

Title: Methodology details, participant recruitment, and data analysis appendix for “A playbook for mapping adolescent interactions with misinformation to perceptions of online harm”

Authors: Gowri S. Swamy (1), Morgan G. Ames (1), Niloufar Salehi (1)

Date: October 30th, 2024

Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

## **Appendix A: Methodology details, participant recruitment, and data analysis**

### *Semi-structured qualitative interviews*

We first conducted 15-minute semi-structured qualitative interviews with participants. Our questions covered two themes: the first focused on how participants established trust or mistrust in their digital interactions, and the second investigated their thoughts and perceptions of misinformation, perceptions of harm, and accountability. Questions included how participants identify misinformation in their online interactions, how their lack of trust arises, who they perceive may encounter harm and or may be accountable, what moderation practices they've encountered, and how their exposure to different media sources influences their perceptions.

### *Semi-structured qualitative interviews*

We first conducted 15-minute semi-structured qualitative interviews with participants. Our questions covered two themes: The first focused on how participants established trust or mistrust in their digital interactions and the second investigated their thoughts and perceptions of misinformation, perceptions of harm, and accountability. Questions included how participants identify misinformation in their online interactions, how their lack of trust arises, who they perceive may encounter harm and/or may be accountable, what moderation practices they've encountered, and how their exposure to different media sources influences their perceptions.

### *Guided artifact retrieval*

Next, we asked participants to navigate through their search history (with the option to select from their mobile phone, tablet, or laptop) to display pieces of misinformation they encountered in the last 30 days. After giving participants 2–3 minutes to navigate through search history, we asked them to explain why the found source was identified as misinformation, how they navigated through the content, and accountability/content moderation practices they engaged with when they encountered the information. We then asked participants to further explain their interpretation of the media policies on the service that influenced their perceptions of trust and harm. This exploration took 15–20 minutes in total.

### *Misinformation newsfeed interaction*

In our third exploration, which took 10–15 minutes, we presented participants with a pre-generated newsfeed of information that included different types of misinformation interspersed with truthful

sources created in Qualtrics. The examples of misinformation we used had been recently trending on various online platforms and spanned multiple types of misinformation. All pieces of misinformation references were assessed based on prevalence on social media platforms and inclusion of certain characteristics when mapping to a specific type of misinformation. This part of the study explored participants' information consumption practices and navigation strategies. We showed participants two to three pieces of each type of misinformation described in Table 2 and Table 3 and two pieces of authentic content for a total of 15 newsfeed interactions (see Figure 1). After each, we asked participants to rank how likely it is that the media they are seeing is not truthfully representing information (on a 5-point scale from 1 = very unlikely to 5 = very likely) and to write down their reasoning in a text field. We also asked them to verbalize their thought processes using "think-aloud" prompting. At the end of this exploration, participants were asked to rank the media they had encountered by the level of harm it caused.

### *Situational mapping*

In the final part of the study, participants were informed that they had interacted with pieces of misinformation (i.e., content that has now been verified as factually incorrect) and they then engaged in a situational mapping activity where they were asked to identify those who were harmed and those who were or should be accountable for misinformation on a collaborative workspace called Mural. First, we presented participants with a chart naming and describing the six different types of misinformation and a thumbnail example. We then guided participants through a large interactive table, and we continued to prompt participants to think aloud as they filled it out. For each type of misinformation, we asked participants to describe:

- 1) whether or not they thought that kind of misinformation caused harm (yes, no, or maybe);
- 2) who they thought the recipients of harm were (from a list of kids/pre-teens, teenagers, politics/political parties, educators, government, differently abled, parents, and other);
- 3) who they thought was accountable for that harm or for preventing future harm (from a list of policymakers, government [with the option to list a specific branch or agency], media other than journalists, journalists, educators, public, technology companies, or other);
- 4) what level of accountability they felt that each stakeholder they listed had (neutral, somewhat accountable, or very accountable); and
- 5) what could be done to mitigate or decrease the risk of harm, based on their interactions with digital media as well as the information they gained through these activities.

### *Recruiting*

Participants were compensated hourly, per the policies set by the campus service and our IRB protocol. Our recruiting message stated that we were looking for participants who use a variety of media channels on a regular basis and are willing to engage in an activity that explores their recent interactions with misinformation.

### *Data analysis*

In the first round of coding, we identified patterns within each participant's interview. In the second round, we identified broader themes across participant interviews and wrote memos on connections to related literature. We then generated descriptive statistics on the data from newsfeed interactions and created affinity mapping diagrams to discover patterns in the participant assessments for reasons behind their trust assignments. Finally, we explored participants' assignments of levels of accountability and harm from the situational mapping exercise, as well as their reasoning, which helped us develop ideas to identify and mitigate harms spread by misinformation.

### *Limitations*

There are a few limitations in our study which are important to note. First, for the purpose of this study, while we worked with participants of ages 18 and 19, we broadly categorized this group as adolescents. Precedence established by prior studies (Laplante et al., 2021; Salac et al., 2023; Xiao et al., 2022) has demonstrated the use of data from similar age groups, and this study aspires to understand information sensibility practices of adolescents by interfacing with 18- and 19-year-old users. However, working with a limited age group poses risks of not being able to authoritatively map to all adolescent users. We also recruited our sample from one university and note that this may limit the overall generalizability of our findings.

Second, while the design of the study intended to best use participant time and thought processes to understand their past experiences with misinformation and notions of content online, it may be the case that the ordering of the tasks may have led participants to think differently or more critically about misinformation than they otherwise would. Sessions were also conducted in a lab setting and, therefore, participant perceptions may change as they were not in their natural environment. In future work, we hope to structure our sessions with participants in such a way that we are able to limit potential biases.

Finally, while the proposed frameworks are based on findings from our study and existing literature, the scope of the paper did not involve the evaluation and testing of the ideas presented. Testing of the playbook will allow for an increasingly detailed and nuanced presentation of the potential uses of the frameworks suggested.

### *Guided artifact retrieval*

Next, we asked participants to navigate through their search history (with the option to select from their mobile phone, tablet, or laptop) to display pieces of misinformation they encountered in the last 30 days. After giving participants 2–3 minutes to navigate through search history, we asked them to explain why the found source was identified as misinformation, how they navigated through the content, and accountability and content moderation practices they engaged with when they encountered the information. We then asked participants to further explain their interpretation of the media policies on the service that influenced their perceptions of trust and harm. This exploration took 15–20 minutes in total.

### *Misinformation newsfeed interaction*

In our third exploration, which took 10–15 minutes, we presented participants with a pre-generated newsfeed of information that included different types of misinformation interspersed with truthful sources, created in Qualtrics. The examples of misinformation we used had been recently trending on various online platforms and spanned multiple types of misinformation. All pieces of misinformation references were assessed based on prevalence on social media platforms and inclusion of certain characteristics when mapping to a specific type of misinformation. This part of the study explored participants' information consumption practices and navigation strategies. We showed participants two to three pieces of each type of misinformation described in Table 2 and Table 3 and two pieces of authentic content for a total of 15 newsfeed interactions (see Figure 1). After each, we asked participants to rank how likely it is that the media they are seeing is not truthfully representing information (on a 5-point scale from 1 = "Very Unlikely" to 5 = "Very Likely") and to write down their reasoning in a text field. We also asked them to verbalize their thought processes using "think-aloud" prompting. At the end of this exploration, participants were asked to rank the media they had encountered by the level of harm it caused.

### *Situational mapping*

In the final part of the study, participants were informed that they had interacted with pieces of misinformation (i.e., content that has now been verified as factually incorrect) and they then engaged in a situational mapping activity where they were asked to identify those who were harmed and those who were or should be accountable for misinformation on a collaborative workspace called Mural. First, we presented participants with a chart naming and describing the six different types of misinformation and a thumbnail example. We then guided participants through a large interactive table, and we continued to prompt participants to think aloud as they filled it out. For each type of misinformation, we asked participants to describe:

- 1) whether or not they thought that kind of misinformation caused harm (yes, no, or maybe);
- 2) who they thought the recipients of harm were (from a list of kids/pre-teens, teenagers, politics/political parties, educators, government, differently-abled, parents, and other);
- 3) who they thought was accountable for that harm or for preventing future harm (from a list of policymakers, government [with the option to list a specific branch or agency], media other than journalists, journalists, educators, public, technology companies, or other);
- 4) what level of accountability they felt that each stakeholder they listed had (neutral, somewhat accountable, or very accountable); and
- 5) what could be done to mitigate or decrease the risk of harm, based on their interactions with digital media as well as the information they gained through these activities.

### *Recruiting*

Participants were compensated hourly, per the policies set by the campus service and our IRB protocol. Our recruiting message stated that we were looking for participants who use a variety of media channels on a regular basis and are willing to engage in an activity that explores their recent interactions with misinformation.

### *Data analysis*

In the first round of coding, we identified patterns within each participant's interview. In the second round, we identified broader themes across participant interviews and wrote memos on connections to related literature. We then generated descriptive statistics on the data from newsfeed interactions and created affinity mapping diagrams to discover patterns in the participant assessments for reasons behind their trust assignments. Finally, we explored participants' assignments of levels of accountability and harm from the situational mapping exercise, as well as their reasoning, which helped us develop ideas to identify and mitigate harms spread by misinformation.

### *Limitations*

There are a few limitations in our study which are important to note. First, for the purpose of this study, while we worked with participants of ages 18 and 19, we broadly categorized this group as adolescents. Precedence established by prior studies (Laplante et al., 2021; Salac et al., 2023; Xiao et al., 2022) has demonstrated the use of data from similar age groups, and this study aspires to understand information sensibility practices of adolescents by interfacing with 18- and 19-year-old users. However, working with a limited age group poses risks of not being able to authoritatively map to all adolescent users. We also recruited our sample from one university and note that this may limit the overall generalizability of our findings.

Second, while the design of the study intended to best use participant time and thought processes to understand their past experiences with misinformation and notions of content online, it may be the case that the ordering of the tasks may have led participants to think differently or more critically about misinformation than they otherwise would. Sessions were also conducted in a lab setting and, therefore, participant perceptions may change as they were not in their natural environment. In future work, we hope to structure our sessions with participants such that we are able to limit potential biases.

Finally, while the proposed frameworks are based on findings from our study and existing literature, the scope of the paper did not involve the evaluation and testing of the ideas presented. Testing of the playbook will allow for an increasingly detailed and nuanced presentation in the potential uses of the frameworks suggested.