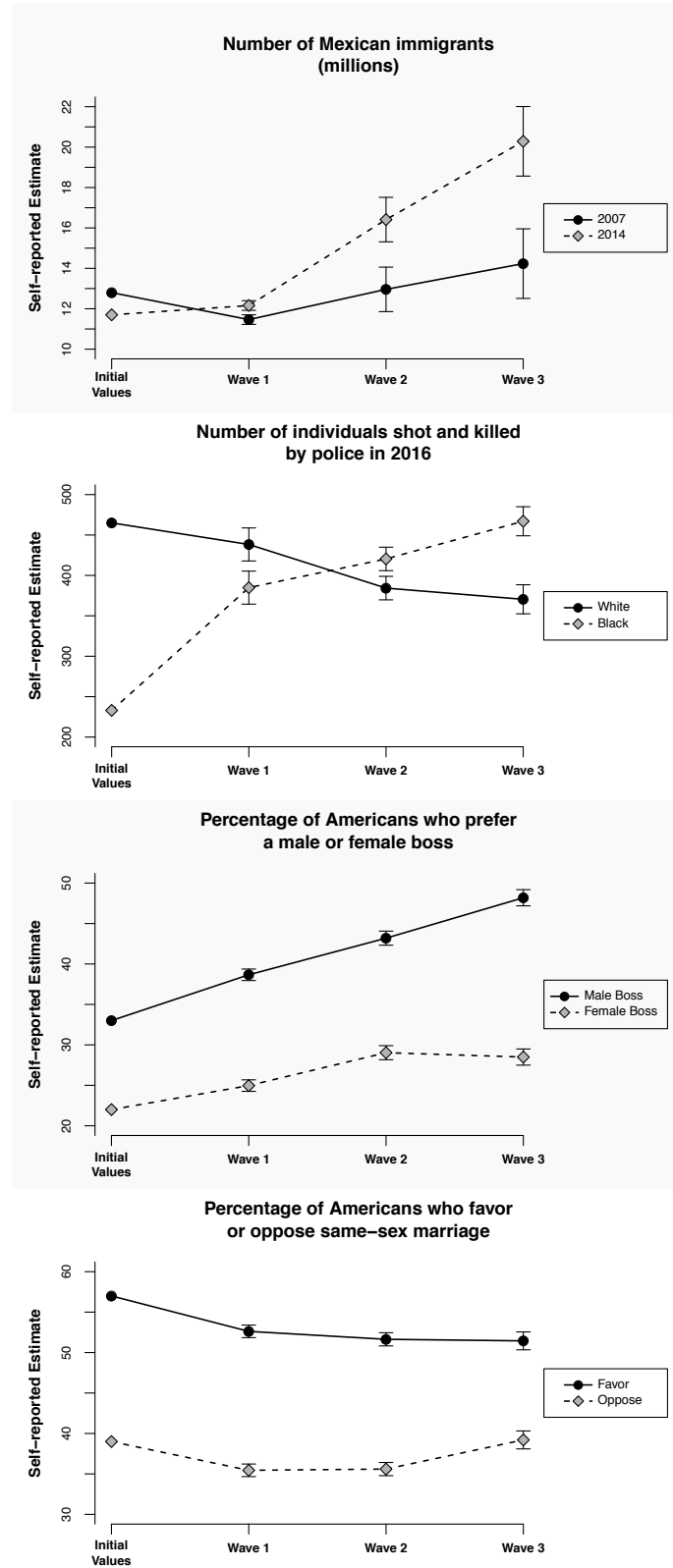


Figure S1. Average numerical estimates for Study 2a combining responses of the original 300 participants and the 101 excluded participants (whose numerical estimated exceeded 100% [for the same-sex marriage or gender preference for bosses issues] or 963 [for the police shooting issue]). Error bars are standard errors adjusted for the within-subjects design.