



The University of Texas at Austin
Center for Media Engagement
Moody College of Communication

STORY LABELS ALONE DON'T INCREASE TRUST

Cynthia Peacock, Gina Masullo Chen, and Natalie (Talia) Jomini Stroud

SUMMARY

The Center for Media Engagement wanted to find out if labeling news stories (using labels such as *news*, *analysis*, *opinion*, or *sponsored content*) increases trust. We tested two different types of labels – above-story labels and in-story labels. We also evaluated whether readers noticed and remembered the labels.

We found that neither type of label increased trust. Also, many people didn't notice or accurately recall the labels, although recall was better for the in-story explainer label. The results suggest that labels aren't effective on their own, but they may be helpful in conjunction with other techniques for conveying trustworthiness.

SUGGESTED CITATION:

Peacock, Cynthia, Chen, Gina Masullo, and Stroud, Natalie Jomini. (2019, October). Story labels alone don't increase trust. Center for Media Engagement. <https://mediaengagement.org/research/story-labels-and-trust>

THE PROBLEM

Public trust in the news media is lagging.¹ Some news organizations have tried to increase trust by using labels – such as *news*, *analysis*, *opinion*, and *sponsored content*.² The idea is that labeling content will make it easier for readers to differentiate, for example, opinion content from news. This project tested two types of labels to see if labeling increases trust in a news story: an above-story label often used by news organizations and an in-story explainer label. The in-story explainer label provided a more detailed description of the type of news story (e.g., “An opinion piece advocates for ideas and draws conclusions based on the author’s interpretation of the facts ...”).

We wanted to find out:

- Will labels affect whether readers trust a story?
- Will readers notice and remember the labels?
- Does the in-story explainer label work better than the above-story label?

KEY FINDINGS

- Labeling stories did not affect trust.
- Nearly half of the participants did not notice whether the story was labeled.
- Those who reported seeing a label were not particularly accurate in recalling the type of label. Of the two labels, recall was better for the in-story explainer label.

IMPLICATIONS FOR NEWSROOMS

Our findings provide little evidence that story labels *alone* affect trust. A majority of our participants either did not notice the above-story labels or incorrectly recalled them. Also, in no case did labeling a story increase participants’ trust in the news. We cannot recommend labeling as a stand-alone method for increasing trust. If news organizations want to use labels to increase awareness and understanding of the type of news being presented, the labels should be designed to explain the type of story and be placed in highly visible locations.

[Our past research](#) suggests that a *combination* of strategies to signal trust – including story labels, author biographies, footnotes for citations, descriptions of how the story was reported, and a description of involvement in the [Trust Project](#), an international consortium of news organizations – can increase trust.³ This study adds to our understanding: Labeling *on its own* does not seem to increase trust, but it may be useful alongside other indicators of trust.



Labeling *on its own* does not seem to increase trust, but it may be useful alongside other indicators of trust.

This study examined only one type of labeling and only two effects – recall and trust. At least for the articles we tested across two experiments with a total of 1,697 participants, trust was not affected. Studying other types of labels and possible effects may produce different outcomes. For example, Facebook reports that labeling stories as “breaking news” increases clicks, likes, comments, and shares of stories in some preliminary tests.⁴ Labeling news stories shared on Facebook as “disputed by third-party fact checkers” makes people think unlabeled stories are more likely to be true – even if they’re not.⁵

Although labels may not uniquely affect trust, our findings show that labels may have value from an educational standpoint. In-story explainer labels were better recalled than the more typical above-story labels. If news organizations want to experiment with labels, they should try creative, highly visible approaches, rather than just putting a label above the news story.

THE EXPERIMENTS

ABOVE-STORY LABELS

The goal of the first experiment was to test labels placed above a story. These labels were designed to resemble those used on news sites but were made more prominent to increase the likelihood people would notice them. One thousand participants were randomly assigned to view a story about electric cars or a story about politics to test if the labels worked similarly across topics.⁶ For each story, people saw the story with no label or with a *news*, *analysis*, *opinion*, or *sponsored* label. The stories were the same, regardless of the label, so we could see if it was the label – not the story – that influenced people. The stories were displayed on a mock news site.

The News Beat

Opinion

Battery That Charges in One Minute Could Beat Tesla in the Race for Tomorrow's Clean Car

Posted on July 23, 2018

By Jim Phipps
The News Beat Staff Reporter

Example of an Above-Story Label for an Opinion Story

Many participants did not recall the label accurately

After reading the story, participants were asked which label they recalled seeing.⁷ Overall, people did not notice the labels.

In fact, what label people remembered had little bearing on what label they were shown.⁸ Overall, 45% reported that they did not notice whether the article was labeled or not. This percentage did not vary depending on whether the article actually was labeled.

Thirty-one percent reported that the article was labeled *news*. Again, this was the same regardless of whether the article was not labeled, was labeled *news*, or was labeled with another category (*opinion*, *analysis*, or *sponsored*). Eight percent thought the article was labeled *analysis*, 6% said the article was not labeled, and 2% said it was labeled *sponsored*. These percentages did not vary depending on how the article was actually labeled.

The only place we see any evidence of labels affecting what people recalled was with the *opinion* label. Here, 13% identified the article as being an opinion article when it was labeled *opinion*. This is higher than the percentage who thought the article was an opinion article when it was actually labeled *analysis* (4.5%) and who thought the article was an opinion article when it actually was not labeled (3.5%).⁹

We analyzed whether the effect of the story labels on the ability to recall the label varied based on participants' backgrounds, including their age, race, ethnicity, education, income, gender, political ideology, and political partisanship.¹⁰ Only one variable seemed to matter: age. The younger the participant, the more likely they were to recall the label correctly when the story was labeled *news* or *opinion*.¹¹

Labels on a story did not increase trust in the news story

Next, we assessed whether labeling influences trust in the news story.¹² After viewing the story, participants were asked to rate the story on a series of 18 items related to trust. These 18 items included ratings of how *fair*, *unbiased*, *balanced*, and *trustworthy* the story was, along with other items.¹³ Ratings for these 18 items were averaged together into one trust score. On average, people rated their trust of the news article as 3.5 on a 1 to 5 scale, with higher values indicating more trust.

Results showed that trust ratings did not vary based on whether people saw an article with a *news* label, an *analysis* label, an *opinion* label, a *sponsored* label, or no label at all.¹⁴ Labeling the story did not affect trust.¹⁵

We analyzed whether the labels affected trust differently depending on people's backgrounds. Few differences emerged.¹⁶

ABOVE-STORY VERSUS IN-STORY LABELS

In a second experiment, we wanted to see if varying the type of news label would affect recall and trust. We compared the above-story label used in the first study to an in-story explainer label and no label at all. The in-story explainer label provided definitions of each label based in part on those proposed by the [Trust Project](#). These labels were included between the byline and the beginning of the article text. The idea was that the in-story label might be more noticeable because it was not where labels are usually placed, and that explaining the type of story may help people to understand what the terms meant.

The 697 participants were randomly assigned to see a story with the in-story label, the above-story label used in the first experiment, or no label. They also were randomly assigned to see one of four stories from real news sites that depicted the four categories of stories: *news*, *analysis*, *opinion*, and *sponsored content*.¹⁷

The News Beat

Gender Parity Needed In Congress

Posted on March 4, 2019

By Jim Phipps
The News Beat Staff Reporter

This article is from our opinion section. An opinion piece advocates for ideas and draws conclusions based on the author's interpretation of the facts and data. Opinion pieces may include reported facts or quotes, but emphasizes the author's own thoughts, personal preferences and conclusions.

Example of an In-Story Label for an Opinion Story

People recall the in-story explainer label better than the above-story label

Overall, in-story explainer labels increased the likelihood that people would recall the correct label compared to those who did not see the label and those who saw above-story labels.¹⁸

As before, many people had no recollection of whether the story they saw was labeled or not. Sixty-three percent of those who saw an article without a label said they did not recall whether there was a label (25% correctly recalled that it was not labeled). Fifty-eight percent of those who saw an article with an above-story label could not recall whether the article was labeled (24% correctly recalled what the article was labeled). More

encouraging, only 24% of those who saw an article with an in-story explainer label could not recall whether the article was labeled, and 66% correctly recalled the article label. These differences in recall are statistically significant.

There was no evidence that the labels worked differently based on participants' backgrounds, including their age, race, ethnicity, education, income, gender, political ideology, and political partisanship.¹⁹

Neither label type increases trust in the news story

Next we examined whether the type of label – above-story or in-story explainer – would affect people's trust in the news story. After viewing the story, participants were asked to rate the story and the news outlet on the same items regarding trust used in the first experiment.

Results showed that neither type of label increased how much people trusted the news story.²⁰

We again examined whether the effect of the labels on trust varied based on the participants' backgrounds. Here, one difference emerged: education.²¹ Among those with less than a college degree, sponsored content with an in-story label resulted in lower trust in the article compared to the above-story label. Those with less than a college degree rated sponsored content a 3.6 when it was accompanied by an above-story label and 2.9 when it was accompanied by an in-story explainer label, on a 1 to 5 scale with higher scores indicating more trust.

METHODOLOGY

Both experiments were embedded in Qualtrics-based online surveys. Participants were recruited in different ways for each study. For the first experiment, participants were recruited using Survey Sampling International²² between July 26 and July 30, 2018. SSI matched our sample to the adult U.S. Internet population for demographics in terms of age, race/ethnicity, gender, and education based on a nationally representative, random sample survey conducted by Pew Research Center. Participants for the second experiment were recruited using TurkPrime on March 17 and 18, 2019. TurkPrime²³ culls participants from Amazon.com's Mechanical Turk, an opt-in online tool that allows participants to complete tasks for payment but blocks suspicious geolocations to guard against bots.

Both surveys and survey procedures were nearly identical across the studies. After participants answered screening questions to verify their age and U.S. residency, they were randomly assigned to experimental conditions. In the first experiment, participants all saw

the same story attributed to a mock news site but it was labeled *news*, *analysis*, *opinion*, *sponsored*, or not labeled at all. Two articles were selected to understand whether the conclusions were robust across topics: a science article about a new type of car battery and a political article about President Trump and the Pentagon looking to end the war in Afghanistan. Both articles were edited for length. No differences emerged across the two articles in the effects of the labels. For the second experiment, participants were randomly assigned to see a story attributed to a mock news site with an above-story label, an in-story explainer label, or no label. They were also randomly assigned to view a news, analysis, opinion, or sponsored story. The stories were based on real stories that reflected each story type, and they were edited to be similar in length and to ensure they reflected the four story types. All the stories related to gender makeup of the U.S. Congress. In both experiments, participants rated the news story they viewed on items that relate to trust and answered questions about their background, including their education, income, gender, and race. We also asked participants to report their ideology (Experiment 1: very conservative 13.3%, conservative 21.0%, moderate 39.4%, liberal 17.1%, very liberal 9.2%; Experiment 2: very conservative 6.7%, conservative 18.5%, moderate 26.5%, liberal 30.4%, very liberal 17.8%) and their political partisanship (Experiment 1: Strong Republican 19.1%, Not strong/leaning Republican 20.4%, Independent 13.4%, Not strong/leaning Democrat 23.3%, Strong Democrat 23.8%; Experiment 2: Strong Republican 11.9%, Not strong/leaning Republican 17.5%, Independent 10.7%, Not strong/leaning Democrat 33.7%, Strong Democrat 26.3%).

PARTICIPANT DEMOGRAPHICS

	U.S. Adult Internet Population	Experiment 1 <i>n</i> = 1,000	Experiment 2 <i>n</i> = 697
Gender			
Male	49%	43%	54%
Female	51	56	45
Race/Ethnicity			
Hispanic/Latino	15%	12%	7%
White	70	79	81
Black/African-American	13	11	9
Asian	5	4	7
Other	12	5	4
Age			
18-29	24%	26%	25%
30-49	36	40	55
50-64	25	28	16
65+	15	5	4
Household Income			
Less than \$30,000 annually	31%	28%	21%
\$30,001 to \$50,000 annually	18	17	25
\$50,001 to \$75,000 annually	14	23	27
\$75,001 or more annually	37	32	27
Education			
High school or less	34%	27%	13%
Some college or Associate's	33	40	33
College degree or More	33	35	55

Note: U.S. Internet population is based on data from Pew Research Center when data were collected in January 2018.

ENDNOTES

¹ Jones, J. (2018). U.S. media trust continues to recover from 2016 low. Retrieved from <https://news.gallup.com/poll/243665/media-trust-continues-recover-2016-low.aspx>; Knight Foundation (2018, June 20) Perceived accuracy and bias in the news media. Retrieved from <https://knightfoundation.org/reports/perceived-accuracy-and-bias-in-the-news-media>

² Micek, J. (2017). We're making it easier to distinguish news from opinion. Retrieved from https://www.pennlive.com/opinion/2017/07/were_making_it_easier_to_disti.html; Stead, S. (2017). Public editor: Lines between opinion, analysis and news need to be clearer. Retrieved from <https://www.theglobeandmail.com/community/inside-the-globe/public-editor-lines-between-opinion-analysis-and-news-need-to-be-clearer/article34310619/>; see also Iannucci, R. & Adair, B. (2017). Reporters' Lab study finds poor labeling on news sites. Retrieved from <https://reporterslab.org/news-labeling-study-results-media-literacy/>

³ Curry, A., & Stroud, N.J. (2017, December 12). Trust in online news. Center for Media Engagement. Retrieved from <https://mediaengagement.org/research/trust-in-online-news/>

⁴ Rhyu, J. (2018, March 3). Facebook for media blog. Retrieved from <https://www.facebook.com/facebookmedia/blog/enabling-publishers-to-label-breaking-news-on-facebook>

⁵ Pennycook, G., Bear, A., Collins, E., & Rand, D. G., (2019, March 15). The implied truth effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings. Available at SSRN: <https://ssrn.com/abstract=3035384> or <http://dx.doi.org/10.2139/ssrn.3035384>

⁶ Of the 1,046 survey completions, we excluded duplicate IP addresses ($n = 12$) and those who reported that they were less than 18 years old ($n = 17$), not U.S. residents ($n = 4$), or unable to view the article page ($n = 13$).

⁷ Participants were asked: "Which label (if any) appeared on *The News Beat* article that you just saw?" They could select one answer from six choices: *news*, *analysis*, *opinion*, *sponsored*, *no label*, or *did not notice whether the article was labeled or not*.

⁸ This was tested using a chi-square statistic, $\chi^2[20] = 30.22$, Cramer's $V = .09$, $p = .06$. When looking at the science and politics article separately, the results are slightly different. For the politics article, the differences are not significant using Bonferroni-corrected pairwise comparisons, but the opinion label results in the highest percentage of people saying that the article was labeled opinion (14.3%), whereas the news label resulted in fewer people saying that the article was labeled opinion (7.1%) as did the analysis (7.0%), sponsored (7.8%), and no label (5.9%) conditions. For the science article, those seeing an opinion label were significantly more likely to say that the article was labeled opinion (11.3%) than those who saw no label (1.0%), but the percentage was not significantly different from those who said that the article was labeled opinion when they saw a news label (7.3%), saw an analysis label (2.0%), or saw a sponsored label (10.1%).

⁹ This was tested using Bonferroni-corrected pairwise comparisons.

¹⁰ We tested this using a logistic regression with stepwise entry for all two- and three-way interactions between the listed background characteristics, the story topic, and the label. Age emerged as a significant predictor, Wald $\chi^2[4] = 19.80$, $p < .01$.

¹¹ When the article was labeled *news*, a 25-year-old had a 46% chance of correctly recalling that it was labeled *news*, whereas a 55-year-old had only a 20% chance. When the article was labeled *opinion*, a 25-year-old had a 19% chance of correctly recalling the label compared to a 6% chance for a 55-year-old. The older the person, the more likely to recall the label correctly when the story was not labeled. When the article was not labeled, a 25-year-old had a 5% chance of correctly recalling that it was not labeled, whereas a 55-year-old had a 13% chance. There were no other differences.

¹² While some scholars view trust and credibility as separate concepts, others use the concepts interchangeably. In general, both concepts rely on similar notions, including, fairness, accuracy, and separating fact from fiction, among other dimensions. (See Kohring, M. & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, 34(2), 231-252. doi: 10.1177/0093650206298071; Johnson, T.J. & Kaye, B. K.(2014). Credibility of social network sites for political information among politically interested Internet users. *Journal of Computer-Mediated Communication*, 19, 957-974. doi: 10.1111/jcc4.12084). We focus on trust because it is a concept more readily accessible to a non-academic audience, but it includes related concepts, such as credibility, fairness, and accuracy, among other dimensions.

¹³ Participants rated the 18 items on a 1 (*strongly disagree*) to 5 (*strongly agree*) scale adapted from Yale, R. N., Jensen, J. D., Carcioppolo, N., Sun, Y., & Liu, M. (2015). Examining first-and second-order factor structures for news credibility. *Communication Methods and Measures*, 9(3), 152-169; Meyer, P. (1988). Defining and measuring credibility of newspapers: Developing an index. *Journalism Quarterly*, 65(3), 567-574. The items were *fair, unbiased, tells the whole story, accurate, respects people's privacy, can be trusted, separates facts from opinion, watches out after my interests, concerned mainly about the public interest, balanced, objective, reports the whole story, honest, believable, trustworthy, up-to-date, current, and timely*. They were averaged together into a composite score.

¹⁴ This was tested using an ANOVA, ($F[4, 994] = 2.26, p = .06$). Although there were differences by topic ($F[1, 994] = 42.22, p < .01$), whereby the science article was rated as more trustworthy than the politics story, the effect of labels did not vary by condition, as an interaction between label and topic was not significant, $F[4, 990] = 0.44, p = .78$).

¹⁵ We also examined if labeling stories would make people trust the news outlet more. How people rated the story was so highly correlated with how people rated the news outlet that we were unable to meaningfully distinguish between the two.

¹⁶ We tested this by entering all two- and three-way interactions between the listed background characteristics, the type of story, and the type of label in an ANOVA. Significant interactions were retained. There was a significant interaction between the label used and ideology $F(4, 960) = 2.39, p = .05$. Those identifying as more conservative rated the article as more trustworthy than did those identifying as more liberal when the article was not labeled. Yet when the article was labeled, people rated the article similarly regardless of whether they were liberal or conservative and regardless of the type of label. The interaction is only present when ideology is treated as a linear predictor. Note that there was also a significant three-way interaction between the type of story, the type of label, and identifying as Hispanic / Latino ($F(4, 960) = 3.02, p < .05$). Examining the differences by label depending on the story type and identifying as Hispanic / Latino and using a Sidak correction turned up no significant differences. There also were conditions with a small sample size. We treat this result as exploratory and note it here in case it persists in future studies.

¹⁷ Of the 808 survey completions, we excluded duplicate IP addresses ($n = 32$) and respondents who were not U.S. residents ($n = 1$), were unable to view the article page ($n = 2$), or failed one of three attention check embedded in the survey ($n = 76$).

¹⁸ This was tested using a logistic regression model, which showed that the in-story explainer label significantly increased the probability people would correctly name the label displayed relative to the no label condition ($B = 1.82, SE = 0.21, p < .001$) and relative to the above-story label condition ($B = 1.84, SE = 0.21, p < .001$). There was a significant main effect of the article type (Wald $\chi^2[3] = 12.84, p < .01$). Overall, the sponsored story was recalled correctly more often than the other story types. There was not a significant interaction between the label type and the story type (Wald $\chi^2[6] = 3.13, p = .79$).

¹⁹ We tested this using a logistic regression with stepwise entry for all two- and three-way interactions between the listed background characteristics, the type of story, and the type of label. None of the interactions were significant.

²⁰ This was tested using an ANCOVA with the label type as the key independent variable and story type as a covariate. Results showed that the type of label did not influence trust in the story ($F[2, 691] = 0.85, p = .43$). There was a significant main effect of the story type ($F[3, 691] = 17.06, p < .001$), whereby the news article was trusted more than the other story types and sponsored content was trusted less than both news and analysis content (differences computed using a Sidak correction).

²¹ We tested this by entering all two- and three-way interactions between the listed background characteristics, the type of story, and the type of label in an ANOVA. Significant interactions were retained. There was a significant three-way interaction, including all lower-order interactions and main effects, between education, the type of story, and the type of label. To aid with interpretation and to ensure sufficient sample size, we collapsed education into two categories: those with less than a college degree and those with a college degree or more. The three-way interaction was significant, $F(6, 673) = 2.89, p < .01$.

²² Survey Sampling International has since been rebranded as Dynata.

²³ Information retrieved from <https://blog.turkprime.com/after-the-bot-scare-understanding-whats-been-happening-with-data-collection-on-mturk-and-how-to-stop-it>.