MIS Quarterly

RESEARCH NOTE

SAYS WHO? THE EFFECTS OF PRESENTATION FORMAT AND SOURCE RATING ON FAKE NEWS IN SOCIAL MEDIA[1]

Antino Kim and Alan R. Dennis
Kelley School of Business, Indiana University,
Bloomington, IN 47405 U.S.A. {antino@indiana.edu} {ardennis@indiana.edu}

News—real or fake—is now abundant on social media. News posts on social media focus users' attention on the headlines, but does it matter who wrote the article? We investigate whether changing the presentation format to highlight the source of the article affects its believability and how social media users choose to engage with it. We conducted two experiments and found that nudging users to think about who wrote the article influenced the extent to which they believed it. The presentation format of highlighting the source had a main effect; it made users more skeptical of all articles, regardless of the source's credibility. For unknown sources, low source ratings had a direct effect on believability. Believability, in turn, influenced the extent to which users would engage with the article (e.g., read, like, comment, and share). We also found confirmation bias to be rampant: users were more likely to believe articles that aligned with their beliefs, over and above the effects of other factors.

Keywords: Fake news, social media, Facebook, source highlighting, source rating

Introduction

Deception has been a long-running problem on the Internet, and it rose to global attention in 2016 with the U.S. presidential election, where deception in the form of "fake news" was deliberately created and spread through social media as part of a disinformation campaign to influence the election results (Allcott and Gentzkow 2017; Barthel et al. 2016; Shane 2017). This is not the first time social media have been used in an attempt to sway elections (Nahon 2015) or influence opinions (Deutsch 2018; English 2017; Grech and Masukume 2017). The prevalence of fake news has shaken the public's belief in journalism and stirred up criticism of social media, such as Facebook, for not taking more proactive measures (Barthel et al. 2016).

News has always been questionable in its reliability; even before the rise of the Internet, some newspapers were known for their biases and potentially distorted reports (Gaziano and McGrath 1986). Today, almost 62% of adults get news from social media (primarily Facebook), and the proportion is increasing (Gottfried and Shearer 2016). Social media have seldom been a neutral platform for news (Nahon 2015), but there are two important differences between social media and traditional journalism. First, on social media, anyone can create "news"—real or fake—and share it, and the posts spread quickly as others read and share them. Second, on most social media platforms, users do not actively choose the source of their stories; rather, the platform presents a mix of news from friends, from sources based on past activities, and from advertisers who have paid to place their content in the user's feed (some with malicious intent). The net effect is to move the quality control function from trained journalists with a putative interest in truth to regular users who have little control over the sources of stories they view and often give little thought to verifying the stories before spreading them. About 23% of social media users report that they have spread fake news (Barthel et al. 2016), and fake news spreads faster than true news on social media, primarily because of people, not bots (Vosoughi et al. 2018). Clearly, helping users make a more informed decision on social media is an important

[1]Gerald Kane was the accepting senior editor for this paper. Michael Chau served as the associate editor.

piece of the solution in curbing the spread of fake news, and this is the key focus of our study.

This paper investigates whether changes to the way in which the source of news is presented can "nudge" (Schneider et al. 2017; Thaler and Sunstein 2008) social media users to make more mindful decisions about whether to believe articles and perform activities that contribute to their spread (e.g., read, like, comment, and share). We investigate two approaches, one subtle and simple (changing the interface to highlight the source of the article) and another that is more expensive (a *source* rating—not to be confused with fact checking individual *articles*). We conducted two studies and found both approaches to have significant effects.

Theory and Hypotheses

Fake news has been defined as "news articles that are intentionally and verifiably false and could mislead readers" (Allcott and Gentzkow 2017). Fake news has long been a problem, but it became increasingly important during the 2016 election in the United States (Cerf 2016), where it may have influenced the outcome (Allcott and Gentzkow 2017; Barthel et al. 2016). And, the influence of fake news is not confined to only elections. A conspiracy theory that came to be known as "Pizzagate" (alleging that a pizzeria in the District of Columbia was the home of a child abuse ring led by the Democratic Party) (Gajanan 2016) went viral on social media (Fisher 2016). The restaurant owners and employees were harassed (Gajanan 2016; Kang 2016), and a man visited the restaurant with a rifle, terrifying customers and workers (Fisher 2016). Fake news can be created for profit or to spread propaganda or disinformation, either for political gain or to advocate beliefs to a wider audience (Deutsch 2018; English 2017; Grech and Masukume 2017; Shane 2017; Subramanian 2017; Sydell 2016).

We begin by discussing information processing in social media, and how the social media context often leads to confirmation bias. Next, we examine how two possible design changes to the social media platform might mitigate confirmation bias. We then consider how confirmation bias and users' beliefs in the articles influence the users' actions that would contribute to the spread of those articles.

Information Processing in Social Media

There are many reasons why people use the Internet, such as accomplishing tasks or simply hedonic enjoyment (Zhou et al. 2011). Most individuals use social media for hedonic purposes (Harsanyi 1977) such as seeking entertainment or connecting with friends (Johnson and Kaye 2015), rather than utilitarian ones such as work tasks. Individuals in a hedonic mindset are less likely to critically consider information than those in a utilitarian mindset, as their consumption is tied to what they desire reality to be, rather than what they know to be real (Hirschman and Holbrook 1982). They are less mindful (Thatcher et al. 2018).

Social media platforms learn users' preferences by tracking what they read and the actions that they take (e.g., like, comment, and share). As commercial entities, the platforms aim to maximize user satisfaction and, thus, they display more content matching the users' choices, so that users see more posts that match their existing beliefs (WSJ 2016). This causes a decrease in the range of information that the users encounter, and, as a result, social media users often exist in small information bubbles—often referred to as echo chambers (Cerf 2016)—that reinforce their beliefs and make them feel that the world is more like them (WSJ 2016).

When individuals encounter information that challenges their preexisting beliefs, they experience cognitive dissonance (Festinger 1957; Mills 1999). When an individual sees two contradictory facts, both of which are plausible (e.g., John is honest, but a story says he lied), he/she must resolve the inconsistency, either by concluding that the two facts are not contradictory (e.g., John lied, but he is still honest because lying is not related to honesty), or by accepting one and rejecting the other (e.g., John is honest, and thus I do not believe he lied; or John lied, and thus I do not believe he is honest) (Festinger 1957).

Resolving such a cognitive dissonance takes cognitive effort, and humans tend to be cognitive misers who resist expending effort (Simon 1979). This tendency is exacerbated when humans are in a hedonic mindset (Hirschman and Holbrook 1982). Because rejecting the new information is simpler than reassessing one's preexisting beliefs, most people retain their existing beliefs and discard the new information as being false (Devine et al. 1990; Koriat et al. 1980; McKenzie 2006). This tendency to favor information that confirms one's preexisting beliefs and ignore information that challenges them is called *confirmation bias* (Devine et al. 1990; Koriat et al. 1980; Nickerson 1998). People are more likely to believe information that matches their preexisting views (i.e., attitude homophily or alignment; Allcott and Gentzkow 2017; Housholder and LaMarre 2014):

H1: *Users are more likely to believe an article that aligns with their preexisting beliefs on the topic.*

Says Who?

In the physical world, people care about who says something—that is, the source of a story (McCracken 1989). When someone says something, we are naturally attuned to the veracity of the speaker. We accept news from those we see as experts (Hovland et al. 1953; McCracken 1989) and discount news from those with a reputation for falsehoods. For example, suppose one of your colleagues told you that the accounting server had been hacked and your firm had lost $2 million. Would you believe it? Part of your belief would be driven by preexisting beliefs about computer security at your firm (i.e., confirmation bias), and part would be influenced by the person who told you the story (e.g., whether he or she has expert knowledge and a reputation for truth-telling).

In the physical world, *who* says something always comes before the content. When we talk with people, we consider who they are before we think about what they say. In this *source-primacy* world, we use our *a priori* knowledge of the source when we consider a story, and our knowledge of the source shapes how we evaluate the content that follows (Hovland et al. 1953). We are more likely to believe information from sources with a reputation for honesty and are less likely to accept information from sources with a poor reputation or sources we do not know (McCracken 1989). Thus, our knowledge of the source influences how we process the information that follows.

On social media, things are different. We focus on Facebook because it is the largest social media platform in the world with over 2 billion users (Statista 2018). On Facebook, users do not choose the sources of news articles; stories from many different sources are intermixed in the same feed. Posts that our Facebook friends share appear. Articles from sources we have read in the past as well as from sources that we have not selected, and may not be cognizant of, appear. Perhaps more insidiously, posts from advertisers paying to get attention (to drive sales or influence opinion) appear intermixed with these more innocuous posts. The posts may also be true or be fake news, designed to deliberately influence opinions and actions—whether they are from foreign agencies or friends who have accidentally or intentionally shared them. This is unlike other news outlets (e.g., TV news, newspapers, online news sites, news apps on mobile phones) where knowledge of the source comes before we view the content.

The current Facebook interface presents articles with little emphasis on who relates the story. The design highlights the article headline and deemphasizes the source (see Figure 1). In the *headline-primacy* world of Facebook, we read the headline first—along with an eye-catching image—and only consider the source as an afterthought, if at all. This interface design influences us to process the headline without consideration of the source. This presentation format disrupts the normal consideration of source that occurs in a source-primacy presentation and makes users more likely to accept the content without the normal and automatic consideration of the source.

When the source precedes the message, the presentation of the source triggers us to think about the source before we read the message (Tormala et al. 2006, 2007). Our thoughts of the source shape how we think about the message that follows. We are biased to think positively about messages from reputable sources and negatively about messages from disreputable sources, so we are more likely to believe messages from reputable sources than disreputable ones (Tormala et al. 2006, 2007). Conversely, when the message precedes the source, our beliefs about the message (i.e., confirmation bias) influence what we think about the source, so the source has less impact on whether we believe the message (Tormala et al. 2006, 2007). Tormala et al. (2006, 2007) studied utilitarian situations in which individuals were motivated to think carefully about messages, rather than the hedonic environment of social media where users are not motivated to think (Gabielkov et al. 2016). Yet, the same processes may be at play in hedonic environments.

We argue that if we can nudge (Schneider et al. 2017; Thaler and Sunstein 2008) social media users to focus first on *says who* (i.e., the source) and then the headline, the source will become more salient. A more salient source will induce users to consider the source before they consider the story. We theorize that this increased source salience, and the increased likelihood of thinking about the source prior to the content, will induce the same mindset that we use when we hear stories in the physical world. Thus, users will adopt a more skeptical mindset.

Changing the headline-primacy interface of Facebook to a source-primacy one is simple: place the source before the headline and highlight it in a way that will induce users to read it before they read the headline. We argue that this interface design will nudge users into a more skeptical mindset by considering the source first and the content second. In short, there will be a main effect:

H2: *Users are less likely to believe an article that is presented in a source-primacy format than one presented in a headline-primacy format.*

A source's reputation influences the extent to which we believe what they say (McCracken 1989). Reputation is based

Figure 1. An Example of a Facebook News Post

on three distinct but interrelated dimensions (Eisenegger and Imhof 2008; Thiessen and Ingenhoff 2011). The first is the functional dimension, which is the ability to achieve certain aims and goals—that is, competence (Eisenegger and Imhof 2008). The second is the social dimension, which is the legitimacy of actions as assessed by social norms (Eisenegger and Imhof 2008). The third is the affective dimension, which is subjective and based on alignment with the appraiser's own beliefs (Eisenegger and Imhof 2008).

As argued above, the source-highlighted format will make the source more salient. Thus, the source's reputation will become more important in assessing the article's believability. Just as in the physical world, news from sources with a good reputation will be more believable than news from unknown sources. Thus, there will be an interaction between the presentation format and the source reputation, such that the source-highlighted format will increase the effects of source reputation.

H3: *The source-primacy format will increase the effect of source reputation on an article's believability.*

Source Ratings

In response to fake news, a number of fact-checking initiatives have been launched which strive to rate the truthfulness of each individual news report (Graves 2016; Lowrey 2017; Wintersieck 2017). An alternative—or complement—to fact checking the *articles* is to rate the original *sources* (e.g., authors or sites). Source ratings based on fact-checking of past articles each source has produced can be used as a predictor of the credibility of future articles, thus producing a source reputation measure that can be attached to new articles at the time of their production. Here, the idea is similar to how eBay uses past purchase ratings to produce seller ratings attached to new products offered for sale.

Source reputation for news articles is the extent to which the source is seen as producing valid statements, and it affects the extent to which we believe a specific report to be credible, although there are article-specific factors as well (Morris et al. 2016; Yang 2012). Source reputation in social media is influenced by past performance and is gradually built by a history of behavior that displays experience, expert knowledge, and reliable information (Teng et al. 2017).

For our research, the mechanism behind the ratings is less important than whether or not source reputation ratings influence users. Here, we assume that the ratings are created by aggregating the ratings from an expert panel assessing prior articles produced by the source. Past research shows that fact checking individual articles influences their believability (Wintersieck 2017), so we argue that the aggregation of these individual items will also have a similar effect:

H4: *Source reputation ratings will directly affect the believability of articles, such that lower source reputation ratings will lead to lower believability of articles from that source.*

Effects on Behavior

Thus far, we have focused on how social media users assess the believability of news articles they see. This believability, in turn, can affect the actions of the users. They can choose to read the details or not, and they also can choose to provide feedback on the post (e.g., like or comment) as well as contribute to the spread of the article (e.g., share). Each of these actions is separate and distinct; you can like, comment on, or share an article without reading it (Gabielkov et al. 2016), although there may be some coherence in behavior: reading before liking, sharing only if one read and liked an article (sometimes without clicking the *Like* button), and so on (Kim and Yang 2017).

We argue that behavior is influenced by preexisting beliefs and the believability of the article. It is possible that presentation format and source ratings may have additional effects over and above confirmation bias and believability, so we include them in our analyses, although we do not hypothesize any effects. A user is more likely to read an article if it is congruent with his or her prior beliefs due to confirmation bias (Ask and Granhag 2005). Confirmation bias often causes selective information search (Ask and Granhag 2005; Klayman 1995), in which people actively seek information that confirms their beliefs and avoid information that does not. Selective information searching will be intensified when people are in a hedonic mindset because they are not seeking to find a correct utilitarian outcome (e.g., determining if a view is correct) but rather are seeking entertainment and enjoyment. Viewing information that supports your beliefs is more enjoyable than viewing information that challenges them (Festinger 1957; Minas et al. 2014), so people will be more likely to read articles that support their preexisting beliefs.

Other actions on Facebook—such as like, comment, and share—are unequally distributed (Hampton et al. 2012; Lee et al. 2016). Most users seldom engage in these behaviors, perhaps because they require more cognitive effort than simply reading (Kim and Yang 2017; Muntinga et al. 2011). Nonetheless, those users engaging in these behaviors do so relatively often (Hampton et al. 2012; Lee et al. 2016). It may be that most users engage in casual reading behavior (which is passive and not observable to other users), while those who are more active on Facebook choose to engage in behaviors that are observable to other users. It may also be that people with certain personality traits are more or less likely to engage in these behaviors (Lee et al. 2016). The choice to act on an article can be influenced by emotion or information, with liking being driven more by emotion, commenting more by cognition, and sharing by both (Kim and Yang 2017).

Therefore, we theorize that one important factor influencing the decision to read, like, comment on or share news articles on social media is the fit with preexisting beliefs. The stronger the fit, the more likely the article is to trigger an emotional reaction leading to a *Like*, or to trigger a cognitive reaction leading to a comment, or to both, leading to sharing. Thus,

H5: *Users are more likely to choose to (a) read, (b) like, (c) comment on, and (d) share an article that aligns with their preexisting beliefs about the topic.*

We theorize that a second important factor influencing the decision to act on an article is the extent to which a user believes it to be true. Believability can be an important factor in the use of social media information (Johnson and Kaye 2015) because if someone does not believe information to be true, they are less likely to engage in it or to encourage its spread by sharing it. Thus,

H6: *Users are more likely to choose to (a) read, (b) like, (c) comment on, and (d) share an article when they believe it to be true.*

Study 1

Method

Participants. We recruited 445 participants—125 through Facebook posts by the authors and the authors' business school, and 320 from a Qualtrics panel of adults in the United States. Half were female, and about 5% were below 24 years of age (college age), 83% between 25 and 64 (working age), and 13% above 65 (retirement age). About 45% did not have a bachelor's degree, 27% had a bachelor's degree, and 28% had a graduate degree. About 36% self-identified as Republican, 49% as Democrat, and 15% as independent.

Task. The participants answered a 15-minute survey that presented 12 news headlines, with 6 designed to appeal to politically left-leaning participants and the other 6 to right-leaning participants (see the Appendix). The headlines were formatted as Facebook posts. The headlines were designed to avoid major differences in the type and magnitude of feelings they would generate (i.e., one shocking headline and the other

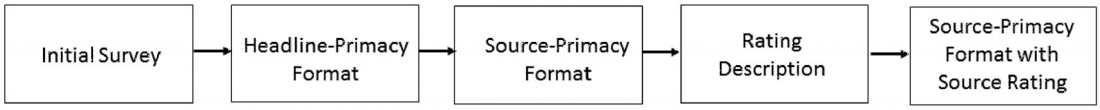

Figure 2. Sequence of the Experiment

bland, one with a celebrity image and the other with none, etc.). We used a gender-neutral name for the *poster*—not to be confused with the original author—and the comment from the poster was a summary of the headline itself. All these efforts were to minimize headline-specific effects, presentation-order effects, and poster-specific effects.

Independent Variables. There were three treatments, and all participants received all treatments, with four headlines presented in each treatment. The headlines seen by each participant were randomly assigned to treatment and presented in random order within treatment, although to prevent confusion, the treatments were always in the order of first to last (see Figure 2).

The first treatment was *headline-primacy* format (the control treatment) designed to mimic the current Facebook style of presentation as closely as possible (see Figure 3(a)). The second treatment was *source-primacy* format with the name of the source presented first (see Figure 3(b)). The third treatment was *source-primacy with source ratings* (see Figure 3(c)); two of the headlines were presented with high source ratings and the other two with low source ratings. A message was inserted before participants saw the headlines with source ratings to explain that the ratings were generated by a group of experts formed by Facebook to ensure that our subjects presumed the ratings to be legitimate.

In a repeated measures design, there are concerns about the effects of an early treatment bleeding over into later ones. This is usually controlled by random treatment order or a fully crossed design in which all treatment orders are used equally. In our study, however, there are meaningful theoretical differences in the likelihood of bleed-over between treatments, so these designs are not appropriate. The control treatment is the current Facebook format, so it is unlikely to influence later treatments. However, the source-primacy treatment is likely to influence the treatments that follow it, because once users are sensitized to think about the source, they are likely to continue to think about it in later treatments. Likewise, once the source rating treatment cues users to think about source reputation, they are likely to think about it in the treatments that follow. Therefore, we placed the control treatment first, the source-primacy treatment second, and the source rating treatment third. We revisit this issue in Study 2.

Two independent variables were self-reported by the participants. The first was the strength of confirmation bias the participant had for the article, which was measured by multiplying the article's importance to the participant (*Do you find the issue described in the article important*? 1 = not at all, 7 = extremely) by the participant's position on the article (–3 = extremely negative to +3 = extremely positive). Thus, it ranged from –21 to +21.

The second was whether the participant knew the source and considered it reputable or not. For unknown sources, we fabricated three names that sounded plausible (*ClickMedium.com*, *NewsUnion.com*, and *MediaNow.com*). All these URLs were verified to be inactive before the experiment. For the known and reputable source, we picked ABC News because it has been ranked among the most reputable and well-known news outlets in the United States across the political spectrum (Engel 2016, 2017). Nonetheless, we also asked each participant whether ABC News was a trusted news source using a seven-point scale, ranging from "strongly disagree" to "strongly agree." We used the word "trusted" because in the popular vernacular, a trusted news source is one with a strong reputation. We coded ABC News as a reputable source for the 325 participants (73%) who rated it 5 or higher. As a robustness check, we ran two additional analyses: (1) using a threshold of 6 or higher, and (2) without subjects who rated 4 or lower. Our results were the same.

Dependent Variables. The believability of each article was measured by taking the average of three 7-point items (*How believable do you find this article? How truthful do you find this article? How credible do you find this article*?). Cronbach's alpha was adequate (0.95). We also measured what actions the participant would take, specifically how likely the participant would be to: *Read, Like, Post a supporting comment, Post an opposing comment,* and *Share*.

Results

The mean believability levels for the three treatments are shown in Table 1. The mean believability for articles in the headline-primacy (i.e., no source highlighting and no rating) is higher than that for those in the source-primacy format (without rating). For the source rating, we exclude scores from

(a) Headline-Primacy

(b) Source-Primacy

(c) Source-Primacy with Source Rating

Figure 3. Study 1 Sample Story

Table 1. Means and Standard Deviations for Believability

Sample Description	Mean	Std. Dev.	N
Effect of Format (All sources)			
Headline-primacy format & no rating	4.806	1.665	1655
Source-primacy format & no rating	4.692	1.739	1416
Effect of Source Rating (Only unknown sources)			
Source-primacy format & no rating	4.482	1.728	905
Source-primacy format & high rating	4.980	1.636	623
Source-primacy format & low rating	4.213	1.918	615

a known and reputable source since it would be more believable regardless of rating.

To test our hypotheses, we performed multilevel mixed-effects linear regression with random intercepts in Stata. About 65% of the participants omitted one or more demographics questions, so including all the demographics items in our analyses would significantly reduce our sample size. We did an initial analysis including all items and found that whether or not participants were from the Qualtrics panel had significant impacts across all dependent variables[2] but no other demographics item was significant in more than one of the seven analyses. Therefore, we retained the Qualtrics variable but omitted other demographics items to retain the maximum sample size. There were no differences in the statistical conclusions with or without the omitted items.

Table 2 presents the estimated coefficients for believability. Confirmation bias has a positive and significant effect on believability. In other words, H1 is supported. Source-primacy has a significant effect, supporting H2. H3 argued that the source-primacy format would strengthen the relationship between the users' perceptions of source reputation and the believability of articles. The interaction of a reputable source and the source-primacy format is not significant, so H3 is not supported. H4 argued that source ratings would affect believability. High ratings have a positive and significant effect whereas low ratings have a negative and significant effect; H4 is supported.

We conducted a *post hoc* analysis to see if the effects of confirmation bias depended on the nature of the user or the headline in any way. We found that the effects of confirmation bias were significantly lower for those of retirement age, but otherwise there were no differences for any demographic groups (age, education, or political affiliation). We also found that left-leaning headlines were less likely to be believed overall, but confirmation bias had a greater effect for them. In other words, regardless of being Democrat, Republican, or independent, our participants were less likely to believe left-leaning headlines, but our left-leaning participants were more likely to believe them.

Table 3 shows the results for users' actions. Confirmation bias has a significant impact across all actions; users are more likely to read, like, and share articles that match their point of view. Commenting is consistent with other types of activities; users are more likely to leave supporting comments for articles that match their opinion and leave opposing comments for articles that they disagree with. In other words, confirmation bias influences users' behaviors; H5a–d are supported.

The same is true for believability; it has a significant positive effect on all actions, indicating that users are more likely to act on articles they believe to be true (see Table 3). H6a–d are supported. The pattern of results suggests that believability mediates most of the effects of source-primacy format and source rating; both have few consistent effects on users' actions over and above their effects on believability, although the source-primacy format does have some effects.

We conducted another *post hoc* analysis to better understand the effects of believability. We coded a believability score of 6 or above as "believe" and 2 or less as "not believe" and similarly, for actions (e.g., read, like, comment, and share), we coded 6 or above as "likely to take action" and 2 or less as "not likely to take action." The results in Table 4 indicate that about 91% of the articles that were perceived to be believable were likely to be read while only 20% of the articles that were unbelievable were likely to be read. The patterns are similar for other action types as well.

Study 2

Method

In Study 1, the treatment order was fixed, which left several open questions. First, it is possible that participants simply became more critical as they read more articles, and our treatments (i.e., source-primacy format and ratings) did not actually have any significant impact. Second, treatment order effect is also a potential concern; perhaps, the source-primacy format is effective only if it appears after the control, and the ratings are effective only if they follow the source-primacy format. Finally, it is also possible that ratings are effective only in the source-primacy format. We investigate all of these issues in our second experiment.

The second experiment was similar to Study 1 with four notable differences. First, we used a 2 × 2 between-subject design with four treatment groups: (1) control treatment, (2) source-primacy treatment, (3) source rating treatment, and (4) source-primacy with source rating treatment. The 12 headlines were randomly assigned to sources and ratings to control for any source-specific or headline-specific effects. For the treatments with source ratings included, the rating description message was inserted at the start of the experiment

[2]This is merely a methodological artifact and makes no theoretical contribution. As a robustness check, we ran additional analyses to verify whether the Qualtrics panel members reacted differently to the treatments than did other participants. We found no such evidence for Source-Primacy treatment and Rated Low treatment, although the Qualtrics panel members did appear more skeptical about the Rated High treatment.

Table 2. Study 1 Results for Believability

Independent Variables	Coefficients
Source-Primacy Format	–0.139*
Rated High	0.389***
Rated Low	–0.329***
Reputable Source	0.396***
Reputable Source × Source-Primacy	0.021
Confirmation Bias	0.477***
Qualtrics Panel Member	0.537***

Notes: Confirmation Bias is standardized. Number of subjects = 445. Number of observations = 4309. ***p < 0.001, **p < 0.01, *p < 0.05.

Table 3. Study 1 Results for User Actions

Independent Variables	Read	Like	Support	Oppose	Share
Believability	0.531***	0.310***	0.219***	0.090*	0.245***
Source-Primacy Format	–0.110	–0.089*	–0.050	–0.156**	–0.116**
Rated High	0.018	–0.090	–0.075	–0.045	–0.027
Rated Low	–0.137*	–0.018	–0.051	–0.027	–0.028
Reputable Source	0.021	0.008	0.020	–0.055	0.010
Reputable Source × Source-Primacy	0.089	0.058	0.018	0.157*	0.066
Confirmation Bias	0.221***	0.712***	0.515***	–0.314***	0.275***
Qualtrics Panel Member	0.672***	1.338***	1.532***	1.683***	1.581***

Notes: Believability and Confirmation Bias are standardized. Number of subjects = 445. Number of observations = 4309. ***p < 0.001, **p < 0.01, *p < 0.05.

Table 4. Study 1 Effects of Believability on User Actions

Actions	Believability ≤ 2	Believability ≥ 6
Read	19.8%	90.6%
Like	1.4%	75.9%
Supporting Comment	0.8%	72.0%
Opposing Comment	7.1%	66.3%
Share	2.5%	73.6%

to explain that the ratings were generated by a group of experts formed by Facebook to ensure subjects presumed the ratings to be legitimate.

The other three changes were minor. We added a control variable to see if there were any differences in believability due to order (i.e., whether headlines in the first half were more or less likely to be believed than those in the second half). We used only unknown sources (i.e., not trusted sources), so we did not test H3; we fabricated 12 names (URLs) that sounded plausible and verified that all were inactive. Finally, we changed the source ratings to use stars, which are more common on e-commerce sites (see Figure 4). For the treatment conditions with source ratings, six headlines were rated high (4.96 to 4.99) with the other six rated low (1.05 to 1.08).

We recruited 501 active Facebook users (those who used Facebook more than once a day) from a Qualtrics panel and randomly assigned them to one of the four treatments. Half were female, and about 9% were below 24 years of age (college age), 77% between 25 and 64 (working age), and 14% above 65 (retirement age). About 70% did not have a bachelor's degree, 21% had a bachelor's degree, and 9% had a graduate degree. About 37% self-identified as Republican, 46% as Democrat, and 17% as independent.

(a) Source Rating

(b) Source-Primacy with Source Rating

Figure 4. Study 2 Sample Story

Results

Table 5 presents the results from Study 2. The control variable shows that there were no differences due to the order of presentation, thus alleviating that concern from Study 1. Once again, H1 (confirmation bias) and H2 (source-primacy format) are supported. H4 (source ratings), on the other hand, is now partially supported, with only low ratings having an effect. Unlike Study 1, subjects who saw articles with the rating treatment did not see articles in the control format. We speculate that without the control format to compare to, the high ratings became the baseline because high ratings are the most common on the Internet (McGlohon et al. 2010; ReviewMeta 2016). Since high ratings were expected, they had little effect on believability. However, low ratings still had significant effects on believability. How users react to low-rated sources is more important because the low-rated sources are the usual culprits in spreading fake news, which is our focus in this work. The coefficients indicate that a low rating had about twice the effect of the source-primacy format, consistent with what we found in Study 1.

One somewhat unexpected result is a significant positive interaction between the source-primacy format and low ratings. The main effects of source-primacy format and low ratings are significant and negative, so a positive interaction indicates a substitution effect. That is, the source-primacy format and low ratings both separately make users less likely to believe the articles they see (with strengths of –0.245 and –0.589, respectively), but when the two are combined, the effect is significantly less than the sum of the two (by about 0.383). This means that the effect of ratings dominates the effect of the source-primacy format, such that once ratings are present, changing to this format may not add much value.

Table 6 shows the analyses for users' actions. Once again, H5a–d (confirmation bias) and H6a–d (believability) are all supported. Format and ratings have no effect over and above these two.

Discussion and Conclusion

The results of our two studies show that presenting articles on social media in a source-primacy format (with the source of the article before the headline) as compared to Facebook's current headline-primacy format (with the headline before the source) makes users less likely to believe them. We expected that users would believe articles from reputable sources more when they were in source-primacy format, but we did not find any support for this hypothesis. The source-primacy format nudges readers to be more skeptical of *all* articles, regardless of their source.

Our results also show that source reputation ratings influenced the believability of articles, with low ratings having more than twice the effect of formatting. When the sources were unknown, a low rating reduced readers' belief. The effects of high ratings were less clear; they were effective only when unrated articles were also presented to the readers. The coefficients in Tables 2 and 5 are standardized, so we can assess the relative effects of presentation format versus source ratings. Low source ratings have more than twice the effect

Table 5. Study 2 Results for Believability

Independent Variables	Coefficients
Source-Primacy Format	–0.245*
Rated High	0.132
Rated Low	–0.589***
Source-Primacy × Rated High	0.140
Source-Primacy × Rated Low	0.383*
Confirmation Bias	0.567***
First Half Articles	0.065

Notes: Confirmation Bias is standardized. Number of subjects = 501. Number of observations = 6012. ***p < 0.001, **p < 0.01, *p < 0.05.

Table 6. Study 2 Results for User Actions

Independent Variables	Read	Like	Support	Oppose	Share
Believability	0.595***	0.397***	0.365***	0.272***	0.462***
Source-Primacy Format	–0.204	–0.146	–0.177	–0.086	–0.100
Rated High	–0.088	–0.055	–0.198	0.040	–0.118
Rated Low	–0.174	0.016	–0.117	–0.052	–0.072
Source-Primacy × Rated High	0.260	0.340	0.334	0.095	0.221
Source-Primacy × Rated Low	0.199	0.186	0.249	0.265	0.137
Confirmation Bias	0.230***	1.050***	0.872***	–0.531***	0.416***
First Half Articles	0.076	0.065	0.024	0.068	0.031

Notes: Believability and Confirmation Bias are standardized. Number of subjects = 501. Number of observations = 6012. ***p < 0.001, **p < 0.01, *p < 0.05.

of presentation format. Low ratings and the source-primacy format are substitutes, so applying the formatting to ratings has little added benefit.

Confirmation bias also affects users' actions. Users are more likely to read, like, post supporting comments, and share articles that they agree with (and more likely to leave opposing comments for articles they disagree with). Our results suggest that altering an article's believability could potentially counterbalance the effect of confirmation bias.

Implications for Research

First and foremost, our results show that the presentation format affects the believability of articles posted on social media, which, in turn, affects users' actions. Placing the source of an article before the headline increased users' skepticism. The prevalence of news consumption on social media (Gottfried and Shearer 2016), combined with the hedonic, entertainment-seeking goals of social media use (Johnson and Kaye 2015), suggests that news consumption may not be mindful, so increased skepticism is a good step forward. Mindfulness can have significant impacts on the way we use technology, for better or worse (Thatcher et al. 2018). Our results show that the way technology is designed to present articles has a significant impact on their believability and the subsequent behaviors that the believability influences. Thus, subtle technology designs can nudge social media users to be more or less skeptical of the articles they read. We need more research on how such simple changes to presentation formats can influence perceptions and behaviors.

We also found that source ratings purportedly from a panel of experts hired by Facebook had significant effects on the believability of articles. There are, of course, many other ways to develop source ratings. If the ratings came from regular users assessing past articles (as users do on Amazon and eBay for past purchases), would they have a stronger or weaker effect on believability than the ratings from a panel of experts? We need more research on ways to compile and present ratings of articles on social media. Another question related to source rating is what would happen if the ratings are not aligned with the users' own judgement. In this work,

ratings were applied to only unknown sources, so the interaction between users' *a priori* assessment of the sources and the ratings remains an open question.

Moreover, our results show that believability and confirmation bias influence the actions a user would take. For reading, believability is the most important factor, being more than twice as important as confirmation bias (see Tables 3 and 6). For sharing, believability has an effect comparable to confirmation bias. However, the pattern is very different for other types of behaviors such as liking or writing a supporting or opposing comment; for these, confirmation bias is the dominant factor, having two or three times the effect of believability. For these actions, which also contribute to the spread of fake news, users care more about whether the article supports (or opposes) their views than whether the article is true. This is curious because liking and commenting have a similar effect as sharing; users on Facebook can view posts that their friends like and comment on, and such posts appear on friends' feeds in a similar manner to shared posts. Hence, it is not clear whether the users are aware of this indirect effect of sharing through liking and commenting. Do they know they could be spreading fake news? If this effect was made clear to users, would it induce them to behave differently? Also, if who will see their activities was explicit, would that make a difference?

Finally, another interesting aspect is the herding effect (Sun 2013). We focused on the original source of the article, not the individual who shared the post on his or her social media feed; we held this constant. Would it matter who shared the article and their relationship to the user, who liked it, and how many people commented on it? Is there any "poster" effect, and if so, how would it interact with the "news source" effect? More research is needed to better understand how who shares an article influences its believability and the actions users take.

Implications for Practice

Our results show that highlighting the source has the effect of making users more critical about the believability of articles. Perhaps, the most important aspect of this finding is that we can nudge people to be more skeptical by making a simple and easy modification to the interface. The effects may not be as strong as other options, but this change is a low-hanging fruit.

In addition, the source reputation ratings (low ratings in particular) can play an important role in how people evaluate the believability of the articles they see, which, in turn, affects their engagement with them (i.e., read, share, etc.). Implementation of a rating system will take time, and further research may be needed to identify the best form of ratings. In practice, many of the fake news articles are from sources that are unknown or even deceptive. The culprits interested in luring traffic to their sites for advertisement revenue may post fake news as "click bait," and to make their scheme more effective, they may intentionally choose names that are very similar to legitimate ones such as ABCnews.com.co (Lu 2014; Murtha 2016). We believe source reputation rating can be an effective countermeasure against such deceptions.

Conclusion

In summary, we investigated two approaches to nudging users to be more skeptical of fake news on social media. Our findings indicate that, although the effect of confirmation bias is salient, both (1) changing the interface to place the source of the article before the headline and (2) presenting low reputation ratings for the source make users less likely to believe an article. Therefore, both approaches can be effective in curbing the spread of fake news on social media.

Acknowledgments

We thank the review team of *MIS Quarterly* and the participants of the 51st Hawaii International Conference on Systems Science for their valuable feedback.

References

Allcott, H., and Gentzkow, M. 2017. "Social Media and Fake News in the 2016 Election," *Journal of Economic Perspectives* (31:2), pp. 211-236.

Ask, K., and Granhag, P. A. 2005. "Motivational Sources of Confirmation Bias in Criminal Investigations: The Need for Cognitive Closure," *Journal of Investigative Psychology and Offender Profiling* (2:1), pp. 43-63.

Barthel, M., Mitchell, A., and Holcomb, J. 2016. "Many Americans Believe Fake News Is Sowing Confusion," Pew Research Center, Washington, DC (http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/).

Cerf, V. G. 2016. "Information and Misinformation on the Internet," *Communications of the ACM* (60:1), p. 9.

Deutsch, J. I. 2018. "Folklore, Politics, and Fake News in the Reception of Rogue One," *New Directions in Folklore* (15:1/2), pp. 109-120.

Devine, P. G., Hirt, E. R., and Gehrke, E. M. 1990. "Diagnostic and Confirmation Strategies in Trait Hypothesis Testing," *Journal of Personality and Social Psychology* (58:6), pp. 952-963.

Eisenegger, M., and Imhof, K. 2008. "The True, the Good and the Beautiful: Reputation Management in the Media Society," in

Public Relations Research: European and International Perspectives and Innovations, A. Zerfass, B. van Ruler, and K. Sriramesh (eds.), New York: Springer, pp. 125-146.

Engel, P. 2016. "The Most and Least Trusted News Sources in America," *Business Insider* (http://www.businessinsider.com/most-least-trusted-news-outlets-sources-channel-2016-12).

Engel, P. 2017. "These Are the Most and Least Trusted News Outlets in America," *Business Insider* (http://www.businessinsider.com/most-and-least-trusted-news-outlets-in-america-2017-3).

English, R. M. 2017. "Anti-Vax Truthiness: How Social Media Echo Chambers Reinforce Anti-Vax Beliefs in Parents," in *Social Media: Practices, Uses and Global Impact*, T. C. Ahern (ed.), New York: Nova Science Publishers Inc., pp. 173-193.

Festinger, L. 1957. *A Theory of Cognitive Dissonance*, Stanford, CA: Stanford University Press.

Fisher, M., Cox, J. W. , and Hermann, P. 2016. "Pizzagate: From Rumor, to Hashtag, to Gunfire in D.C.," *The Washington Post*, December 6.

Gabielkov, M., Ramachandran, A., Chaintreau, A., and Legout, A. 2016. "Social Clicks: What and Who Gets Read on Twitter?," *ACM SIGMETRICS Performance Evaluation Review* (44:1), pp. 179-192.

Gajanan, M. 2016. "Now Pizzagate Conspiracy Theorists Are Targeting a Pizzeria in New York City," *Time*, December 7 (http://time.com/4594323/pizzagate-robertas-new-york-conspiracy-theory/).

Gaziano, C., and McGrath, K. 1986. "Measuring the Concept of Credibility," *Journalism Quarterly* (63:3), pp. 451-462.

Gottfried, J., and Shearer, E. 2016. "News Use across Social Medial Platforms 2016," Pew Research Center, Washington, DC (http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/).

Graves, L. 2016. "Boundaries Not Drawn: Mapping the Institutional Roots of the Global Fact-Checking Movement," *Journalism Studies* (19:5), pp. 613-631.

Grech, V., and Masukume, G. 2017. "Fake News of Baby Booms 9 Months after Major Sporting Events Distorts the Public's Understanding of Early Human Development Science," *Early Human Development* (115), pp. 16-17.

Hampton, K. N., Goulet, L. S., Marlow, C., and Rainie, L. 2012. "Why Most Facebook Users Get More Than They Give," *Pew Internet & American Life Project* (3), pp. 1-40.

Harsanyi, J. C. 1977. "Morality and the Theory of Rational Behavior," *Social Research* (44:4), pp. 624-656.

Hirschman, E. C., and Holbrook, M. B. 1982. "Hedonic Consumption: Emerging Concepts, Methods and Propositions," *Journal of Marketing* (46:3), pp. 92-101.

Housholder, E. E., and LaMarre, H. L. 2014. "Facebook Politics: Toward a Process Model for Achieving Political Source Credibility through Social Media," *Journal of Information Technology & Politics* (11:4), pp. 368-382.

Hovland, C. I., Janis, I. L., and Kelley, H. H. 1953. *Communication and Persuasion: Psychological Studies of Opinion Change*, New Haven, CT: Yale University Press.

Johnson, T. J., and Kaye, B. K. 2015. "Reasons to Believe: Influence of Credibility on Motivations for Using Social Networks," *Computers in Human Behavior* (50), pp. 544-555.

Kang, C. 2016. "Fake News Onslaught Targets Pizzeria as Nest of Child-Trafficking," *The New York Times*, Technology Section, November 21.

Kim, C., and Yang, S.-U. 2017. "Like, Comment, and Share on Facebook: How Each Behavior Differs from the Other," *Public Relations Review* (43:2), pp. 441-449.

Klayman, J. 1995. "Varieties of Confirmation Bias," *Psychology of Learning and Motivation* (32), pp. 385-418.

Koriat, A., Lichtenstein, S., and Fischhoff, B. 1980. "Reasons for Confidence," *Journal of Experimental Psychology: Human Learning and Memory* (6:2), pp. 107-118.

Lee, S.-Y., Hansen, S. S., and Lee, J. K. 2016. "What Makes Us Click 'Like' on Facebook? Examining Psychological, Technological, and Motivational Factors on Virtual Endorsement," *Computer Communications* (73), pp. 332-341.

Lowrey, W. 2017. "The Emergence and Development of News Fact-Checking Sites: Institutional Logics and Population Ecology," *Journalism Studies* (18:3), pp. 376-394.

Lu, J. 2014. "What Is 'Click Bait' and Why Facebook Wants to Display Less of It," *The Washington Post*, August 26, (https://www.washingtonpost.com/news/technology/wp/2014/08/26/what-is-click-bait-and-why-facebook-wants-to-display-less-of-it).

McCracken, G. 1989. "Who Is the Celebrity Endorser? Cultural Foundations of the Endorsement Process," *Journal of Consumer Research* (16:3), pp. 310-321.

McGlohon, M., Glance, N. S., and Reiter, Z. 2010. "Star Quality: Aggregating Reviews to Rank Products and Merchants," in *Proceedings of the AAAI International Conference on Web and Social Media*, Washington, DC.

McKenzie, C. R. 2006. "Increased Sensitivity to Differentially Diagnostic Answers Using Familiar Materials: Implications for Confirmation Bias," *Memory & Cognition* (34:3), pp. 577-588.

Mills, J. 1999. "Improving the 1957 Version of Dissonance Theory," in *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology,* J. Mills and E. Harmon-Jones (eds.), Washington, DC: American Psychological Association, pp. 25-42.

Minas, R. K., Potter, R. F., Dennis, A. R., Bartelt, V., and Bae, S. 2014. "Putting on the Thinking Cap: Using Neurois to Understand Information Processing Biases in Virtual Teams," *Journal of Management Information Systems* (30:4), pp. 49-82.

Morris, J. D., Choi, Y., and Ju, I. 2016. "Are Social Marketing and Advertising Communications (SMACS) Meaningful? A Survey of Facebook User Emotional Responses, Source Credibility, Personal Relevance, and Perceived Intrusiveness," *Journal of Current Issues & Research in Advertising* (37:2), pp. 165-182.

Muntinga, D. G., Moorman, M., and Smit, E. G. 2011. "Introducing Cobras: Exploring Motivations for Brand-Related Social Media Use," *International Journal of Advertising* (30:1), pp. 13-46.

Murtha, J. 2016. "How Fake News Sites Frequently Trick Big-Time Journalists." *Columbia Journalism Review*, May 26 (https://www.cjr.org/analysis/how_fake_news_sites_frequently_trick_big-time_journalists.php).

Nahon, K. 2015. "Where There Is Social Media There Is Politics," in *The Routledge Companion to Social Media and Politics,* A. Bruns, G. Enli, E. Skogerbo, A. O. Larsson, and C. Christensen (eds.), New York: Routledge, pp. 39-55.

Nickerson, R. S. 1998. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises," *Review of General Psychology* (2:2), pp. 175-220.

ReviewMeta. 2016. "Analysis of 7 Million Amazon Reviews: Customers Who Receive Free or Discounted Item Much More Likely to Write Positive Review," ReviewMeta, June 29 (https://reviewmeta.com/blog/analysis-of-7-million-amazon-reviews-customers-who-receive-free-or-discounted-item-much-more-likely-to-write-positive-review/).

Schneider, C., Weinmann, M., and vom Brocke, J. 2017. "Digital Nudging–Influencing Choices by Using Interface Design," *Communications of the ACM* (61:7), pp. 67-73.

Shane, S. 2017. "The Fake Americans Russia Created to Influence the Election." *The New York Times*, Politics Section, September 7 (https://www.nytimes.com/2017/09/07/us/politics/russia-facebook-twitter-election.html).

Simon, H. A. 1979. "Rational Decision Making in Business Organizations," *The American Economic Review* (69:4), pp. 493-513.

Statista. 2018. "Number of Monthly Active Facebook Users Worldwide as of 2nd Quarter 2018 (in Millions)" (https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/)

Subramanian, S. 2017. "Inside the Macedonian Fake-News Complex," *Wired Magazine*, February 15 (https://www.wired.com/2017/02/veles-macedonia-fake-news/).

Sun, H. 2013. "A Longitudinal Study of Herd Behavior in the Adoption and Continued Use of Technology," *MIS Quarterly* (37:4), pp. 1013-1041.

Sydell, L. 2016. "We Tracked Down a Fake-News Creator in the Suburbs. Here's What We Learned," *National Public Radio*, November 26 (http://www.npr.org/sections/alltechconsidered/2016/11/23/503146770/npr-finds-the-head-of-a-covert-fake-news-operationin-the-suburbs).

Teng, S., Khong, K. W., Chong, A. Y. L., and Lin, B. 2017. "Persuasive Electronic Word-of-Mouth Messages in Social Media," *Journal of Computer Information Systems* (57:1), pp. 76-88.

Thaler, R. H., and Sunstein, C. R. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, CT: Yale University Press.

Thatcher, J. B., Wright, R. T., Sun, H., Zagenczyk, T. J., and Klein, R. 2018. "Mindfulness in Information Technology Use: Definitions, Distinctions, and a New Measure," *MIS Quarterly* (42:3), pp. 831-847.

Thiessen, A., and Ingenhoff, D. 2011. "Safeguarding Reputation through Strategic, Integrated and Situational Crisis Communication Management: Development of the Integrative Model of Crisis Communication," *Corporate Communications: An International Journal* (16:1), pp. 8-26.

Tormala, Z. L., Briñol, P., and Petty, R. E. 2006. "When Credibility Attacks: The Reverse Impact of Source Credibility on Persuasion," *Journal of Experimental Social Psychology* (42:5), pp. 684-691.

Tormala, Z. L., Briñol, P., and Petty, R. E. 2007. "Multiple Roles for Source Credibility under High Elaboration: It's All in the Timing," *Social Cognition* (25:4), pp. 536-552.

Vosoughi, S., Roy, D., and Aral, S. 2018. "The Spread of True and False News Online," *Science* (359:6380), pp. 1146-1151.

Wintersieck, A. L. 2017. "Debating the Truth: The Impact of Fact-Checking During Electoral Debates," *American Politics Research* (45:2), pp. 304-331.

WSJ. 2016. "Blue Feed, Red Feed," *The Wall Street Journal*, May 19 (http://graphics.wsj.com/blue-feed-red-feed/).

Yang, T. 2012. "The Decision Behavior of Facebook Users," *Journal of Computer Information Systems* (52:3), pp. 50-59.

Zhou, Z., Jin, X.-L., Vogel, D. R., Fang, Y., and Chen, X. 2011. "Individual Motivations and Demographic Differences in Social Virtual World Uses: An Exploratory Investigation in Second Life," *International Journal of Information Management* (31:3), pp. 261-271.

About the Authors

Antino Kim is currently an assistant professor at the Kelley School of Business, Indiana University. He received his Ph.D. from the Foster School of Business, University of Washington. Antino's research interests include misinformation and social media, digital piracy and policy implications, supply chain of information goods, and IT and worker displacement. His recent papers appear in *Management Science*, *MIS Quarterly*, and *Journal of Management Information Systems*.

Alan R. Dennis is Professor of Information Systems and holds the John T. Chambers Chair of Internet Systems in the Kelley School of Business at Indiana University. He was named a Fellow of the Association for Information Systems in 2012. Alan has written more than 150 research papers, and has won numerous awards for his theoretical and applied research. His research focuses on three main themes: team collaboration; fake news on social media; and information security. He is President Elect of the Association for Information Systems and co-Editor-in-Chief of *AIS Transactions on Replication Research*. He also has written four books (two on data communications and networking, and two on systems analysis and design).

Appendix

Headlines Used

The 12 Headlines Used in the Experiments

- The Humane Society Foundation Donates $100,000 to Planned Parenthood After Women's March in DC
- A Republican GOP Senator Will Not Vote to Defund Planned Parenthood
- Planned Parenthood Receives a Sum of $1,000,000 Donation from Crowd Sourcing
- Girl Scouts are Planning an Organization-Wide Fundraiser for Planned Parenthood
- Planned Parenthood Visits Campuses to Educate Young Women about the Importance of Having a Choice
- Universities Connect their Healthcare Systems with Planned Parenthood to Provide Better Care to Coeds
- Republicans Pledge to Only Fund National Pregnancy Care Center That Does Not Perform Abortions
- Pro-Life Supporters Rally in Front of Planned Parenthood Nationwide
- State Republicans Introduce New Bills to Allow Abortion Only After a Long Monitoring Period
- The State of Nevada Strengthens the Restriction on Abortion and Contraception
- Planned Parenthood Now Required to Provide Classes on Abortion Before Getting Consent for the Procedure
- On-Campus Pro-Life Supporters Significantly Reduce the Number of Abortions among Coeds