

Truth & Lies on Social Media during COVID-19: Applying Mediated Statement Analysis (MSA) for Digital Deception Detection

Margaret C. Stewart, Ph.D

Associate Professor, Communication Studies
School of Communication

Christa L. Arnold, PhD.

Associate Professor, Communication Studies
School of Communication

F. Dan Richard, PhD.

Professor, Psychology

University of North Florida
1 UNF Drive, Jacksonville, FL United States

Abstract: Our research introduces Mediated Statement Analysis (MSA) as a developing research methodology comprised of nine (9) categories of textual & linguistic cues to assist with detecting potentially deceptive social media posts online. Our pilot data of 1,414 truthful-untruthful social media posts uncovered a subset of data and subsequent preliminary findings related to the COVID-19 pandemic. We apply MSA to the COVID pilot data subset (n = 130) to analyze patterns of textual and linguistic cues in truthful-untruthful social media posts across three (3) MSA categories. Exploring MSA's application to social media posts about COVID is relevant in the present-day crisis of misinformation online, related to health communication narratives and lived experiences during the global pandemic.

Keywords — Deception online, mediated communication, crisis communication, mediated statement analysis (MSA), social media

SUGGESTED CITATION: Stewart, M.C., Arnold, C.L., & Richard, F.D. (2025). Truth & lies on social media during COVID-19: Applying mediated statement analysis (MSA) for digital deception detection. *Proceedings of the 2025 International Crisis and Risk Communication Conference*, 13(1) 159-162. <https://www.doi.org/10.69931/001c.142918>

INTRODUCTION

Mediated Statement Analysis, or MSA, is introduced to the International Crisis & Risk Communication Association (ICRCA) community as an applied research methodology to detect potentially deceptive social media posts. Preliminary analysis of the pilot data comprised of 1,414 truthful-untruthful social media posts uncovered a subset of data related to the COVID-19 pandemic. Applying MSA to this specific data subset consisting of 130 truthful-untruthful social media posts to analyze patterns of textual and linguistic cues reveals meaningfulness of the emerging methodology for deception detection online in addition to findings about truthful-untruthful social media content related to the pandemic. The rationale supporting this research inquiry stems from the present-day crisis of misinformation online and, more specifically, across

ISSN: 3068-6539

© 2025 Copyright is held by the owner/author(s).

Publication rights are licensed to ICRCA.

<https://doi.org/10.69931/001c.142918>

social media platforms related to health communication narratives and lived experiences during the global pandemic.

PURPOSE & DEVELOPMENT OF MEDIATED STATEMENT ANALYSIS (MSA)

“The major moral crisis of our times” describes the presence and persistence of misinformation on the Internet by researchers [1, 2]. Despite the pervasiveness of misinformation online, identifying deceptive content online is effectively challenged given the lack of a standard methodology to detect untruthful digital content [3]. According to Petrescu et al. [4] deceptive digital messaging tends to manipulate key elements of communication, such as the level of detail, description, accuracy, relevance, or clarity as common tactics used by individuals for purposes of online deception [4]. Digital deception and misinformation are not uniform, further complicating the ability for adequately detection. The continuum of digital deception continuum ranges from posting content and ideas containing subtle inaccuracies or exaggerations, to passing off personal opinions as fact, to innocently and unknowingly sharing incorrect information, to intentionally disseminating lies for purposes of harmful intent. The evolution of social media algorithms can compound the spread of misinformation by learning about the traits, interests, values, and preferences of Internet users to signal and serve more engaging and preferential content to online consumers and social media actors [5]. At the beginning of 2025, 5.56 billion Internet users and over 5 billion social media users are reported across the globe, which accounts for roughly 68% and 64% of the world’s population respectively [6]. Social media users average an upwards of 2 hours per day engaging in digital activities and are active on 6+ unique social media platforms [7].

For these reasons, coupled with the present digital culture where mediated communication technologies and social media are ever-increasing and diversifying, our research team identified an opportunity and need to develop a method to aid in digital deception detection. Mediated Statement Analysis (MSA) emerges as an applied research method for online deception data analysis. MSA derives from Statement Analysis (SA), a method of analysis used in law enforcement to examine textual artifacts and transcripts of criminal testimonies, interview statements, and fraud investigation to identify suspicious content that may benefit from additional probing and evaluation. The original SA method examines exactly what people say or write in text to evaluate the statements for deviations in language cues or patterns that may indicate the presence of untruthful or deceptive information. Although SA has not been used in mediated communication contexts nor research, our team’s review and critique of the original method afforded a springboard on which to adapt select strategies to begin developing MSA. Through the adoption of fully applicable SA techniques, coupled with modified strategies that are specific and relevant to digital environments and mediated communication venues, such as social media, we developed a total of nine (9) MSA categories to use when evaluating social media posts to detect truthful-untruthful content [8].

MSA METHOD & PROCEDURES

METHODOLOGY

Mediated Statement Analysis identifies categories, combinations of categories, and patterns of potential deception in social media textual posts. MSA consists of nine (9) uniquely defined coding categories, including: (1) Word count, (2) Detail & Description, (3) Direct Numerical References, (4) Emotional & Affective Language, (5) +/- paradigm/pro-social & anti-social language, (6) Excessive emphasis, (7) Use of Absolutes, Hedges, & Qualifiers, (8) Exaggerated claims, and (9) combination of categories. For the current project, trained coders have labeled (identified) statement pairs that demonstrate three (3) key elements of MSA, specifically word count (length of statement), level of detail & description, and combination of these two (2) categories. The definitions for these coding categories used to analyze data cross the MSA pilot study and the COVID-19 data subset are presented below in Figure 1:

Figure 1: Description of MSA Coding Categories for Pilot Data Analysis

Category	Description
Word Count	Length of truthful or untruthful statement based upon count of words
Detail & Description	Messages containing details and descriptions in the form of explanations, elaboration, personalization, specificity, etc. (i.e., names of individuals & specific locations), direct references to people & places).
Combination of Cues	Indicates posts that fall into both of these coding categories; represents a cluster of potentially deceptive cues.

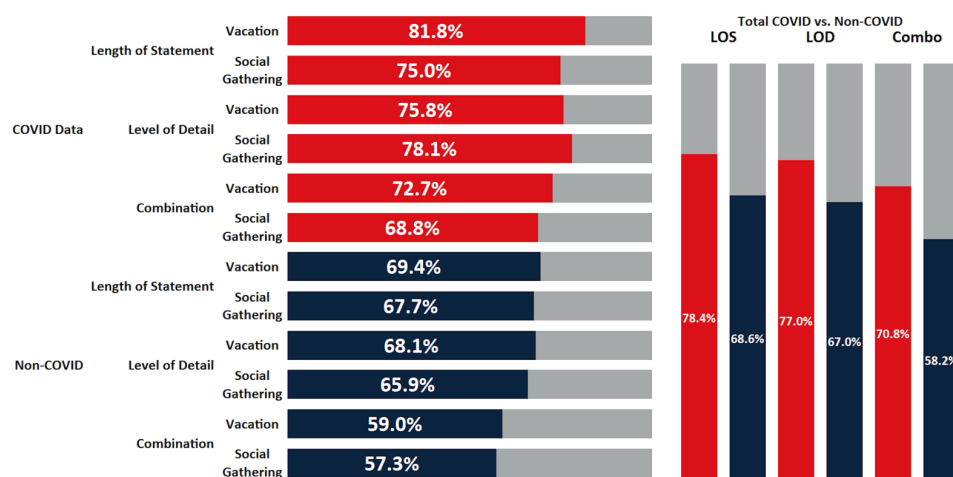
PROCEDURES

First, pilot data was gathered through an IRB approved (IRB#1654153-1) Qualtrics survey from 245 participants. The survey yielded 1,414 truthful-untruthful social media across three individual scenarios: (1) a recent vacation, (2) a recent social gathering, and (3) the last time they procrastinated. Second, a mixed methods approach was used to analyze data to test our application of the noted initial three (3) MSA coding categories. Qualitative data analysis uncovered the subset of 130 truthful-untruthful social media posts (65 pairs of posts) that related to COVID-19, mostly surrounding the survey topics of vacation and social gathering; only two (2) procrastination posts were related to the pandemic.

RESULTS

In our pilot data analysis, we applied MSA to 245 participants' and 1,414 truthful-untruthful social media posts for analysis. Overall findings reveal differences between truthful-untruthful posts in that truthful posts contain (1) significantly more descriptive and detailed content, (2) contain significantly more words than untruthful posts, and (3) contain combinations of longer word count and detailed/descriptive content together in posts. For this current analysis, we sought to apply MSA to the COVID-19 subset of pilot data and compare the findings, both to uncover potential insights about digital deception during times of pandemic crisis and to continue to validate the MSA methodology as a means to detect untruthful or suspicious textual content in the form of social media posts online. The results summary is presented below in Figure 2:

FIGURE 2: COVID-19 SUBSET V. PILOT DATA MSA RESULTS SUMMARY



Analysis of the COVID data subset exclusively identified that word count of truthful social media posts was longer than in untruthful social media posts in 51 out of 65 (78.4%), 27 out of 33 (81.8%) of posts related to vacation and 24 out of 32 (75%) related to social gatherings. Going forward, we refer to the word count category findings as LOS, or length of statement. Regarding level of detail and description, 50 out of 65 (76.9%) cases of truthful posts are more detailed and descriptive than their untruthful counterparts, with 25 out of 33 (75.75%) of cases related to vacation and 25 out of 32 (78.1%) about social gathering, or LOD in future reference. Truthful-untruthful posts that contained a combination, or COMBO, where both LOS and LOD categories appeared together in 46 out of 65 (70.8%) cases, with 24 out of 33 (72.7%) about vacation and 22 out of 32 (68.75%) about social gathering.

Comparative analysis with the NON-COVID pilot data findings revealed that LOS was a factor in truthful more than untruthful social media posts in 259 out of 387 (66.9%) (vacation: 134 of 199 (67.3%); social gathering: 125 of 188 (66.5%)). LOD was identified in 253 out of 387 (65.4%) (vacation: 133 of 199 (66.8%); social gathering: 120 of 188 (63.8%)) truthful versus untruthful social media posts. Across the pilot data set (less pairs of posts gathered for the procrastination survey item, COMBO of MSA categories was present in 217 out of 387 (56.1%) (vacation: 113 of 199 (56.8%); social gathering: 104 of 188 (55.3%) of truthful-untruthful social media posts.

DISCUSSION

KEY FINDINGS

Findings from these preliminary MSA categories (LOS, LOD, COMBO) reveal that MSA trends uphold in crisis-related (COVID) social media posts when compared with overall pilot data. According to the data, the MSA categories analyzed in COVID data subset appear more frequently than across the pilot data set. Additional LOS (Word Count) data reveals that both truthful & untruthful COVID-related social media posts are longer than their non-COVID counterparts: LOS means & SD differences in the COVID data subset are closer than compared with the overall pilot data. Perhaps individuals have more to express or share with their online social networks during a crisis or include more information in

crisis-related social media posts. The slightly larger percentage of MSA patterns in the COVID data subset, in addition to the pilot data analysis findings, suggest that MSA may be a reliable method to examine textual messages and posts online related to crisis narratives & information. This is worthy of additional exploration in various crisis contexts (i.e. during natural disasters or other health-crises).

NEXT STEPS FOR MSA RESEARCH

Our research team continues to explore MSA across multiple data sets. We are currently coding and analyzing the remaining six (6) MSA categories across the full data set of 1,414 truthful-untruthful social media posts. Additionally, at present we are examining differences/similarities of MSA patterns across Internet user demographics among truthful-untruthful social media posts in the pilot data. Further, we have already gathered a second, larger data set comprising 4,110 truthful-untruthful social media posts that we will begin coding and analyzing using our MSA methodology. Our upcoming phase of this research will explore the application of MSA via AI using machine-learning (ML) models to this larger data set. As we continue to explore and validate MSA as an effective method to examine textual online posts and messages, we look forward to future explorations surrounding crisis narratives and misinformation online. This multi-phase, multi-year research agenda investigates applying our novel method of MSA longitudinally and using innovative methods such as AI to continuously evaluate deceptive patterns and behaviors in textual content online.

CONCLUSION

The key objectives for this research as presented to academic and professional colleagues at the 2025 ICRC conference include to (1) provide an overview of utilizing MSA for digital deception detection in crisis communication contexts, (2) share our initial findings from the pandemic pilot data subset, (3) reveal observations and highlights related to truthful-untruthful social media posts about COVID-19 and compare these results with those of our MSA pilot study, and (4) update about our larger in-progress, multi-phase investigation using MSA to analyze truthful-untruthful social media posts. This research contributes to crisis communication research methodologies and practices when analyzing online textual content for digital deception detection in online mediated venues, such as social media platforms. Practical implications include the potential for MSA to be developed into an artificial intelligence (AI) algorithm and/or digital application powered by machine-learning models that may foster detection deception for online users in the future.

Author Biographies

Margaret C. Stewart, Ph.D., Indiana University of Pennsylvania (IUP). Associate Professor of Communication Studies, School of Communication, University of North Florida (UNF). m.c.stewart@unf.edu.

Christa L. Arnold, Ph.D., University of Florida. Associate Professor of Communication Studies, School of Communication, University of North Florida (UNF). christa.arnold@unf.edu.

F. Dan Richard, Ph.D., Texas Christian University. Professor of Psychology, University of North Florida (UNF). drichard@unf.edu.

REFERENCES

- [1] Al-Rodhan, N. (2017). Post-truth politics, the fifth estate and the securitization of fake news. *Global Policy*. Retrieved from <https://www.globalpolicyjournal.com/blog/07/06/2017/post-truth-politics-fifth-estateand-securitization-fake-news>
- [2] Effron, D.A. & Raj, M. (2019). Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share. *Psychological Science*, 31(1). <https://doi.org/10.1177/09567976198878>
- [3] Tsikerdekis, M., & Zeadally, S. (2014). Online deception in social media. *Communications of the ACM*, 57, 72-80. <https://dl.acm.org/doi/10.1145/2629612>
- [4] Petrescu, M., Ajjan, H., & Harrison, D. L. (2023). Man vs machine – Detecting deception in online reviews. *Journal of Business Research*, 154. <https://doi.org/10.1016/j.jbusres.2022.113346>
- [5] Roberts, N., & Qahri-Saremi, H. (2023). Tragedy, truth, and technology: the 3t theory of social media-driven misinformation. *Journal of the Association for Information Systems*, 24(5), 1358-1377. <https://aisel.aisnet.org/jais/vol24/iss5/3>
- [6] Petrosyan, A. (2023) Worldwide Digital Population 2023. Retrieved: <https://www.statista.com/statistics/617136/digital-population-worldwide/>
- [7] Sprout Social (2024). 50+ Must-know social media marketing statistics for 2024. Retrieved: <https://sproutsocial.com/insights/social-media-statistics/>
- [8] Stewart, M.C. & Arnold, C. L. (2022). Bridging the gap for online deception detection: Uncovering methodology to identify deceptive content in mediated communication. *Proceedings of the International Crisis and Risk Communication Conference*, Volume 5 (pp. 42-44). Orlando FL: Nicholson School of Communication and Media. <https://doi.org/10.30658/icrcc.2022>.