SLOCIAL slow social media

Anshuman Dhar, Ed Paradis, Leo Salemann, & Madison Stemmler

Dec. 15, 2021

Introduction

Abstract

We are interested in the interaction design of social media and the emotional and psychological effects it has on users. Our ideas are based on the concept of "less is more" - instead of adding something to the social media experience, what could we take away? Examples include temporarily hiding the platform's "reshare button" (Twitter's Retweet, Facebook's Share, etc.) or removing certain content from the user's under certain conditions.

Modern social media is built for speed. It prioritizes first impressions, reactionary reshares, and emotional comments over deliberate consideration, lateral reading, and mindful interactions with others. While this emphasis on fast interactions increases user engagement (and thus exposure to advertising), it may also contribute to the propagation of mis-, mal-, and disinformation across the platform and to society at large.

In the Background section we begin with an exploration of "slowness" from the slow design movement, to mindfulness, to the potential for "speed brakes" to prevent the re-share of disinformation without thoughtful engagement. We move on to examine current interventions in social media platforms and techniques for interrupting "doom scroll" behaviors. We finish the section with a discussion of paternalism in social media design.

During Ideation, we map out ways people engage with social media and assess the disinformation propagation risk of each. We consider resharing content as high risk because it directly affects how far a piece of disinformation can propagate, while reading content is considered low risk because the activity in and of itself does not affect anyone but the reader. After brainstorming dozens of interventions for each engagement activity, we chose to focus on a *scrolling* intervention because it can lead to *reshares* of and

reactions to disinformation. We believe delaying an intervention until these higher-risk activities would be too late.

Our scrolling intervention prototype is patterned after Twitter, with an intervention that desaturates the colors of the interface to grayscale over time. Our hypotheses are that this intervention will (1) slow down the scrolling experience, (2) encourage more analytical thinking, and (3) reduce the spread of disinformation.

We developed a between-subjects experiment in which the treatment group of subjects received the intervention while the control group received the same prototype, showing the same content, but without the fade-to-grayscale intervention. A pilot study of 8 subjects (4 in each group) validated that the prototype is usable and data can be collected to support quantitative and qualitative analysis.

Further experiment runs with a larger population controlled for basic demographics and social media proficiency would be required to test hypotheses 1 and 2 in a statistically significant manner. Hypothesis 3 would require a new experimental design to measure actual propagation of disinformation rather than the accuracy of a proxy task.

Background

Slow Design

Slow design focuses on the well-being of individuals and society by considering the environment that users exist in when interacting with interfaces and tangible objects. "It encourages people [to do] things at the right time and with the right speed and helps them to understand and reflect on their actions" (Grosse-Hering et al., 2013). While the case studiesBarbara Grosse-Hering looks at are not directly related to misinformation, they provide lessons which can prove to be very helpful to social media platform design. The most helpful is that "Slow Design principles can be used to create more 'mindful'

interactions that stimulate positive user involvement" (Grosse-Hering et al., 2013). They go on to say these principles don't mean designing interactions which take more time, but simply allowing the user to focus more on the parts which are meaningful; less on the parts which are not. This is especially relevant for social media companies which claim to help users engage more "meaningfully," but most users aren't happy with the way they consume content and feel like they're losing time (Saindon, 2021).

Slow & Mindful

In Information Disorder, Warle and Derakshan identify nudges as a primary technological solution to the proliferation of "polluted" (i.e. mis- dis- and mal-) information on social media platforms. (Council of Europe et al., 2017, pp. 59–60). Nudges are combinations of human and algorithmic persuasion (Matias, 2017) intended to encourage people to change their behavior while still preserving their liberty. While Warle and Derkshan describe the primary benefit of such nudges as "slowing down the spread of polluted information," most of them do so by giving the user more things to do, such as commenting or down-voting (Bilton et al., 2017), rather than passive contemplation.

In their post mortem *Stop the Steal and Patriot Party: The Growth and Mitigation of an Adversarial Harmful Movement*, Facebook cited caps, rate-limiters, and other "slow down" techniques to control the runaway growth and posting-rate identified as a root causes that led to the spread and incitement of the Stop the Steal movement on the platform (Mac, 2021).

Schwarz and Jalbert (2021) along with Oyserman & Dawson warn that "when thoughts flow smoothly, people are likely to agree [with mis-, dis-, mal- information] without much critical analysis." They go on to cite multiple studies in which color contrast (Reber and Schwarz, 1999), font choice (Song and Schwarz, 2008), and audio/video quality (Newman & Schwarz, 2018) are deployed to disrupt the smooth flow of thoughts. These disruptions, which can be thought of as user friction, are found to lead to more critical analysis among the users tested. While user friction is normally thought of as an anti-pattern in UX design (Krug, 2000), we assert it can also be deployed constructively as a "speed brake" to encourage

more mindful social media consumption as advocated by some social media practitioners (Tatiana, 2018), the mindfulness movement (Mindful Staff, 2015; Willard, 2018), and the slow movement (Köhler et al., 2019).

Current interventions by social media platforms

Most interventions currently deployed by social media platforms can be classified as "soft" such as warning labels, or "hard" such as removing a user's content or suspending their account. Looking at the warnings used by Twitter during the election, studies found a backfire effect in which "Trump's tweets that were marked as containing misinformation spread further than the tweets that received no intervention from Twitter" (Swann, S., 2021). The same backfire effect has been observed for hard interventions where "tweets that were removed from the platform spiked in engagement on other social media outlets, namely Facebook, Instagram and Reddit" (Swann, S., 2021).

According to new research (Kaiser et al., 2021), interstitial warnings are better than fact-checks or contextual warnings which show up beside or below the content. The authors go on to claim that using interstitial warnings resulted in far better results where users "noticed the warnings, considered the warnings, and then either declined to read the flagged content or sought out alternative information to verify it."

Alternatively, a study from the School of Computing at the National University of Singapore (von der Weth et al., 2020) concluded that, as opposed to increasing skepticism among users consuming social media content, targeting the propagators of misinformation by encouraging more conscious posting was effective at slowing the spread of misinformation. The researchers utilized nudging, explained by the authors as "a form of soft paternalism to guide users by suggesting, instead of enforcing a certain behavior" (von der Weth et al., 2020). Through an experimental browser plugin dubbed ShareAware, users' posts were analyzed for factuality and credibility. If concerns were raised, the user was notified and "nudged" to review their post more carefully and intentionally. While the researchers found this method

can significantly reduce the spread of misinformation, it is not without its limitations—particularly that ShareAware relies on a semantic understanding of post content, technology that will have to be considerably improved prior to being widely implemented outside research environments.

It's worth noting that creating technology to detect and reduce the spread of information has the potential to encourage ever more sophisticated forms of mis- and disinformation. As the authors stated in the discussion on ShareAware, "these tasks are inherently very challenging and can be viewed as a cat-and-mouse game – any improvements in detection will result in the development of improved bots or better fake content" (von der Weth et al., 2020). Yet technologies like ShareAware show preliminary promising results to begin chipping away at the problem.

Interrupting the Doom Scroll

In *On Reading and Being Read in the Pandemic: Software, Interface, and The Endless Doomscroller*, Grosser describes how the combination of algorithmic social media feeds, global anxiety, and the gaps in the flow of time that come with the sudden transition living and working at home form the perfect conditions for excessive consumption of bad news on social media sites. This often leaves users feeling stuck in the endless scroll of their social media feeds, "unable to disrupt [their] own behavior". The activity became so pervasive that it has gained a new term, "doomscrolling" (Grosser, 2020). While the effects of doomscrolling have been well-documented, it is interesting to consider when interrupting the doomscroll can be against the user's best interests. In the same paper the author notes that, for many, mindlessly scrolling was "'meditative' and a 'distraction from uncertainty' in the face of 'existential anxiety'" (Grosser, 2020). In such cases, there exists a conflict between the intervention designer's intent and the user's reality.

Studies have investigated if adding intentional friction into a mobile application has any effect on users' level of satisfaction (Mejtoft et al., 2019). Although the authors mainly look at an onboarding experience for a meditation app as opposed to the scrolling experience of a social media platform, it's interesting that

many of the participants actually preferred to have intentional friction in their experience. They felt "more satisfied when they had a clear understanding of the goal of each [onboarding] task" (Mejtoft et al., 2019). In other words, the onboarding experience has benefited from a constructive deployment of user friction as described in our Slow and Mindful subsection.

Paternalism in design and media

Intervening into the experience of a user for what the designer perceives as the user's best interest runs the risk of inappropriate paternalism. Schell calls for "personal responsibility" to be taken by designers for their creations (Schell 2008, p456). Norman says designers have a "moral obligation", and points out that "design … takes on a political significance" - a particularly compelling idea when the design at hand is for a social media platform (Norman 2013, p. 291). Soft paternalism interventions such as nudges have been shown as a promising method to discourage unintended disclosures on Facebook (Wang et al., 2014). However, Thaler and Sunstein point out that paternalism does not require coercion and suggest, similarly to Norman, that "organizational decisions are inevitable … and a form of paternalism cannot be avoided" (Acquisti, 2009). Given the tendency for social media's intrinsic properties to encourage the belief of falsehoods (Schwarz & Jalbert, 2020), Schell's call for responsibility can be interpreted as a call for the inclusion of explicit design elements to discourage the belief of falsehoods.

The design community has identified ways to encourage voluntary compliance with desired behavioral norms in an online community (Kraut et al., 2011, pp. 140–166). Our area of interest seeks to identify an analogous set of design concepts that encourage voluntary forms of personal engagement with social media that lead to better psychological and emotional health. The Slow Design (Grosse-Hering et al., 2013) and Mindfulness (Mindful Staff, 2015) communities suggest that users might accept some amount of friction in their social media experiences. Research into disinformation on social media suggests that such friction can encourage critical thinking and reduce the spread of such disinformation (Schwarz and Jalbert, 2021). For these reasons, we have adopted a design strategy to slow the spread of disinformation

on social media based on the deployment of constructive user friction while preserving user agency and avoiding designer paternalism.

Problem Statement

Our problem is how to reduce the spread of disinformation on social media platforms. Modern social media is designed to maximize user engagement, which can be measured several ways: the amount of time spent on the platform, the number of user interactions per unit time, and/or the amount of content (and thus advertising) observed by a user in a particular demographic. All of these metrics are improved by encouraging short, speedy interactions. Relying on first impressions, reactionary reshares, and writing emotional comments take less time than deliberate consideration, lateral reading, and mindful interactions with others. In short, social media is built for speed, prioritizing fast intuitive "system 1 thinking" over slower deliberative "system 2 thinking."

The problem with the emphasis on system 1 thinking is that "People are likely to agree [with mis-, dis-, mal- information] without much critical analysis" (Schwarz & Halbert, 2020). Disinformation can spread quickly in environments that encourage little contemplation or careful thinking. Instead of each point of reshare being a possibility to correct or filter, it is a broadcast to a user's followers.

Our goal is to disrupt this system of dissemination by introducing design interventions that promote system 2 thinking and thus slow the spread of disinformation and misinformation. These interventions are not, however, based on the content being shared. They apply equally across all users and content such that they are robust against changes in tactics of disinformers and whichever current events arise for manipulation of the public's attitudes. Thus we aim to avoid many of the biases and difficulties of content moderation techniques.

Design Preview

Solution Space

Our solution centers on a platform designed to encourage more mindful consumption and propagation of social media content. We believe this can be achieved by slowing down the consumption experience to promote more thoughtful (i.e. system 2) engagement with content and more intentional sharing.

We focus on the consumption and resharing experiences. Resharing is a point in the social media user experience where an individual must make a decision about taking a specific action. Therefore we expect a clear metric of impact in the number of times a piece of disinformation is shared or not. Additionally, we measure the time spent completing a given task for possible further insights. Ideally, our intervention should guide users to consider the veracity of a post and its effect on their communities.

By increasing the time to share a post, the viral spread of information will decrease. Therefore in cases where disinformation is reshared, the overall speed of its distribution will be slower. This will benefit the platform by decreasing the propagation of disinformation.

Stakeholders

As with any media platform, stakeholders include content creators, content consumers, advertisers, third party developers, and the social media company itself. We realize slowness will be a hard sell in a metrics-driven industry that values rapid growth, rapid reshares, and maximized engagement. Considering the extent to which opinions and ideas shared through social media influence peoples' off-line behavior, we include non-users, politicians, and society at large as extended stakeholders. We hope that through encouraging more mindful social media engagement, both direct and extended stakeholders will have a more thoughtful, less divisive, and more peaceful society.

We offer the following value propositions:

- **For Users**: A social media experience that leaves them feeling more connected and less stressed.

- **For Content Creators**: An assurance that their work will be engaged with deeply.

- **For Advertisers**: A chance to build a more loyal customer base and confidence that users have interacted with them for their product offering, rather than an inattentive state of mind while scrolling.

- **For society at large**: A less polarized populace more resistant to disinformation.

Table 1 - Stakeholders

Stakeholder	Potential Negative Impacts	Potential Positive Impact	Notes
User who often re-shares	May feel constrained or censored	May feel more empowered	These stakeholders are the primary focus of our project
User who consumes but rarely re-shares	Less content, no instant gratification, may feel their feed is "filtered" or empty	Receives high-quality content	Less impact overall
Influencers (people who want to persuade)	Move to other platforms with less restrictions, taking their followers with them	Followers engage more deeply with their content, become more loyal	Includes politicians, entertainers, other public figures

White-hat content creator (posts truthful content)	Might slow the spread of important information, not getting to people on time.	Incentivised to produce higher-quality content which is consumed more thoughtfully	Includes journalists, activists
Black-hat content creator (posts disinformation , other malicious content)	Ignore platform interventions, refuse to use the platform/move to other social platforms	Slow or stop the spread of mis- and disinformation	Includes RU-IRA, misinformation agents. Negative impacts to black-hat creators are considered a net-positive to everyone else.
Advertisers	Reduction of "quantity" impressions	Increase in "quality" impressions	May necessitate change in business model
Society at large	"Censorship" or limited access to information	A less polarized populace more resistant to disinformation.	Societal-level impacts are hard to attribute to specific changes

Approach

Our design approach began with a whiteboard ideation session in which we mapped out different ways people interact with social media platforms and how each of these activities can contribute to the propagation of disinformation. From there, we brainstormed a series of interventions that could be applied to each interaction, voted a top three, then voted a single winner. The winning intervention was implemented as a Twitter-like experience in Figma, which was used to conduct a between-subjects pilot study. All participants in the study received a simulated Twitter feed with the same content, and a task to

find and reshare cat pictures. The treatment group also received the intervention, which de-saturated the colors of the feed over time until it was grayscale by the end. We recorded time to complete the task and the number of "cat Tweets" correctly identified.

Method

Method overview

Our design process consisted of the following elements:

- **Ideation** - Map out the ways people interact with social media. For each interaction, identify the risk that it could lead to disinformation propagation and brainstorm interventions to mitigate such risk.
- **Prototyping** - Select a single intervention and develop a medium to high fidelity prototype.
- **Experiment** - Using the prototype, design an experiment and perform a pilot study. Validate the usability of the prototype and the ability to collect and analyze qualitative and quantitative data.

During ideation, we brainstormed and sketched possible interventions with a theme of "slow social media". We then synthesized a framework of social media user interaction, focusing on the user flows which could lead to the spread of misinformation. With this framework, we identified specific interventions and placed them into our framework. We then presented our concepts for feedback.

Using this feedback we settled on a single intervention design - a gradual desaturation of colors while scrolling a social media feed. We then designed our prototype and an experimental process to gain feedback from users. With both the qualitative and quantitative data collected from user testing, we evaluated our hypothesis and formed preliminary design conclusions. These were then presented to class.

Our final steps were analyzing the results of our study and identifying possible improvements.

Design

Our primary concern is the spread of disinformation on social media platforms, which comes from:

- Posting new disinformation.

- Re-sharing other people's posts containing disinformation.

- Adding a positive reaction (liking, upvoting) posts containing disinformation.

We believe that by slowing down the entry points to these states, we can reduce the spread of disinformation.

Ideation

Social media engagement loop

Our first whiteboard session focused on the different ways people engage with social media, where we dentified 9 activities.
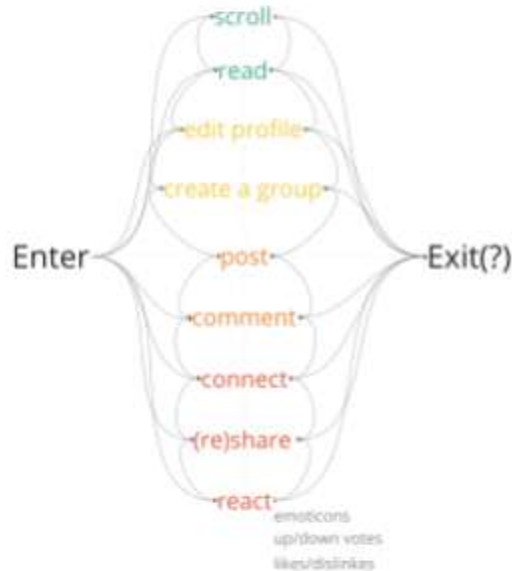
Figure 1 - Whiteboard the Social Media Engagement Loop

1. Scrolling - Looking at a list of content, often in compressed (e.g. headline) form.

2. Reading - Reading a specific piece of content in detail. "Opening" a headline to read the actual article. Also generalizes to viewing, watching, or listening for platforms that focus on audio-visual rather than textual content.

3. Posting - Creating a new piece of content to post to the platform.

4. Commenting - Writing comments on someone else's post.

5. Reacting - Similar to commenting but via icons or emojis (likes, upvotes, smile, frown, …).

6. Connecting - Following other users, accepting connection requests, joining groups.

7. Re-Sharing - Taking content seen on the platform, and forwarding it to other users (i.e. a user's followers).

8. Group formation - Starting a new interest group.

9. Profile Creation & Update - Creating or modifying your representation of yourself.

These activities are nonlinear in that people can start with any one of them when they first enter the platform and switch between them at will, forming a "Social Media Engagement Loop" in which users engage continuously with no clear exit point. For each activity we assigned a low (green), moderate (yellow), high (orange) or severe (red) degree of risk based on the likelihood of the activity directly contributing to the propagation of disinformation on the platform.

Social Media Engagement Loop

- scroll
- read
- edit profile
- create a group
- post
- comment
- connect
- (re)share
- react
  emoticons
  up/down votes
  likes/dislikes

Enter — Exit(?)

Reading & scrolling do not directly spread disinformation.

Reading & scrolling do not directly spread disinformation.

Group formation can be the first step in a disinformation campaign

Posting, commenting can spread disinformation, but takes time.

Posting, commenting can spread disinformation, but takes time.

Figure 2 - Final Social Media Engagement Loop

Journey mapping and brainstorming

With the activities identified, we "unrolled" the social media engagement loop into one of multiple possible linear journeys, in which the user:

1. *Enters* the platform

2. *Scrolls* through content headlines.

3. *Reads* a piece of content.

4. *Chooses* to "engage", meaning to *comment, reshare, or react.*

5. *Exits.*

Figure 3 - Whiteboarding the User Journey

For completeness, we also added *connect* and *write* before brainstorming invention ideas with sticky notes.
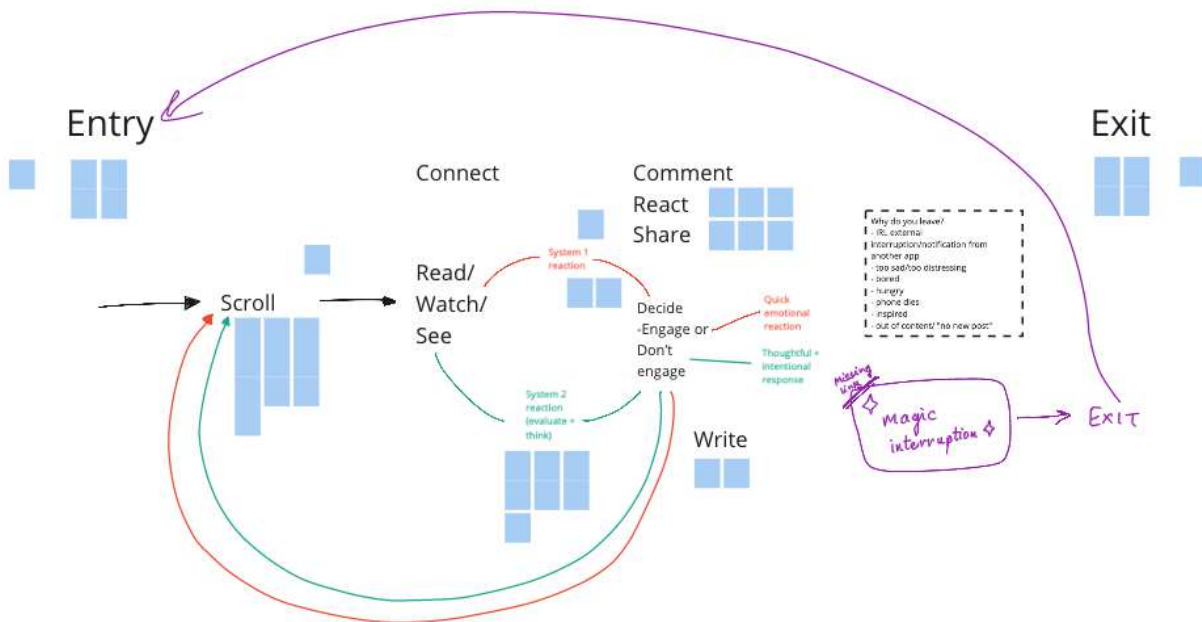

Figure 3 - Final User Journey

Collaborative voting yielded three candidate interventions:

- **Scrolling Interface (4 votes)** - Interventions which change the appearance and interaction design of the scrolling experience, such as introducing friction (slow scrolling, reduced text density), decreasing fidelity (desaturated colors), unexpected affordance changes (moving/resizing buttons).

- **Behavior-Based Limiting (3 votes)** - Interventions which restrict or deny a user the ability to perform possibly problematic actions based on behavior in the user's current session. Examples include daily or hourly thresholds of posts or shares; and/or frequency limits (time since last share, etc).

- **Mindful Re-Share (2 votes)** - Interventions which prompt users to enter a more mindful state before taking an action, composed of affordances to (1) allow people to set their intentions before engaging, (2) engage in sharing, and (3) reflect on the affect the reshares have caused.

Mapped to the user journey, these interventions land in low, to moderate, and high risk areas:
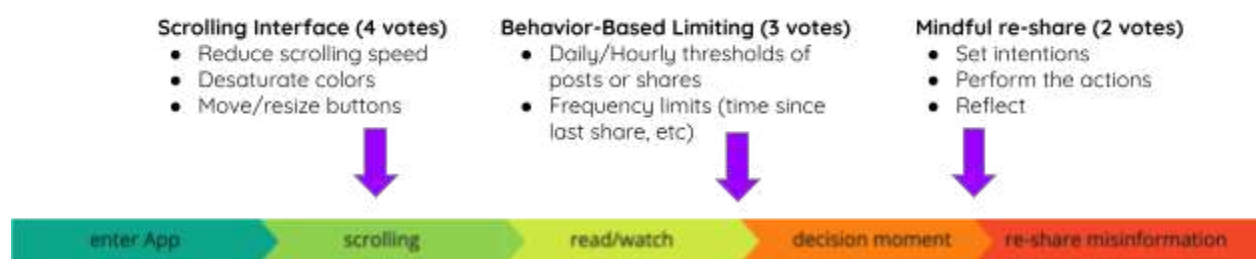

Figure 4 - Aligning Interventions to the User Journey

Down-Selecting a Single Intervention - Scrolling

Although scrolling is considered a low-risk activity, we chose to intervene here as a preemptive strategy. As users read content and choose to engage, it may be too late to slow the spread of disinformation if their

mind has already been made up to share. Intervening at the scrolling stage allows us to influence user behavior before the fast, intuitive system 1 thinking takes hold.

Prototyping

User Scenario

Our prototype is built around a *feature* (fading the content to gray over time) and a *scenario* (an activity that a participant will perform with the prototype as part of an experiment. We wanted to look at how many people view information in an analytical/critical nature before they share it, so we based our scenario on the "Noah's Ark"/"Moses' Ark" riddle used in similar experiments (Song & Schwarz, 2008) . We wanted to replicate the Noah/Moses puzzle in visual and textual form, so we looked for artifacts in the form of pictures, emojis, and text.

Finding Artifacts

For each medium we needed to render the same concept with "Noah style" artifacts that were clearly true as well as "Moses style" artifacts that were deceptive in some way. For the concept of "baby", Noah-style pictures would show a baby as the main figure of the image, while Moses-style would conceal the baby in some way, such as utilizing negative space.

Figure 5 - Baby Images detectable by System 1 thinking (left) and System 2 (right)

A Noah-style emoji would be an actual baby emoji such as 👶 while Moses-style would be a non-baby emoji which looks similar at first glance, such as 🐤. For text we consider the word "baby" as Noah-tyle, while Moses-style would be a slight misspelling such as bacy, beby, daby, etc.

In addition to the availability of Noah-style and Moses-style imagery, text, and emojis, artifacts for the experiment also had to be:

- Apolitical (eg. avoid Trump tweets or partisan issues).

- Something which did not evoke an emotional response (eg. anger, sorrow).

- Something which works across text as well as images.

- Something which would not feel out of place on people's social media feeds.

- Something which exists across different types of media that populate people's feeds (eg. photographs, graphics, videos).

- Something with lots of easy/hard identifiable images online.

In the end, we selected cats as the artifact for our main experiment with racoons (plus lemurs plus red pandas) for our practice run.

Example Tweets

A total of 210 tweets were developed, 44 of which depicted the concept of cat in some way.

Table 2 - Example Tweets

	Easy/Noah	Hard/Moses
image		
emoji		

			(emoji is a fox rather than a cat)
text	thoughts of a cat @feline_feelings ·30s dear. humam. sleeping in. is not. allowed. you must wake up. whenever. i decide. i want food. regardless. of how. early. that may be. 💬 2 🔁 8 ♡ 107 ⬆	

Figma Prototype

The tweets were assembled into a single feed in Figma and embedded into a simulated iPhone. Users could scroll up and down through the feed in a web browser - either on desktop or mobile (both simulated an iPhone). Three feeds were created -- a full-color racoon feed (Practice), and a cat feed that came in full-color (Control) and fade-to-gray (Treatment)

Figure 6 - Prototype Examples

The tweets in the Control and treatment group were identical, except for the color desaturation.

Figure 6 - Prototype with Scroll Desaturation

Experiment & Results

Methods

We conducted user interviews over Zoom and used the built-in recording feature to aid in post-interview analysis. After an introductory statement, participants were introduced to a simplified version of our prototype and asked to perform a task with it in order to set a baseline for familiarity bias. Then participants were presented with one of two versions of our prototype: a control with no design changes or a version with our color desaturating intervention. They were asked to scroll through our prototype, find tweets of a certain type, and reshare them. Upon reaching the end of the prototype, we asked questions to gather qualitative information on their experience and thoughts. Our study script can be found in Appendix A.

Quantitative Results

Participants 1-4 were presented with the control prototype: all posts in full color.

Table 3 - Quantitative Results, Control Group

	Participant 1	Participant 2	Participant 3	Participant 4	Average
Target tweets found	27	32	40	27	31.5
Target tweets not found	17	12	4	17	12.5
Group	Control	Control	Control	Control	Control
Accuracy (Found/Total)	61%	73%	91%	61%	72%
Time to Complete (mm:ss)	14:14	17:11	19:58	11:34	15:44

Participants 5-8 were presented with the experimental treatment prototype: posts gradually desaturate to grayscale.

Table 4 - Quantitative Results, Treatment Group

	Participant 5	Participant 6	Participant 7	Participant 8	Average
Target tweets found	34	26	32	37	32.3
Target tweets not found	10	18	12	7	12.8
Group	Treatment	Treatment	Treatment	Treatment	Treatment
Accuracy (Found/Total)	77%	59%	73%	84%	73%
Time to Complete (mm:ss)	15:10	11:54	9:23	8:43	11:17

Statistically, the results are nearly the same. Some of the variation in accuracy and completion time could be attributed to variations in participants' familiarity with social media in general and Twitter in particular. Testing more subjects and controlling for these variables would let us identify and remove outliers and arrive at more statistically significant results.

A chart of successful content identification over the tested sequence which compares the control and experimental and treatment groups allows visualization of patterns related to the number of tweets seen. Separating the target tweets by their difficulty shows that the control group more commonly identified easy tweets than the group with color desaturation.
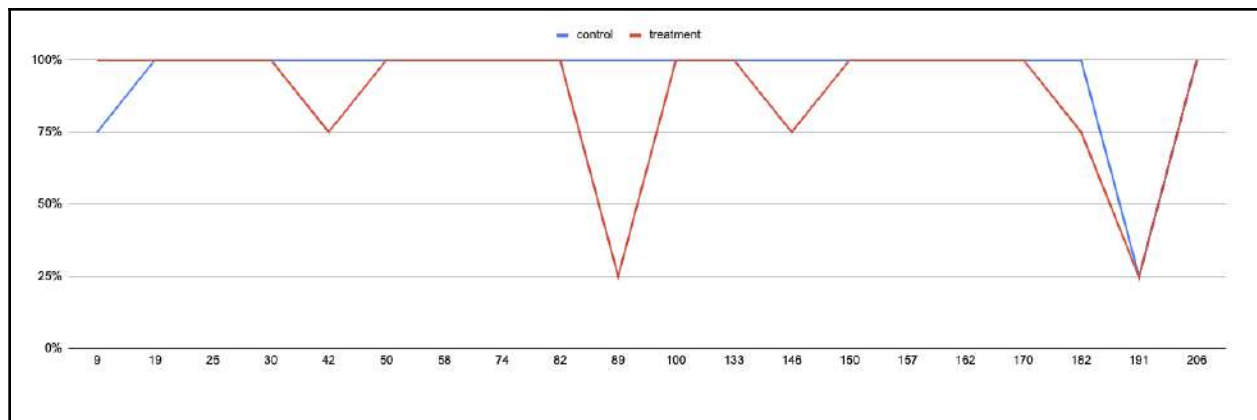
Figure 7 - Low-difficulty target tweet identification by sequence order

Medium-difficulty tweets showed relatively similar behavior between the two groups.
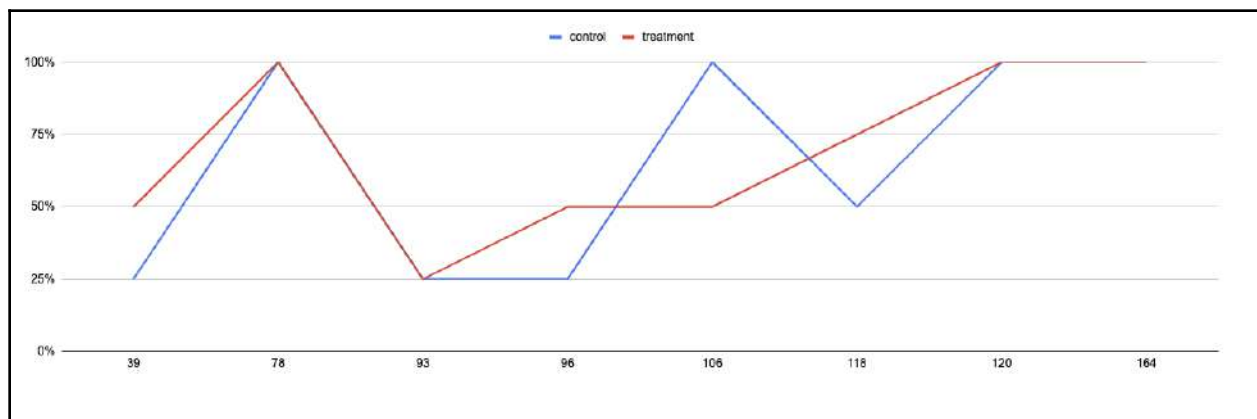
Figure 8 - Medium-difficulty target tweet identification by sequence order

The tweets with difficult to identify target content show an interesting trend. Later in the presentation sequence, the experimental treatment group more often correctly identified the target content. This is notable because content seen by the experiment group towards the end of their sequence was deeply desaturated.
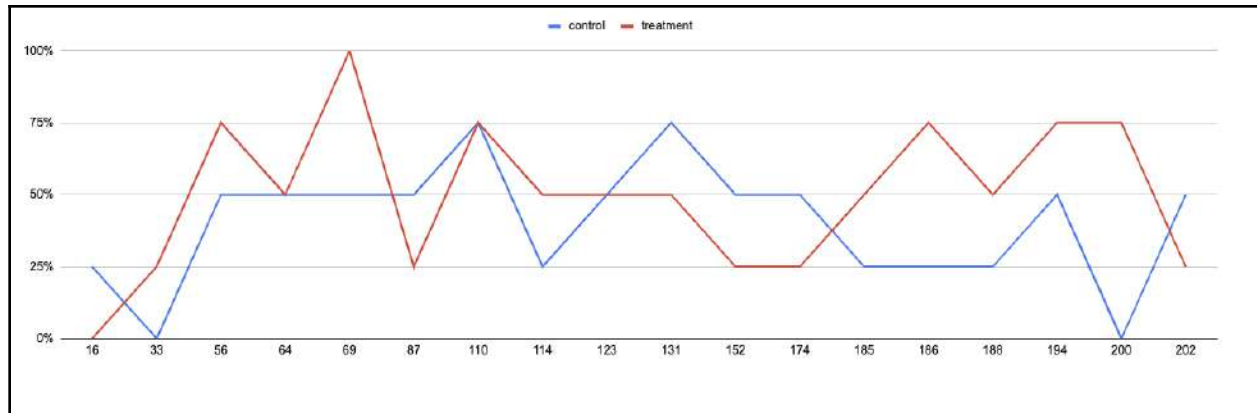
Figure 9 - High-difficulty target tweet identification by sequence order

Full details of our results are recorded in Appendix B.

Qualitative Results

We asked six questions of the participants after they performed the task with our prototypes.

Table 5 - Qualitative Questionnaire

	Question Text
Q1	*When using this design, how difficult was it to share tweets about cats on a scale of 1 to 5, with 1 being very easy and 5 being very hard?*
Q2	*Did you notice any difference in the content between this design and the previous? How did that make you feel?*
Q3	*Did you find any part of the design confusing? Which parts?*
Q4	*How often do you retweet things on Twitter?*
Q5	*For what reasons do you retweet things on Twitter?*
Q6	*Do you have any questions for me about the tasks, designs, this study, or the project?*

Participants of both control and treatment groups responded that the experimental task was not difficult (Q1), although one participant separated the mechanical retweeting process from the decision to retweet. They noted that the decision process was harder than the act of retweeting.

All participants with the experimental treatment noticed the colors desaturating (Q2). Some found it made the task of identifying images easier, while others found it harder. It seemed to increase focus on the content.

Most participants did not find the design confusing (Q3) regardless of which design they were presented with. Two reported the simulated feed's content was confusing as it did not appear realistic due to its jumble of topics.

Reported retweet behavior was noticeably different between the two groups (Q4). The experiment group was made up of fairly active Twitter users, while the control group was not. This identified a significant unintentional bias in our procedure.

When discussing the reasons why a participant would share something on social media (Q5), the content of the post was the most common consideration. This included both a consideration of the content to the participant themselves as well as the perceived value of the content the recipients of the reshare.

Participants showed interest in the purpose of the study (Q6). Participants given the control design were often curious as to if there was any actual difference.

Qualitative Insights

Discussion of the participants' experience with the prototype and with social media in general continued beyond the study questions. When discussing reasons why something would be shared on social media,

participants mentioned a concern with the curation of an online identity. This included both the desire to create a specific identity as well as the desire to avoid creating an identity at all.

> "I retweet stuff if I really really thought it was great or if it was something I wanted to associate myself with." - P7

> "Why bother to share if it's not something of your own? … I don't want to expend my energy doing that." - P4

All participants with the experimental treatment noticed the colors desaturating. Some found this made it easier to perform the task. Others found it harder. Participants mentioned the desaturation made them examine the content more closely in order to complete their task.

> "Grayscale [may] have made it easier to find cats" - P5

> "[As colors desaturated], it got harder to know what was going on in pictures" - P6

Participants who reported being familiar with Twitter commented that the prototype did not meet their expectations of functionality or content. One noted the lack of button animations and another noted the lack of proper thread behavior. Those unfamiliar with Twitter did not comment on the prototype's fidelity.

Study Limitations

The limitations of our user study fall into three categories of decreasing importance: lack of screener survey, inappropriate task context, and prototype fidelity.

Lack of screener survey

Participants were selected from a population of convenience, which is not uncommon for an early stage study of this size. However, we did not make use of a screener survey which would have identified familiarity with social media, familiarity with Twitter specifically, technology familiarity in general, and

attitudinal position towards social media. Our methodology selected participants for control or experimental groups at random, with some balancing occurring to make even numbers towards the end of recruitment. This led to an unintentional bias, unknown to our group until afterwards, that the entire group of the experimental design were familiar with Twitter, while a majority (3 of 4) of the control group were unfamiliar with Twitter. This lop-sided sampling is an obvious flaw in our experiment.

Inappropriate task context

In our study script (see Appendix A), we explicitly directed participants to perform a task using the prototype. While it was our intent to give participants a context to use the prototype, the situating of their behavior as a "task [the investigator would] like you to complete is to find and share tweets with cat(s) in it" was not appropriate (Appendix A). This prompt seemed to encourage a very analytical search through the content of the prototype which both investigators and participants identified as being "different than how [they] normally use Twitter" (Participant 6). We also suspect that this analytical search behavior artificially encouraged System 2 in both the control and experiment group, potentially obscuring the effects of our design intervention.

An improvement in script prompting would encourage a more naturalistic response from the participants. We have considered a role-play situation such as imagining one is a cat-lover who is interested in sharing tweets about cats. We would then use the participants' think-aloud responses to better understand how they conceptualized how a motivated social media user would share content.

Prototype fidelity

Our prototype was designed to test a single interaction - scrolling. The familiarity of some participants with Twitter created expectations of a richer interaction experience. The high number of suggested affordances in the design (buttons, profile icons, links, embedded content, etc) led to several participants commenting that their experience felt artificial. While some degree of artificiality is always present in a

prototype testing situation, we could have mitigated this behavior in two ways. We could have added simple placeholder interactions (button tapping animations, or a "coming soon" screen to navigate to, etc) or we could have framed the experience by setting expectations about the level of fidelity in the prototype. Either method would have helped the participants engage with our prototype in a more natural way.

Conclusion

A body of research suggests people are better at critical thinking when they slow down. The Mindfulness and Slow Design movements provide an alternative to the frictionless speed-above-all experiences promoted by *Don't Make Me Think* (Krug, 2000). Based on these insights, we sought to design an intervention to introduce constructive friction into the social media experience, in hopes of slowing the spread of mis-, mal-, and disinformation on social media platforms. Exploring the Social Media Engagement loop revealed multiple intervention opportunities, from which we selected a scrolling intervention in which color desaturates to grayscale over time. The resulting prototype demonstrates a simple content-independent design intervention that could be applied to any social media platform with minimal software development.

Our pilot study tested eight people on a find-the-cat exercise in a between-subjects experiment, in which the treatment group received the fade-to-gray intervention while the control group received the same social media feed in full-color.

Although this pilot study suggests our prototype is usable and that our basic experiment design is sound, we can not conclude that the treatment increases or decreases accuracy and time-to-complete in statistically significant manners. We suggest testing a larger population and controlling for prior social media experience and content sharing behaviors in order to improve the experimental design used here.

Additionally, a separate experiment would be required to verify if improved performance on the find-the-cat task generalizes to a reduced spread of disinformation.

While our study results were not conclusive, they suggest that the effects of modifying visual design elements of social media apps warrant further investigation.

Sources

Acquisti, A. (2009). Nudging Privacy: The Behavioral Economics of Personal Information. IEEE

    Security & Privacy Magazine, 7(6), 82–85. https://doi.org/10.1109/msp.2009.163

Barbara Grosse-Hering, Jon Mason, Dzmitry Aliakseyeu, Conny Bakker, and Pieter Desmet.

    2013. Slow design for meaningful interactions. In Proceedings of the SIGCHI Conference

    on Human Factors in Computing Systems(CHI '13). ACM, New York, NY, USA,

    3431–3440.

Bilton, R., Tameez, H., Scire, S., & Ordway, D. (2017, February 2). Reddit's /r/worldnews

    community used a series of nudges to push users to fact-check suspicious news. Nieman

    Lab. Retrieved October 23, 2021, from

    https://www.niemanlab.org/2017/02/reddits-rworldnews-community-used-a-series-of-nudge

    s-to-push-users-to-fact-check-suspicious-news/

Council of Europe, Wardle, C., Derakhshan, H., Burns, A., & Dias, N. (2017, September).

    INFORMATION DISORDER : Toward an interdisciplinary framework for research and

    policy making (DGI(2017)09). Council of Europe.

Delft Institute of Positive Design | Slow Design Book of Inspiration. (2012). Delft Institute of

    Positive Design. Retrieved October 24, 2021, from

    https://diopd.org/slow-design-book-of-inspiration/

Grosser, B. (2020). On Reading and Being Read in the Pandemic: Software, Interface, and The

    Endless Doomscroller. School of Art + Design | University of Illinois at Urbana-Champaign.

    https://elmcip.net/sites/default/files/media/critical_writing/attachments/ben_grosser_on_rea

    ding_and_being_read_in_the_pandemic_software_interface_and_the_endless_doomscroller

    _217.pdf

Kaiser, B., Mayer, J., & Nathan Matias, J. (2021, October 16). Warnings That Work: Combating

    Misinformation Without Deplatforming. Lawfare.

    https://www.lawfareblog.com/warnings-work-combating-misinformation-without-deplatfor

    ming

Köhler, B., Sabria, D., & Blumtritt, J. (2019, April 25). *The Slow Media Manifesto – Slow*

    Media. Slow Media. Retrieved October 23, 2021, from http://en.slow-media.net/manifesto

Kraut, R. E., Resnick, P., Kiesler, S., Chen, Y., Burke, M., Kittur, N., Ren, Y., Konstan, J., &

    Riedl, J. (2011). Building Successful Online Communities. Amsterdam University Press.

Krug, S. (2000). *Don't Make Me Think*. New Riders Press.

Mac, R. (2021, April 29). *Read Facebook's Internal Report About Its Role In The Capitol*

    Insurrection. BuzzFeed News. Retrieved October 23, 2021, from

    https://www.buzzfeednews.com/article/ryanmac/full-facebook-stop-the-steal-internal-report

Matias, N. J. (2017, February 1). *Persuading Algorithms With an AI Nudge*. CivilServant.

    Retrieved October 23, 2021, from

    https://civilservant.io/persuading_ais_preserving_liberties_r_worldnews.html

Mejtoft, T., Hale, S., & Söderström, U. (2019). Design Friction. Proceedings of the 31st

    European Conference on Cognitive Ergonomics. Published.

    https://doi.org/10.1145/3335082.3335106

Norman, D. (2013). The Design of Everyday Things: Revised and Expanded Edition

    (Revised ed.). Basic Books.

Mindful Staff. (2015, June 9). *Mindful Social Networking*. Mindful. Retrieved October 23, 2021,

    from https://www.mindful.org/mindful-social-networking/

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: How audio quality influences

perceptions of the researcher and research. *Science Communication*, 40(2), 246–257.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth.

Consciousness and Cognition: An International Journal, 8(3), 338–342.

https://doi.org/10.1006/ccog.1999.0386

Saindon, Jacob, "THE USE OF DISTRACTION: DOOMSCROLLING, LOSING TIME, AND

DIGITAL WELL-BEING IN PANDEMIC SPACE-TIMES" (2021). Theses and

Dissertations--Geography. 73.

https://uknowledge.uky.edu/geography_etds/73

Schell, J. (2008). The Art of Game Design: A Book of Lenses (1st ed.). CRC Press.

Schwarz, N., & Jalbert, M. (2021). When (fake) news feels true: Intuitions of truth and the

acceptance and correction of misinformation. In R. Greifeneder, M. E. Jaffé, & E. J.

Newman (Eds.), *The Psychology of Fake News: Accepting, Sharing, and Correcting*

Misinformation (1st ed., p. 75). Routledge. https://doi.org/10.4324/9780429295379

Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low

processing fluency attenuates the Moses illusion. *Social Cognition,* 26, 791–799.

Swann, S. (2021, October 6). Study finds warning labels did little to halt spread of Trump's

misinformation. The Fulcrum.

https://thefulcrum.us/big-picture/trump-twitter-misinformation

Tatiana, T. (2018, May 28). *Slow Social: The Call for a Mindful Digital Revolution*. Medium.

Retrieved October 23, 2021, from

https://medium.com/womentorship/slow-social-the-call-for-a-mindful-digital-revolution-d9b

50960eca2

Wang, Y., Leon, P. G., Acquisti, A., Cranor, L. F., Forget, A., & Sadeh, N. (2014). A field trial of

    privacy nudges for facebook. Proceedings of the SIGCHI Conference on Human Factors in

    Computing Systems. Published. https://doi.org/10.1145/2556288.2557413

Willard, C. (2018, June 14). *Before You Scroll, Try This Mindful Social Media Practice*. Mindful.

    Retrieved October 23, 2021, from

    https://www.mindful.org/before-you-scroll-try-this-social-media-practice/

Appendix A

Research Study Script

Pre-Study Setup

Meet with the subject via Zoom
Give them the link to the Figma <link>
Use Zoom record

Study Script

Introduction

Hello, I'm [YOUR NAME], and as part of a graduate class at the University of Washington, I'd like to present you with a few designs for a social media interface and get your thoughts and opinions. Your responses will not be connected with your identity and will be only used in the course of this project.

In this study, I will show you an interface and ask you to perform a task using the interface. Afterwards, I will ask you a few questions to better understand how we could improve the design.

As you complete the task, speak aloud what is going through your mind as it comes up. Sometimes people get so into the task, they forget to do so – so I might annoyingly keep reminding you to think out loud. Don't mind me!
Try to use the interface as you would normally use a social media platform. If you have any questions, feel free to ask.

There are no incorrect answers or responses. We are interested in your thoughts. Please feel free to be as frank or honest as you want. I haven't designed this, I am just interested in looking at your interaction with it. I'll be recording the call and taking notes throughout the study to review later. All the information will be anonymised, and on completion of the project, will be erased.

We will go through two tasks in this study. Before we begin, I will show you an example design so you can become more familiar with it.

Any questions before we begin?

Practice Run

[Tell subject to go to this [figma prototype link](#) and walk them through the practice run prototype]

This is an example design. The task I'd like you to complete is to **find and share a tweet about racoons.** When you find the racoon tweet, click the reshare icon:
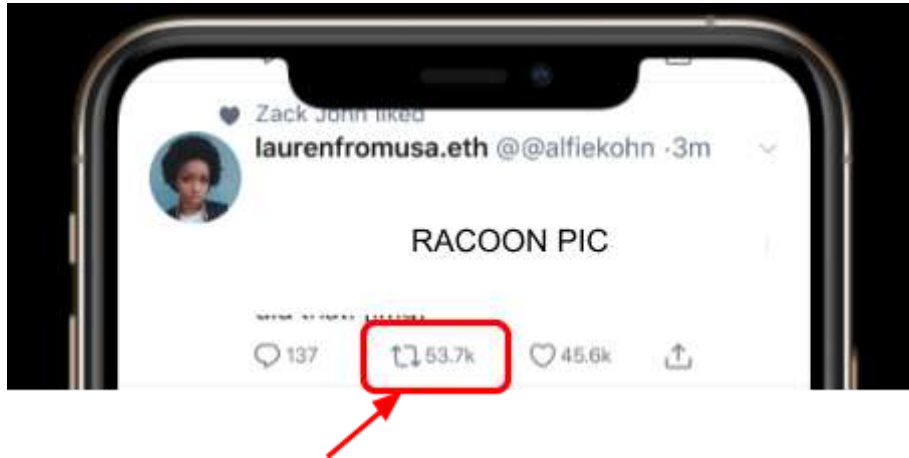
Main Task

Now we'll do the first task.

The task I'd like you to complete is to **find and share tweets with cat(s) in it**. Please speak aloud what is going through your mind as you complete the task.

[Flip a coin]
[If Heads, tell subject to go to this [figma prototype link](#)]
[If Tails, tell subject to go to this [figma prototype link](#)]

[Remind the subject to think out loud, notice their facial reactions/cursors when they go across "fox" tweets and when they go across "cat" ones]

Post-task questions

Q1 - When using this design, how difficult was it to share tweets about cats on a scale of 1 to 5, with 1 being very easy and 5 being very hard?

Q2 - Did you notice any difference in the content between this design and the previous? How did that make you feel?

Q3 - Did you find any part of the design confusing? Which parts?

Post-study questions

Q4 - How often do you retweet things on Twitter?

Q5 - For what reasons do you retweet things on Twitter?

Q6 - Do you have any questions for me about the tasks, designs, this study, or the project?

Thank you for participating in our study!

Appendix B

Study Results

Table B.1: Percentage of participants who identified a target tweet

Tweet	Difficulty	control	treatment	both		Tweet	Difficulty	control	treatment	both
9	EZ	75%	100%	88%		114	HARD	25%	50%	38%
16	HARD	25%	0%	13%		118	MEDIUM	50%	75%	63%
19	EZ	100%	100%	100%		120	MEDIUM	100%	100%	100%
25	EZ	100%	100%	100%		123	HARD	50%	50%	50%
30	EZ	100%	100%	100%		131	HARD	75%	50%	63%
33	HARD	0%	25%	13%		133	EZ	100%	100%	100%
39	MEDIUM	25%	50%	38%		146	EZ	100%	75%	88%
42	EZ	100%	75%	88%		150	EZ	100%	100%	100%
50	EZ	100%	100%	100%		152	HARD	50%	25%	38%
56	HARD	50%	75%	63%		157	EZ	100%	100%	100%
58	EZ	100%	100%	100%		162	EZ	100%	100%	100%
64	HARD	50%	50%	50%		164	MEDIUM	100%	100%	100%
69	HARD	50%	100%	75%		170	EZ	100%	100%	100%
74	EZ	100%	100%	100%		174	HARD	50%	25%	38%
78	MEDIUM	100%	100%	100%		182	EZ	100%	75%	88%
82	EZ	100%	100%	100%		185	HARD	25%	50%	38%
87	HARD	50%	25%	38%		186	HARD	25%	75%	50%
89	EZ	100%	25%	63%		188	HARD	25%	50%	38%
93	MEDIUM	25%	25%	25%		191	EZ	25%	25%	25%
96	MEDIUM	25%	50%	38%		194	HARD	50%	75%	63%
100	EZ	100%	100%	100%		200	HARD	0%	75%	38%
106	MEDIUM	100%	50%	75%		202	HARD	50%	25%	38%
110	HARD	75%	75%	75%		206	EZ	100%	100%	100%
						AVERAGE	EZ	95%	89%	92%
						AVERAGE	MEDIUM	66%	69%	67%
						AVERAGE	HARD	40%	50%	45%

Table B.2: Percentage of participants who identified an easy target tweet

Tweet	Difficulty	control	treatment	both
9	EZ	75%	100%	88%
19	EZ	100%	100%	100%
25	EZ	100%	100%	100%
30	EZ	100%	100%	100%
42	EZ	100%	75%	88%
50	EZ	100%	100%	100%
58	EZ	100%	100%	100%
74	EZ	100%	100%	100%
82	EZ	100%	100%	100%
89	EZ	100%	25%	63%
100	EZ	100%	100%	100%
133	EZ	100%	100%	100%
146	EZ	100%	75%	88%
150	EZ	100%	100%	100%
157	EZ	100%	100%	100%
162	EZ	100%	100%	100%
170	EZ	100%	100%	100%
182	EZ	100%	75%	88%
191	EZ	25%	25%	25%
206	EZ	100%	100%	100%
AVERAGE	EZ	95%	89%	92%

Table B.3: Percentage of participants who identified a medium target tweet

Tweet	Difficulty	control	treatment	both
39	MEDIUM	25%	50%	38%
78	MEDIUM	100%	100%	100%
93	MEDIUM	25%	25%	25%
96	MEDIUM	25%	50%	38%
106	MEDIUM	100%	50%	75%
118	MEDIUM	50%	75%	63%
120	MEDIUM	100%	100%	100%
164	MEDIUM	100%	100%	100%
AVERAGE	MEDIUM	66%	69%	67%

Table B.4: Percentage of participants who identified a hard target tweet

Tweet	Difficulty	control	treatment	both
16	HARD	25%	0%	13%
33	HARD	0%	25%	13%
56	HARD	50%	75%	63%
64	HARD	50%	50%	50%
69	HARD	50%	100%	75%
87	HARD	50%	25%	38%
110	HARD	75%	75%	75%
114	HARD	25%	50%	38%
123	HARD	50%	50%	50%
131	HARD	75%	50%	63%
152	HARD	50%	25%	38%
174	HARD	50%	25%	38%
185	HARD	25%	50%	38%
186	HARD	25%	75%	50%
188	HARD	25%	50%	38%
194	HARD	50%	75%	63%
200	HARD	0%	75%	38%
202	HARD	50%	25%	38%
AVERAGE	HARD	40%	50%	45%