

A Behavior Analytic Approach to Promote Fact-Checking on Social Media

Carol E. Williams, MA, BCBA
The Chicago School of Professional Psychology.

Introduction

The purpose of this study was to examine factors that may have an effect on the behavior of social media users, specifically relating to fact checking information and preventing the sharing of false information.

- Proliferation of false information suggest that society is devolving from the information age to the disinformation age.
 - 93% of Americans report use of online sources to access news (Mitchell et al. 2017)
 - Current limits on social media behavior are platform driven policies that are unilateral and reactive.
 - Once posted, information is forever, unlimited by topic or platform.
 - Social media platforms have no viable means to permanently remove false information, (Khosravi & Unger, 2015; Kullar et al. 2020).
- Many accept information as true rather than question it as false (Lee, et al., 2006; Dahlstrom, 2010; Gilbert, et al., 1990).
 - Coherence / biases exacerbate condition of acceptance
 - Rule governed behavior less sensitive to contingency management (e.g., consequences for proliferation of false information)
- Social media behavior is verbal behavior.
 - Subject to motivations, contingencies, rule-governance, environment, etc.
 - Behavior best addressed from a bilateral approach (platform policies and user accountability)
 - Contingency management
 - Response effort
 - Response cost / publicizing performance
 - Other variables: Reduced sensitivity to contextual learning, and influence of rule governance and coherence

Methodology

Quasi-behavioral experiment: Single subject reversal ABACAD design used to prevent possible carryover effects

- Total of 30 participants (adults, 18 years or over, social media users)
- Post content collected from the top factually accurate and nonpartisan online sources and top nonfactual / fake news sources. Content was cleansed of identifying source data prior to administration to participants.
- Data collection procedures
 - Physical manipulatives used to simulate Twitter environment: 6 index cards per person per condition
 - Equal number of false/true content provided to all users though randomly dispersed.
 - Baseline conditions: Content provided with no accompanying verifying information
 - Intervention conditions:
 - B: Verifying information provided along with the content
 - C: Verifying information is provided if user opts to fact-check by checking a box and flipping over the index card.
 - D: Same as C and user behavior of fact-checking/sharing posted as % value and clearly visible to all other participants
- Dependent variables:
 - Fact-checking behavior and Sharing behavior
- Independent variables:
 - Provision of convenient fact-checking resource (condition C)
 - Fact-checking source and publicizing performance (D)
- Reliability and Validity
 - Experimental effect of changes in DVs that covary based on manipulation of IVs
 - Use of multiple settings and multiple participants represents an inter-subject replication design
 - Experiments limited to one setting administration with same constructs for all participants.
 - Demographics survey held to end of experiment to avoid potential bias
 - Participants solicited from general population of social media users.

Implications

- Implications for practice: The reversal design used in this study allowed for the demonstration of changes in sharing behavior between conditions. Where fact checking was conveniently available, sharing behavior increased, and the sharing behavior decreased in the return to each of the baseline conditions, suggesting a relation between a fact checking resource and sharing behavior.
- Implications for research: Coherence is implicated in this study through sharing of information consistent with the participant's own views despite providing information that the content is false. A bilateral approach to stemming the flow of false information on social media platforms would include user accountability measures and consequences related to fact checking prior to sharing information. Research shows that users are more attuned to visual information. Thus, Twitter, could consider a simple color-coding method for users to quickly identify users/sources that are reputable and implement stringent fact checking prior to posting.
- Key contribution: Single subject design studies in the area of social media, demonstrating that ABA is applicable to examine, and effect behavior change in all populations.
- Key feature: Use of both statistical analysis and a single subject research design. The contention that statistical analysis is not appropriate for research using single subject design is based in the premise that the clinical significance in single subject design is in the individual differences, observable through visual inspection. Conversely, statistical analysis is the preferred method to provide generalization to larger human populations. In this study, statistical analysis was used to augment visual graphical results, and support, through detailed inferential analysis, exacting differences in behaviors. Statistics were not used to deduce causal inferences, but rather, to more specifically analyze differences and similarities that were apparent based on visual analysis.

Results

Research Question 1: Does providing a convenient fact-checking source change user behavior with regard to sharing information?

Figure 1. Ratio: Liked, Shared, True; Baseline to Intervention

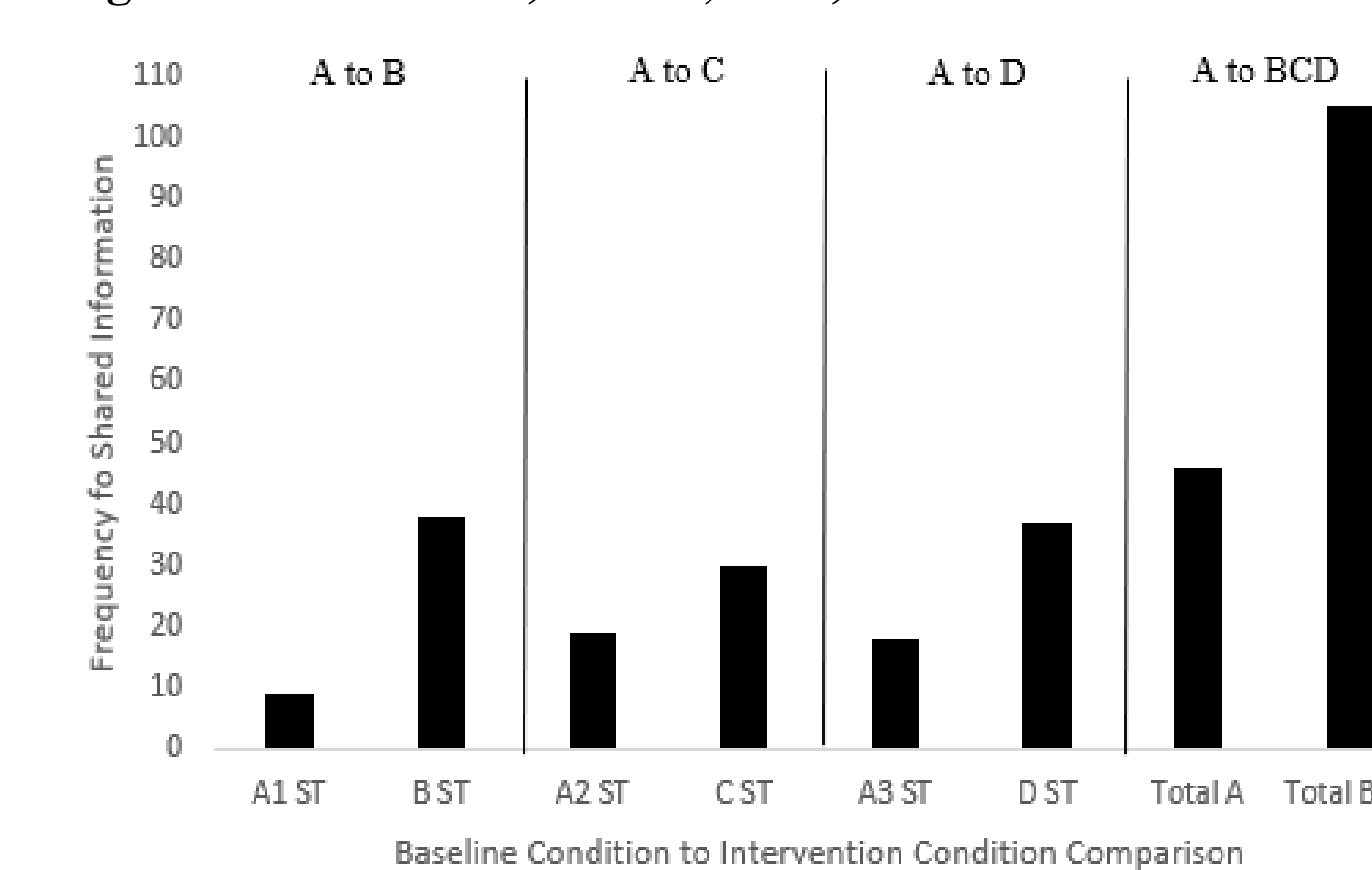


Figure 2. False Content, True Content, Shared, Liked

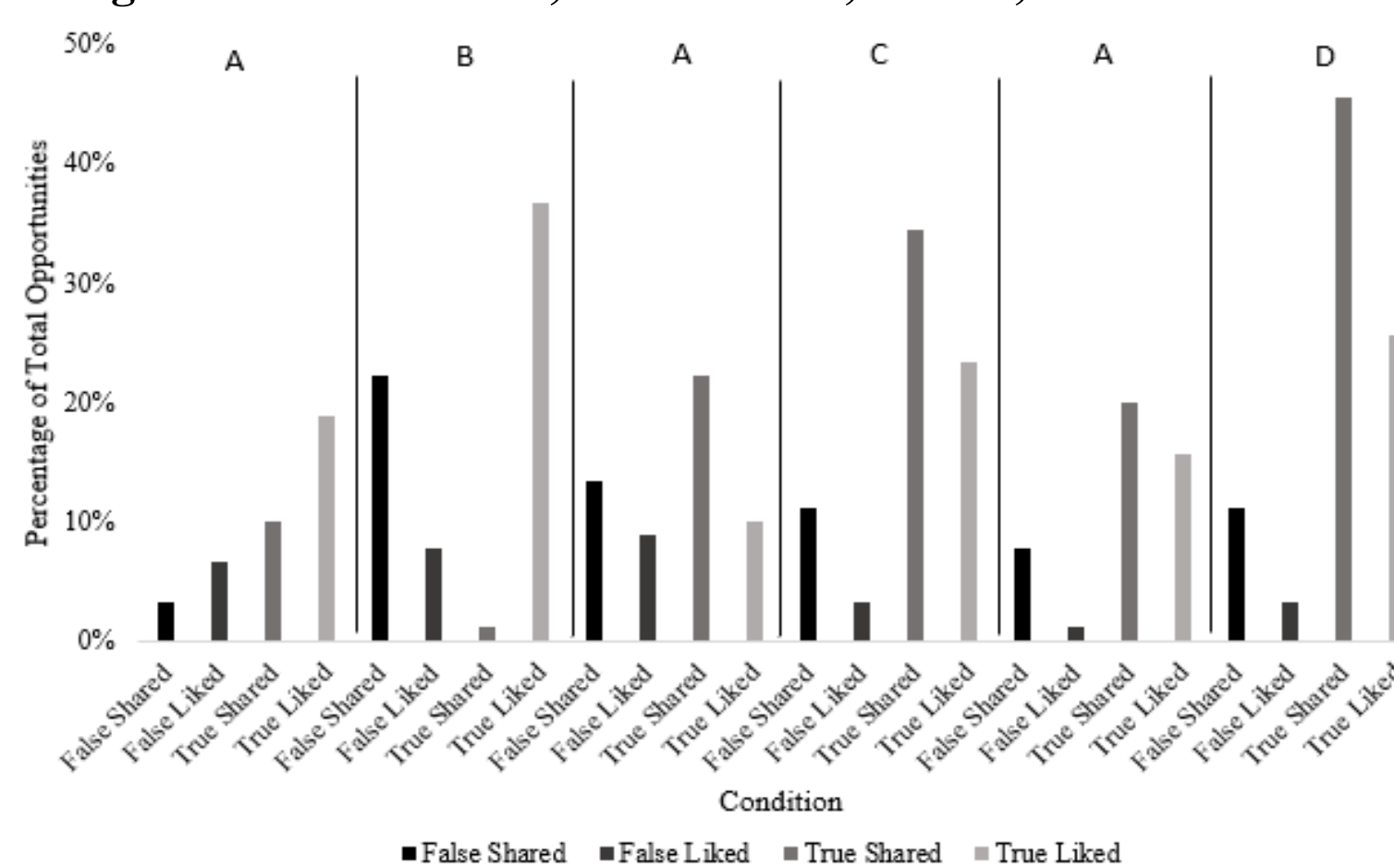


Figure 3. Liked/Not Liked Fact-Checked/Not Fact-Checked

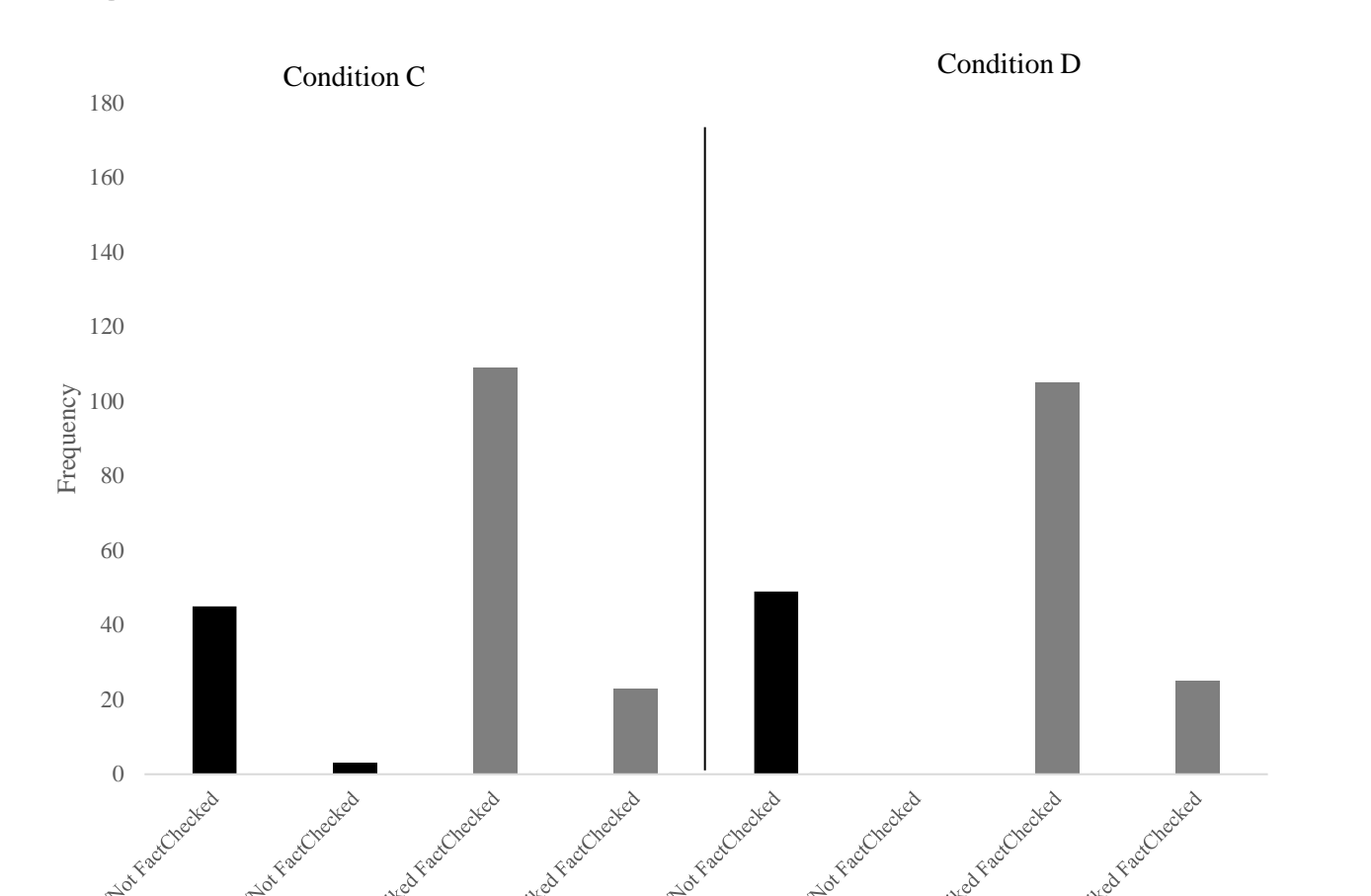


Figure 4. Liked/Not Liked Shared, Not Fact-Checked

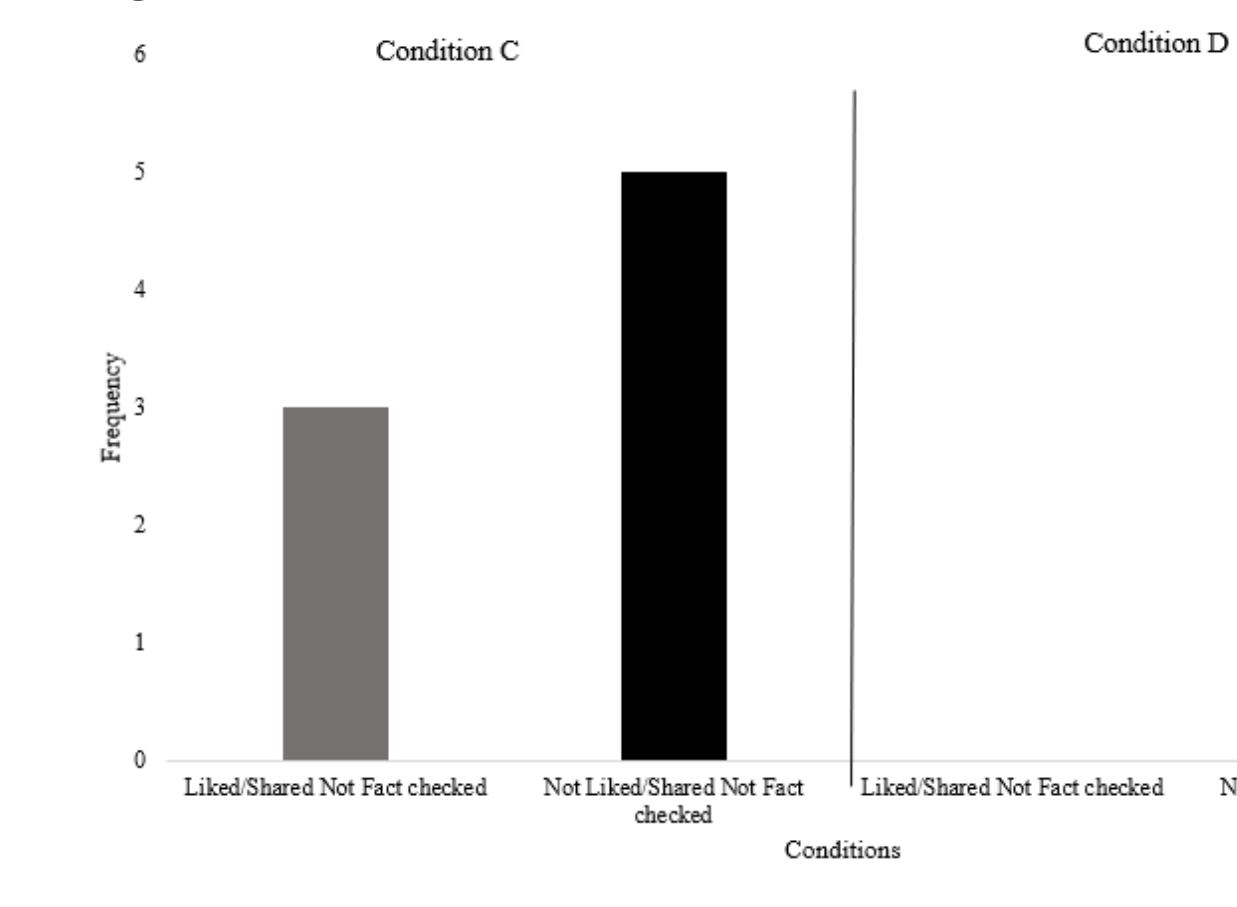


Figure 5. Not Liked / Liked Shared False Content

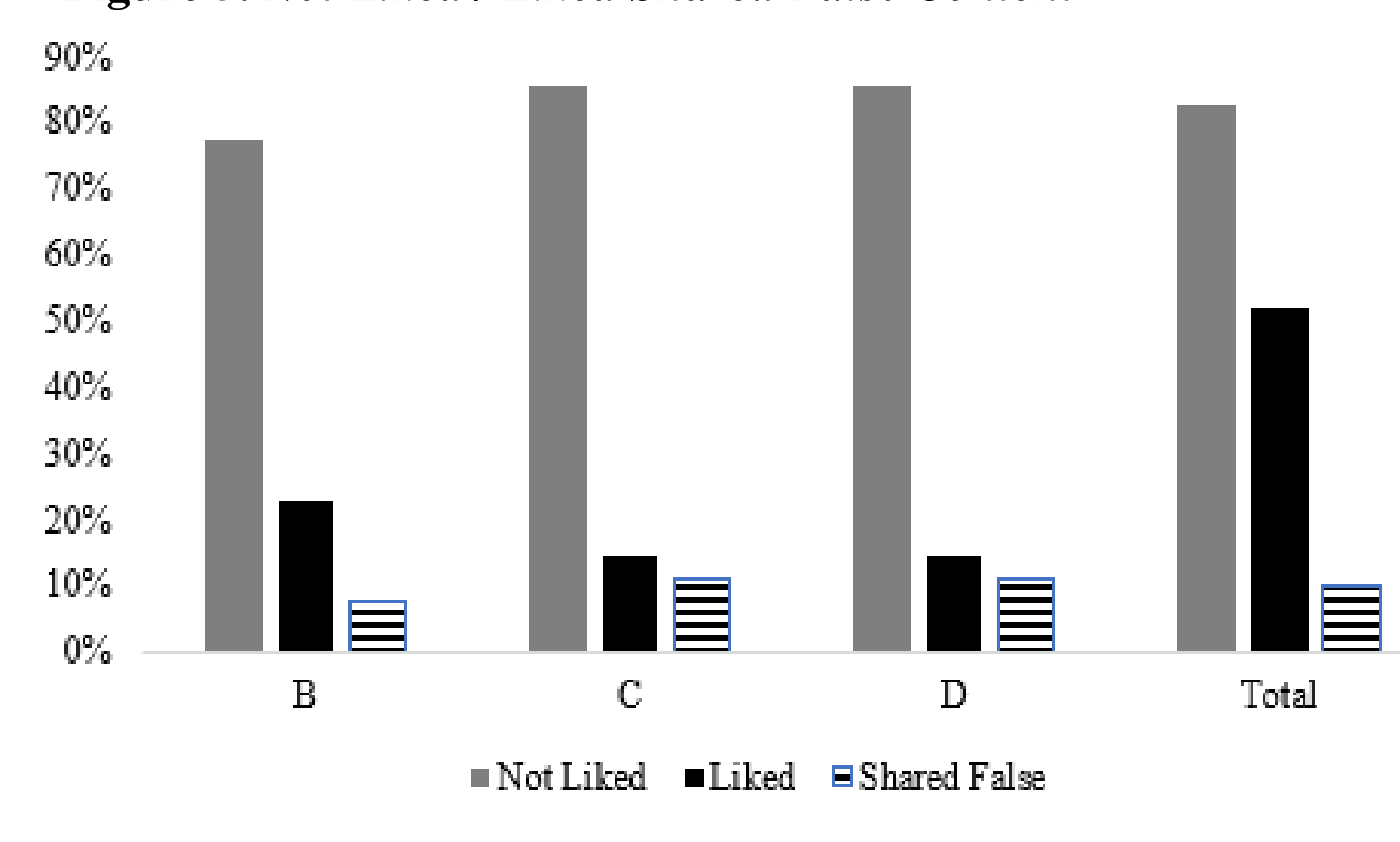
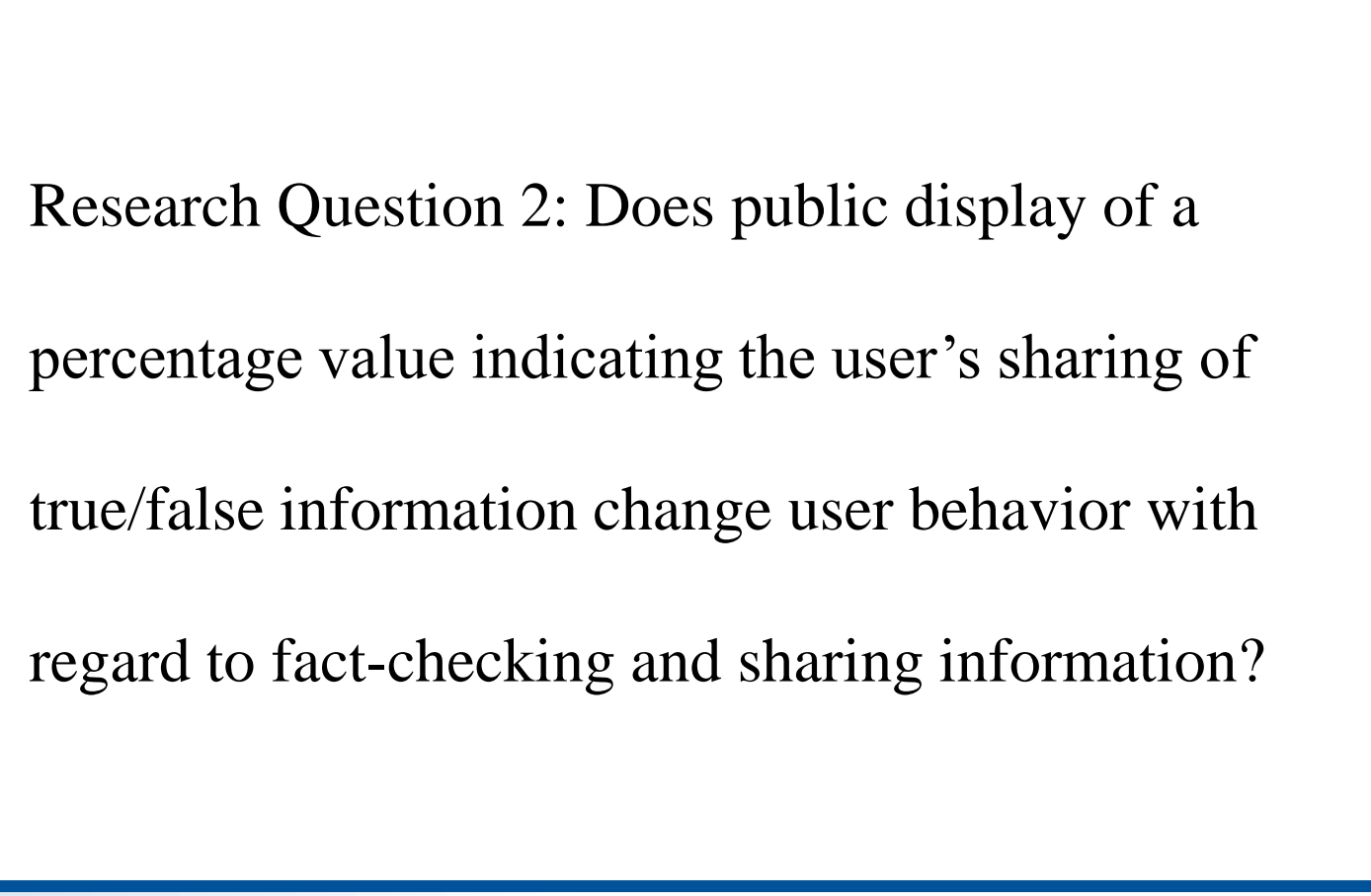
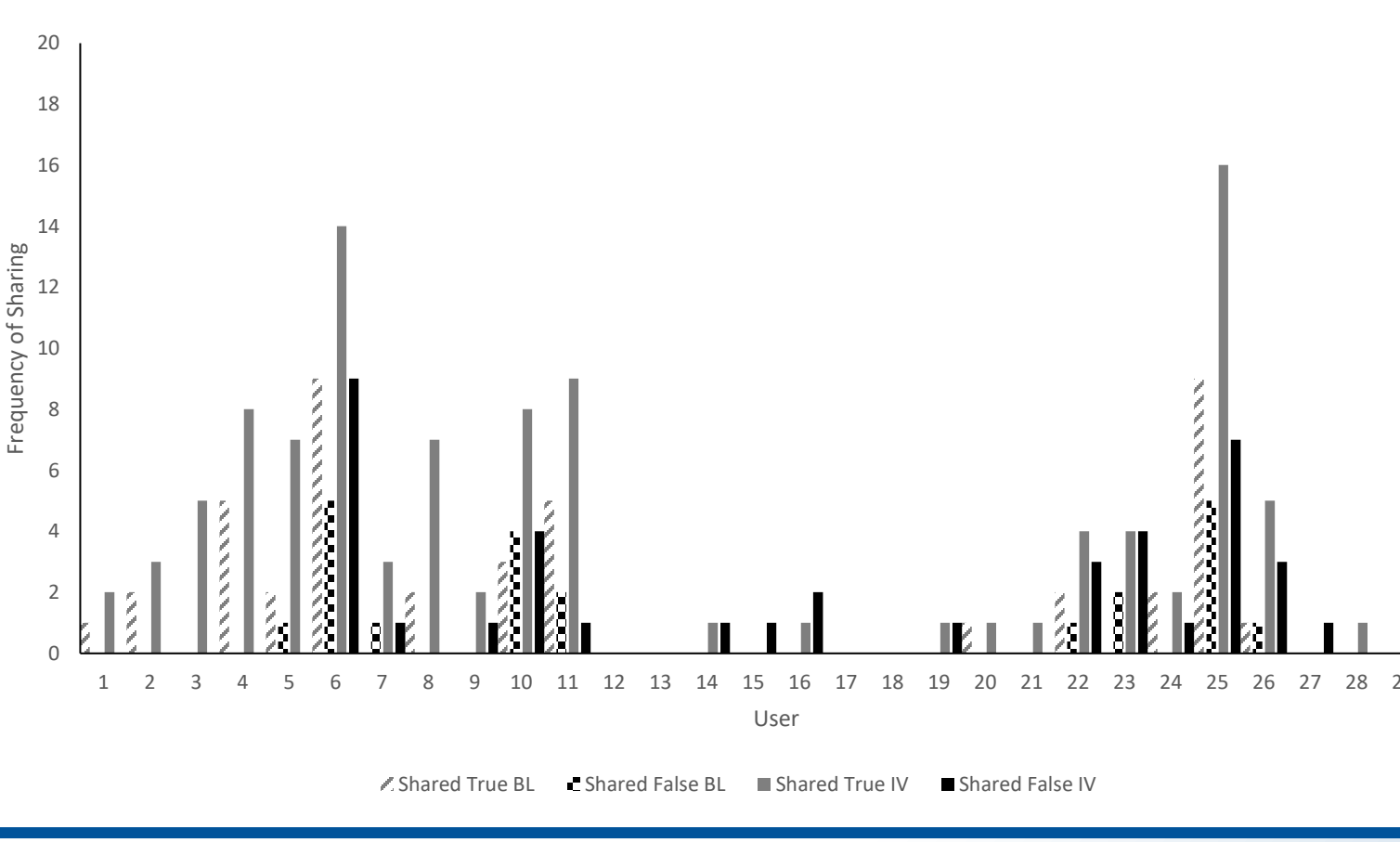


Figure 6. Sharing Fact-Checked False Content



Research Question 2: Does public display of a percentage value indicating the user's sharing of true/false information change user behavior with regard to fact-checking and sharing information?

Figure 7. False / True Content Shared in Baselines and Interventions



Acknowledgements

- Special thanks for the guidance and assistance provided by the members of my dissertation committee:
 - Dr. Siri Ming (Chair)
 - Dr. Ruth Anne Rehfeldt
 - Dr. Frankie Ballard
- Additional thanks to Dr. Mary Caruso-Anderson for guidance in the development of the conceptual formation of this project.
- The author declares no conflict of interest
- There has been no applicable funding

Discussion

Providing a convenient means to fact check information significantly increased the rate of sharing true content compared to baseline (H1₁) (Figure 1)

- There was a statistical difference between the sharing of true information between baseline (A1) with no fact-checking (M=.05, SD=.22) and condition B (M=.01, SD=.08), t(2.57), df=179, p=.01; and condition A3 (M=.11, SD=.32) and condition D (M=.23, SD=.42), t(-3.46), df=179, p=.001. There was not a statistical difference in sharing true information between condition A2 (M=.11, SD=.32) and condition C (M=.17, SD=.38), t(-1.82), df=179, p=.07.

Providing a convenient means to fact-check information did not significantly decrease sharing of false information. (H1₂) (Figure 2).

Compared to users who did not endorse content, users who endorse (like) content were not significantly less likely to fact check that content. (H1₃) Users who liked content did not fact check. (Figure 3).

Compared to users who did not endorse content, users who endorse (like) content were not significantly more likely to share that content without fact-checking. (H1₄) (Figure 4)

Compared to users who did not endorse content, users who endorse (like) content were significantly more likely to share content despite information verifying the content to be false (H1₅). Users who endorsed false content were more likely to share the false content (Figure 5).

- There was a statistical difference between liking and sharing false content (M=.07, SD=.25) and not liking and sharing false content (M=.13, SD=.34), t(-9.300), df=178, p=.05

Providing convenient fact-checking and a display of a percentage value on the user's ratio of sharing verified accurate information versus false information, did not significantly increase fact-checking compared to fact-checking alone. (H2₁) (Figure 6)

Compared to users who did not endorse false content, users who endorsed (like) false content were significantly more likely to share that content despite a negative percentage value rating. (H2₂) (Figure 6).

Overall sharing of false content occurred at low rates compared to overall sharing and sharing of true content (Figure 2).

Of note, false content was shared a total of 62 times across all conditions out of a possible 540 opportunities (11.49%). Individual analysis shows that two users were responsible for sharing and endorsing false content in excess of 40% of opportunities; sharing false content in 14 out of 90 and 12 out of 90 opportunities, respectively (Figure 7).

While sharing of false content did not significantly decrease overall, rates were trending lower following condition B and rates of sharing true content significantly increased.

Fact-checking was utilized when offered on most occasions. Frequency statistics determined that in condition C participants chose to fact-check in 132 out of the 180 opportunities (73.3%), and in condition D, participants chose to fact-check in 139 out of the 180 opportunities (77.2%).

References

Dahlstrom, M. F. (2010). The role of causality in information acceptance in narratives: An example from science communication. *Communication Research*, 37(6), 857-875. <https://doi.org/10.1177/0093650210362683>

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59(4), 601-613. <https://doi.org/10.1037/0022-3514.59.4.601>

Lee, L., Frederick, S., & Ariely, D. (2006). Try it, you'll like it: The influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, 17(12), 1054-1058. <https://doi.org/10.1111/j.1467-9280.2006.01829.x>

Mitchell, A., Gottfried, J., Shearer, E., & Lu, K. (2017, February 09). *How Americans encounter, recall, and act upon digital news*. Pew Research Center. <https://www.journalism.org/2017/02/09/how-americans-encounter-recall-and-act-upon-digital-news/>